Я. Базарсад, Р. Энхбат

Магадлалын онол математик статистик

Гарын авлага

Редактор доктор (Ph.D),
дэд профессор Я. Базарсад

Улаанбаатар–2008 он

Магадлалын онол математик статистик

Ядамын Базарсад, Рэнцэнгийн Энхбат

Номын эхийг LaTeXсистем ашиглан бэлтгэв.

Магадлалын онол, санамсаргүй хэмжигдэхүүн, математик статистикийн хичээл заадаг багш болон их дээд сургуулийн оюутан сурагчид, судлаач, сонирхогч нарт зориулав.

Магадлалын онол, математик статистикийн үндсэн ойлголтууд түүний дотор дисперс, корреляц, регрессийн шинжилгээний арга ойлголтууд, сэдэв бүрт хамаарах бодолт тайлбар бүхий практикийн жишээ бодлогуудыг багтааж, лабораторийн хичээлийн гарын авлага болохуйц программуудыг хавсаргасан.

ISBN

Агуулга

Өмнөх үг

"Магадлалын онол, математик статистик" хэмээх энэхүү номын анхны хэвлэл 1994 онд гарч олны хүртээл болсноор багш, судлаачид, оюутан сурагчдын эх хэл дээрээ ашиглах сурах бичгийн хэрэгцээг бага боловч хангаж ирсэн билээ. Номын хүрэлцээг нэмэгдүүлэх, хэвлэлтийн чанарыг сайжруулах өргөн олны хүсэлтийг харгалзан, гарын авлагын зарим бүлгийг нэмэлт материал, жишээ бодлогоор өргөтгөн, үг үсэг, найруулга, өрөлтийн алдааг шүүн тунгааж, зураг графикийг компьютерээр үйлдэн, шинэчилсэн хувилбарыг толилуулж байна.

Номын I бүлэгт комбинаторикийн үндсэн ойлголтуудыг түүврийн хэл дээр өгч, буцаалттай болон буцаалтгүй түүврийн тоог олох аргуудыг олон тооны жишээ бодлогоор тайлбарлан үзүүлсэн.

II-IV бүлэгт санамсаргүй үзэгдэл түүний магадлал, магадлал олох аргууд, нийлмэл үзэгдлийн магадлал олох томъёо, Бернуллийн томъёо түүний хязгаарын тохиолдлууд зэрэг магадлалын үндсэн арга, ойлголтуудыг холбогдох жишээ бодлогын хамт орууллаа.

V-VII бүлэгт санамсаргүй хэмжигдэхүүн түүний тархалтын хууль, тоон үзүүлэлтүүд, түгээмэл ашиглагдах дискрет болон тасралтгүй тархалтууд зэрэг санамсаргүй хэмжигдэхүүнтэй холбоотой сэдвүүд зураг, тайлбар, жишээний хамт багтсан.

VIII-IX бүлэгт математик статистикийн онолын асуудал болон практикт байнга хэрэглэгддэг тархалтын хуулиуд тэдгээрийн тархалт, тоон үзүүлэлтийг тусгайлан оруулж, их тооны хуулийн теоремуудыг хамруулав.

X-XIII бүлэгт математик статистикийн үндсэн ойлголтууд, тархалтын параметрийг үнэлэх аргууд, статистик таамаглал шалгах, дисперс, корреляци, регрессийн шинжилгээ зэрэг статистик боловсруулалт, шинжилгээний бүх аргуудыг дэлгэрэнгүй авч үзсэн.

XIV бүлэгт түүврийн статистикийн боловсруулалтын зарим тоон параметрийг олох программуудыг паскаль хэл дээр бичиж оруулсан. Эдгээр программыг C, C++, Java гэх мэт программчлалын өөр хэл дээр хөрвүүлэн ашиглах боломжтой. Төгсгөлд нь статистик боловсруулалтанд зайлшгүй шаардлагатай 9 таблицыг хавсаргасан болно.

Номын зураг, график дизайныг хийхэд туслалцаа үзүүлсэн магистр Я. Лутбатад талархал илэрхийлье. Номын талаарх санал, дүгнэлтээ ШУТИС-ийн КТМС-ийн эконометрик, үйлдлийн судалгааны багийн нэр дээр болон Bazarsad@csms.edu.mn гэсэн e-mail хаягаар ирүүлнэ үү.

Зохиогчид

⋆⋆
⋆ Энэхүү номыг ШУТИС-ийн 40 жил, ⋆
⋆ Эдийн засаг, менежментийн мэргэжилтэн бэлтгэсний 50, ⋆
⋆ Компьютерийн ухааны мэргэжилтэн бэлтгэсний 25 ⋆
⋆ жилийн ойд зориулав. ⋆
⋆⋆

Бүлэг 1

Комбинаторикийн элементүүд

Төгсгөлөг олонлогийн элементүүдийг элдэв янзаар нэгтгэх, хослуулах, байрлуулах, хослол байрлалын тоог олох асуудал практикт өргөн тохиолддог. Ийм хэлбэрийн бодлогуудыг комбинаторикийн бодлогууд гэх ба түүнийг судлах математикийн салбар ухааныг комбинаторик гэнэ.

1.1 Үржвэр, нийлбэрийн дүрэм. Түүвэр

Теорем 1. *$a_1, a_2, \ldots, a_m$ ба $b_1, b_2, \ldots, b_n$ гэсэн харгалзан m, n элементүүд бүхий олонлог тус бүрээс нэг нэг элемент оролцуулан зохиосон (a_i, b_j) $i = \overline{1, m}, j = \overline{1, n}$ хэлбэрийн хосын тоо $m \cdot n$ байна.*

Одоо энэ теоремийг арай өргөн тохиолдолд томъёолъё.

Теорем 2. *$a_1, a_2, \ldots, a_{n_1}$ гэсэн n_1 элементтэй 1-р олонлог, $b_1, b_2, \ldots, b_{n_2}$ гэсэн n_2 элементтэй 2-р олонлог, гэх мэтчилэн $c_1, c_2, \ldots, c_{n_k}$ гэсэн n_k элемент бүхий k-р олонлог тус бүрээс нэг нэг элемент оролцуулан зохиосон $(a_i, b_j, \ldots, c_l)$ хэлбэрийн хослолын нийт тоо $n_1, n_2, \ldots n_k$-тай тэнцүү байна.*

Энэхүү теорем нь комбинаторикийн онолын үндсэн теоремүүдийн нэг бөгөөд түүнийг үржвэрийн дүрэм гэж нэрлэдэг.

Жишээ-1.1 Цэцгийн үйлдвэрт 6 ягаан, 5 хөх, 4 шар, 4 цагаан өнгийн цэцгүүд байжээ. Нэг өнгийн цэцгүүд нь өөр өөр гэвэл өнгө бүрээс нэг цэцэг оролцуулсан цэцгийн баглаа хэдэн янзаар хийж болох вэ?

Ягаан өнгийн цэцгийг 6 янзаар, хөх өнгийн цэцгийг 5 янзаар, шар ба цагаан өнгийн цэцгийг тус бүр 4 янзаар сонгон авч болох учраас өмнөх теорем ёсоор $6 \cdot 5 \cdot 4 \cdot 4 = 480$ янзаар цэцгийн баглаа хийж болно.

Жишээ-1.2 Бүх цифрүүд нь ялгаатай 5 оронтой тоо хэдийг зохиож болох вэ?

5 оронтой тоо 0 гэсэн цифрээр эхлэж болохгүй тул I цифрийг 9 янзаар сонгон авч болно. Цифрүүд нь ялгаатай гэсэн учраас II цифрийг мөн 9 янзаар сонгон авч болно. Үүнчлэн III цифрийг 8 янзаар (өмнөх 2 цифрийг давтаж

авахгүй), IV цифрийг 7 янзаар, V цифрийг 6 янзаар сонгон авч болно. Тэгвэл өмнөх теорем ёсоор ялгаатай цифрүүд бүхий 5 оронтой тоо $9\cdot9\cdot8\cdot7\cdot6 = 27216$ янз байна.

Теорем 3. *Хэрэв a_1 элементийг n_1 янзаар, a_2 элементийг n_2 янзаар гэх мэтчилэн дараагийн сонголт бүр нь өмнөхөөсөө ялгаатай байхаар a_k элементийг n_k янзаар сонгох боломжтой бол $a_1, a_2, \ldots, a_k$ элементүүдийн аль нэгийг $n_1+n_2+\ldots+n_k$ янзаар сонгон авч болно.*

Энэ теоремийг нийлбэрийн дүрэм хэмээн нэрлэнэ.

Жишээ-1.3 Хүүхдийн номын дэлгүүрт байгаа зурагтай номнуудын 5 нь англи хэлээр, 3 нь орос хэлээр, 4 нь хятад хэлээр, 2 нь герман, мөн 2 нь франц хэлээр тус тус хэвлэгдсэн байв. Эдгээрээс таамгаар нэг ном авахад англи хэлээр юмуу эсвэл манай хөрш улсуудын хэлээр хэвлэгдсэн байх боломжийн тоо хэд вэ?

Нийт 16 номноос англи хэлээр хэвлэгдсэн ном сонгох боломж 5, хятад болон орос хэлээр хэвлэгдсэн ном сонгох боломж харгалзан 4 ба 3 тул дээрх теорем ёсоор $5 + 4 + 3 = 12$ янзаар сонгон авч болно.

Тодорхойлолт *$a_1, a_2, \ldots a_n$ гэсэн n элементтэй олонлогийг n хэмжээт эх олонлог гэе. Энэ олонлогоос тодорхой эрэмбээр сонгон авсан k ширхэг элемент бүхий*

$$\{a_{i_1}, a_{i_2}, \ldots, a_{i_k}\} \tag{1.1}$$

олонлогийг k хэмжээт түүвэр гэж нэрлэнэ.

Таамгаар авсан элементийг эх олонлогт буцааж хийх, эс хийхээс шалтгаалан түүвэр нь буцаалттай ба буцаалтгүй гэж 2 янз байна.

1.2 Буцаалттай түүвэр

n элемент бүхий эх олонлогоос зохиосон k хэмжээт буцаалттай түүврийн тоо n^k байна. Энэ тоог n элементтэй олонлогоос k элементээр зохиосон давталттай гүйлгэмэл гэж нэрлээд $\widetilde{A}_n^k$ гэж тэмдэглэнэ. Өөрөөр хэлбэл,

$$\widetilde{A}_n^k = n^k \tag{1.2}$$

Буцаалттай түүврүүд (давталттай гүйлгэмлүүд) хоорондоо элементүүдийн эрэмбээрээ, эсвэл элементээрээ, эсвэл хоёулангаараа ялгагдана. Буцаалттай түүвэрт элемент давтагдаж болно.

Жишээ-1.4 9 давхар байрны I давхраас 4 хүн лифтээр өгсчээ. Хэн нь ч аль ч давхарт буух боломжтой гэвэл эдгээр 9 хүн хичнээн янзаар бууж болох вэ?

Хүмүүсийн бууж болох давхрын тоо 8 учир эх олонлог 8 элементтэй, түүврийн хэмжээ 4 болно. Хэн нь ч аль ч давхарт буух боломжтой тул буцаалттай түүвэр юм. Иймд бууж болох нийт боломжийн тоо
$\widetilde{A}_8^4 = 8^4 = 4096$ болно.

1.3 Буцаалтгүй түүвэр

n элементтэй эх олонлогоос зохиосон k хэмжээт буцаалтгүй түүврийн тоо $n(n-1)\dots(n-k+1)$ байна. Энэ тоог n элементтэй олонлогоос k элементээр зохиосон гүйлгэмэл гэж нэрлээд A_n^k гэж тэмдэглэнэ. Өөрөөр хэлбэл,

$$A_n^k = n(n-1)\dots(n-k+1) = \frac{n!}{(n-k)!} \tag{1.3}$$

Нэг буцаалтгүй түүвэр нөгөөгөөсөө эсвэл элементийн эрэмбээрээ эсвэл элементээрээ ялгагдана. Буцаалтгүй түүвэрт элемент давтагдахгүй.

Жишээ-1.5 Нэг ангийн 20 оюутнаас ангийн ахлагч, оюутны зөвлөлийн гишүүн, спортын багийн ахлагч нарыг хэдэн янзаар сонгож болох вэ?

Эхлээд ангийн ахлагчийг хэн нэгээр нь сонгосон гэвэл үлдсэн 19 оюутнаас оюутны зөвлөлийн гишүүнийг сонгоно. Үүний дараа спортын багийн ахлагчийг үлдсэн 18 оюутнаас сонгоно. Иймд бидний бодлого нь 20 элементтэй олонлогоос 3-аар зохиосон буцаалтгүй түүврийн тоог олох бодлого болно: $A_{20}^3 = 20\cdot 19\cdot 18 = 6840$.

Сэлгэмэл *$n=k$ байх гүйлгэмлийг сэлгэмэл гэж нэрлээд P_n гэж тэмдэглэнэ.*

$$P_n = n(n-1)\dots(n-n+1) = n! \tag{1.4}$$

Нэг сэлгэмэл нөгөөгөөсөө зөвхөн элементийнхээ эрэмбээр ялгагдана. Иймд бүрэлдэхүүнээрээ ижил, эрэмбээрээ ялгаатай гүйлгэмлийг сэлгэмэл гэнэ.

Жишээ-1.6 Оюутан Дорж, Оюун, Бат нар зэрэгцээ гурван сандал дээр хэдэн янзаар сууж болох вэ?

Нийт боломжийн тоо нь 3 элементтэй олонлогийн сэлгэмэлийн тоо $P_3 = 3! = 6$-тай тэнцүү байна.

Хэсэглэл *n элементтэй эх олонлогоос зохиосон k хэмжээт буцаалтгүй түүврүүд дотроос ядаж нэг элементээрээ ялгагдах түүврийн тоо $\frac{A_n^k}{k!}$ болох ба энэ тоог n элементтэй олонлогооос k элементээр зохиосон хэсэглэл гэж нэрлээд C_n^k гэж тэмдэглэнэ.*

$$C_n^k = \frac{A_n^k}{k!} = \frac{n!}{k!(n-k)!} \tag{1.5}$$

n-ээс k-аар авсан хэсэглэл нь n хэмжээт эх олонлогийн k элементтэй дэд олонлогуудын нийт тоо юм. $C_n^0, C_n^1, \dots, C_n^{n-1}$ тоонууд нь $(a+b)^n$ гэсэн хоёр гишүүнтийн Ньютоны биномын коэффициентүүд бөгөөд

1. $C_n^k = C_n^{n-k}$,
2. $C_n^0 + C_n^1 + \dots + C_n^n = 2^n$,
3. $1 \le k \le n$ байх дурын k тооны хувьд $C_n^k = C_{n-1}^k + C_{n-1}^{k-1}$ чанаруудыг

хангана. $C_n^0 = 1, \quad 0! = 1$ гэж тооцно.

Жишээ-1.7 Ангийн 30 оюутны 12 нь эмэгтэй бол 3 эмэгтэй 5 эрэгтэй оролцсон 8 хүнтэй спортын багийг хэдэн янзаар байгуулж болох вэ?

12 эмэгтэйгээс 3 эмэгтэй сонгон авах боломжийн тоо C_{12}^3, 18 эрэгтэйгээс 5 эрэгтэй сонгон авах боломжийн тоо C_{18}^5 байна. Үржвэрийн дүрэм ёсоор, 8 хүнтэй баг байгуулах боломжийн нийт тоо $n = C_{12}^3 \cdot C_{15}^8 = 188496$.

Давталттай хэсэглэл $a_1, a_2, \ldots, a_n$*—эх олонлогийн зарим элементүүд (эсвэл бүгд) давтагдах тохиолдолд түүнээс зохиосон хэсэглэлийг давталттай гэх бөгөөд* $\widetilde{C}_n^k$ *гэж тэмдэглэнэ:*

$$\widetilde{C}_n^k = C_{n+k-1}^k \tag{1.6}$$

Жишээ-1.8 Цагаан толгойн үсгийн сангийн 21 картын 5 нь A үсэгтэй, 6 нь 5-ын цифртэй, 10 нь $\triangle$ дүрстэй байв. Эдгээрээс 4 картыг хичнээн ялгаатай аргаар сонгон авч болох вэ?

Тус бүр нь ялгаатай давталт бүхий 3 өөр төрлийн картаас, эрэмбэ нь үл хамаарах 4 карт авч буй тул давталттай хэсэглэл болно.
Иймд, $\widetilde{C}_3^4 = C_{4+3-1}^4 = C_6^4 = 15$ ялгаатай аргаар сонгон авч болно.

1.4 Өгөгдсөн бүрэлдэхүүний гүйлгэмэл (Давталттай сэлгэмэл)

$X = \{a_1, a_2, \ldots, a_n\}$ *олонлогийн элементүүдээс давталттайгаар сонгон авсан* k *элементийг* ($k \geq n$ *байж болно!*) k *урттай мөр гэе.* k *урттай мөрд* a_i *элементийн давтагдсан тоог* k_i *гэж тэмдэглээд* k_i*-г эрэмбээр нь бичвэл* $(k_1, k_2, \ldots, k_n)$ *гэсэн шинэ мөр үүснэ. Энэ мөрийг өгөгдсөн* k *урттай мөрийн бүрэлдэхүүн гэнэ.* $(k_1 + k_2 + \cdots + k_n = k)$.
Жишээлбэл, $X = \{a_1, a_2, a_3, a_4\}$ бол $\alpha = (a_1, a_3, a_1, a_1, a_2, a_4, a_3)$ нь 7 урттай мөр болно. α мөрийн бүрэлдэхүүн нь (3,1,2,1) болно.
Өгөгдсөн $(k_1, k_2, \ldots, k_n)$ *бүрэлдэхүүний гүйлгэмлийн тоог* $A(k_1, k_2, \ldots, k_n)$ *гэж тэмдэглээд давталттай сэлгэмэл гэнэ:*

$$A(k_1, k_2, \ldots, k_n) = \frac{(k_1 + k_2 + \cdots + k_n)!}{k_1! k_2! \ldots k_n!} \tag{1.7}$$

(1.7) томъёо нь бүрэлдэхүүнээрээ ижил боловч элементүүд нь давтагдах, зөвхөн элементийнхээ эрэмбээр ялгагдах гүйлгэмлийн тоо юм.
Ньютоны биномын өргөтгөл болох томъёог тэмдэглэе.

$$(a_1 + a_2 + \cdots + a_n)^k = \sum \frac{k!}{k_1! k_2! \ldots k_n!} a_1^{k_1} a_2^{k_2} \ldots\ldots a_n^{k_n} \tag{1.8}$$

Үүнд, томъёоны нийлбэр нь $k_1 + k_2 + \cdots + k_n = k$ нөхцөлийг хангах бүх $k_1, k_2, \ldots, k_n$-ээр авагдана. (1.8) томъёог полинимал томъёо гэнэ.

Полинимал томъёоны гишүүдийн коэффициентүүд нь $(k_1, k_2, \ldots, k_n)$ бүрэлдэхүүний гүйлгэмлийн тоо юм.

$k = 2$ үед (1.8) томъёоноос Ньютоны биномын томьёо мөрдөнө.

Жишээ-1.9 Оюутны номын санд профессор Хироногийн лекцийн эмхэтгэл 3ш , зах зээлийн эдийн засаг, менежментийн сурах бичиг гэсэн номууд тус бүр 2 байжээ. Эдгээр номуудыг тавиур дээр хэдэн ялгаатай янзаар байрлуулж болох вэ?

Тавиур дээр байрлуулах нийт 7 номын бүрэлдэхүүн нь (3,2,2) болно. Иймд эдгээр бүрэлдэхүүний тоо (1.7) томъёо ёсоор

$$A(3,2,2) = \frac{(3+2+2)!}{3!2!2!} = 210.$$

Жишээ-1.10 "ОЛОНЛОГ" гэсэн үгийн үсгүүдийн байрыг солих замаар, үсгүүдийн дарааллыг хичнээн өөрөөр үүсгэж болох вэ?

Энэ бодлого нь О, Л, Н, Г гэсэн 4 өөр үсгээр, 7 үсэгтэй үсгүүдийн дараалал хэдийг үүсгэж болох вэ? гэсэн бодлоготой эквивалент юм. Зохиох үгийн бүрэлдэхүүн нь (3,2,1,1) учир нийт тоо нь

$$A(3,2,1,1) = \frac{(3+2+1+1)!}{3!2!1!1!} = 420.$$

Бүлэг 2

Үзэгдэл, түүний магадлал

2.1 Санамсаргүй үзэгдэл ба магадлалын онол

Байгаль ертөнц хүний нийгэмд явагдаж байгаа аливаа үзэгдлийг ямар нэг туршилтын үр дүн гэж үзэж болно. Эдгээр туршилтыг байгаль өөрөө хийхээс гадна шинжлэх ухааны судалгаа, техник эдийн засаг, үйлдвэрлэлийн процесст хүн өөрөө хийж болно. Туршилт давтагдахад үзэгдэл дахин давтагдана.

Ийнхүү, магадлалын онолд аливаа туршилтын үр дүнг үзэгдэл гэж нэрлэнэ. Тухайн туршилтанд явагдаж ч болох, явагдахгүй ч байж болох үзэгдлийг санамсаргүй үзэгдэл гэнэ. Өөрөөр, хэлбэл явагдах ба эс явагдахыг нь урьдчилан хэлэх боломжгүй үзэгдлийг санамсаргүй гэнэ.

Давтагдах, санамсаргүй үзэгдлийг судлаж үзвэл өөрийн дотоод нарийн зүй тогтолтой байдаг байна. Ямарваа туршилтыг ижил нөхцөлд олон дахин давтан хийхэд нэг удаагийн туршилтын үр дүнг урьдчилан хэлж чадахгүй ч гэсэн олон удаагийн туршилтын дүнг яг хэлж болдог байна.

Туршилтыг n удаа давтахад бидний сонирхсон үзэгдэл m удаа явагдсан (илэрсэн) гэвэл $\dfrac{m}{n}$ харьцааг уг үзэгдлийн илрэх харьцангуй давтамж гэнэ. Туршилтын тоог хүрэлцээтэй олшруулахад үзэгдэл явагдах харьцангуй давтамж $\dfrac{m}{n}$ нь тодорхой нэг тооны орчинд хэлбэлздэг болохыг олж тогтоосон байна. Тухайлбал, зоос хаях туршилтанд сүлдээрээ буух үзэгдлийг сонирхъё. Туршилтыг олон дахин давтан хийхэд сүлдээрээ буух үзэгдлийн харьцангүй давтамж $\dfrac{1}{2}$ гэсэн тооны орчинд тогтворждог байна. Үүнийг францын байгаль судлаач Бюффен, английн математикч К.Пирсон нар практик дээр шалгажээ. Шалгасан үр дүнг дараах хүснэгтээр харуулав.

	Зоос хаясан тоо	Сүлд буусан тоо	Сүлд буух үзэгдлийн харьцангуй давтамж
Бюффен	4040	2048	0.50693
К.Пирсон	12000	6019	0.5016
К.Пирсон	24000	12012	0.5005

Санамсаргүй үзэгдлийн давтамж нь тогтворждог энэ чухал чанарыг магадлалын онол судална. Өөрөөр хэлбэл, магадлалын онол нь санамсаргүй үзэгдлүүдийн өвөрмөц шинж чанар, ерөнхий зүй тогтлыг судлаж тэдгээрт элдэв тоон үнэлгээнүүд тогтоож өгнө. Магадлалын онол, математик статистикийн аргууд нь физик техникийн шинжлэх ухаан, тооцон бодох техник, астрономи, автомат удирдлагын онол, цаг уур судлал, цэрэгжилтийн ухаан, эдийн засаг, хэл шинжлэл, нийгэм судлал, биологи гэх мэт орчин үеийн мэдээллийн бүх салбарт нэвтэрсний гадна мэдээллийн онол, нийтийн үйлчилгээний онол, найдварын онол, тоглоомын онол, нөөцийн онол, сүйрлийн онол зэрэг шинэ шинэ салбарууд үүсэн хөгжиж байна.

2.2 Санамсаргүй үзэгдэл. Эгэл үзэгдлийн огторгуй

T-ямар нэг туршилт гэж үзээд энэ туршилтын боломжит бүх дүнгүүдийг

$$\omega_1, \omega_2, \dots, \omega_i, \dots \tag{2.1}$$

гэж тэмдэглэе. ω_i-тус бүрийг эгэл үзэгдэл гэж нэрлэнэ.

Бүх эгэл үзэгдлийн олонлогийг тухайн туршилтын эгэл үзэгдлийн огторгуй гэнэ. Эгэл үзэгдлийн огторгуйг Ω гэж тэмдэглэвэл:

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_i, \dots\} \tag{2.2}$$

болно. Тухайн туршилтын эгэл үзэгдлийн огторгуй төгсгөлөг, тоологдом, төгсгөлгүй элементтэй (чадалтай) байж болно. Тухайлбал, шоог нэг удаа орхих туршилтанд $\Omega = \{1, 2, 3, 4, 5, 6\}$, зоосыг нэг удаа хаях туршилтанд Ω=\{с,т\}. Үүнд: с-сүлд, т-тоо. Зоосыг хоёр удаа хаях туршилтанд Ω =\{сс, тт, тс, ст\} гэсэн төгсгөлөг элементүүдтэй байхад, зоосыг анх удаа сүлдээрээ буутал хаях туршилтанд Ω нь төгсгөлөг буюу тоологдом байж болно.

Ω =\{с, тс, ттс, тттс, …\}

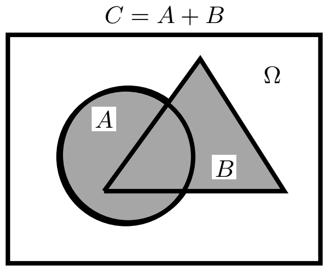
Эгэл үзэгдлийн огторгуй Ω-ийн дурын дэд олонлог E бүрийг ($E \subset \Omega$) үзэгдэл гэнэ. Цаашид санамсаргүй үзэгдлийг $A, B, C \dots$ гэх мэтчилэн латин цагаан толгойн том үсгүүдээр тэмдэглэх болно.

Хэрэв туршилтанд $\omega \in E$ эгэл үзэгдлийн ядаж нэг нь илэрвэл (явагдвал) E үзэгдлийг явагдсан гэнэ. Тухайлбал, шоо орхих туршилтанд: i-ээр i нүд буух үзэгдлийг тэмдэглэвэл ($i = \overline{1,6}$). $\Omega = \{i \mid i = \overline{1,6}\} = \{1,2,3,4,5,6\}$ болно. Ω-ийн дараах дэд олонлогуудыг авч үзье:
$\Omega \supset A = \{2,4,6\}$, $\Omega \supset B = \{3,6\}, \Omega \supset C = \{2,3,5\}$. Тэгвэл A нь шоон дээр тэгш тоо буух үзэгдэл, B нь шоон дээр 3-д хуваагдах тоо буух үзэгдэл, C нь шоо анхны тоогоор тусах үзэгдлүүд болох жишээтэй.

2.3 Санамсаргүй үзэгдэл дээр хийх үйлдлүүд

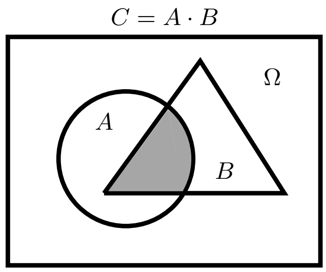
A, B нь $A \subset \Omega$, $B \subset \Omega$ байх дурын үзэгдлүүд байг.

1. A, ба B үзэгдлийн ядаж нэгд нь харъяалагдах эгэл үзэгдлүүд бүхий C-үзэгдлийг тэдгээрийн нийлбэр гээд $C = A + B$ буюу $C = A \cup B$ гэж тэмдэглэнэ. Өөрөөр хэлбэл, A ба B үзэгдлүүдийн ядаж нэг нь явагдах үзэгдлийг тэдгээрийн нийлбэр гэнэ. A, B үзэгдлийн эгэл үзэгдлүүдийг хавтгайн цэгүүдээр дүрсэлбэл, зураг-2.1-д C үзэгдлийн эгэл үзэгдлүүдийг зурааслан тэмдэглэв.



Зураг 2.1:

2. A, B үзэгдлүүдэд нэгэн зэрэг харъяалагдах эгэл үзэгдлүүд бүхий C үзэгдлийг тэдгээрийн үржвэр гээд $C = A \cdot B$ буюу $C = A \cap B$ гэж тэмдэглэнэ. Өөрөөр хэлбэл, A, B үзэгдлүүд нэгэн зэрэг явагдах үзэгдлийг тэдгээрийн үржвэр гэнэ. Үзэгдлүүдийн нийлбэр ба үржвэр гэсэн ойлголтыг төгсгөлөг ба тоологдом тооны үзэгдлүүд дээр дэлгэрүүлж болно. (Зураг 2.2)

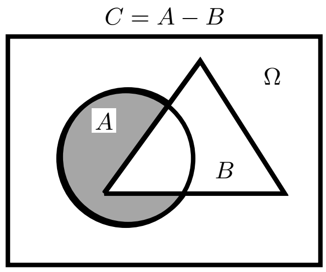


Зураг 2.2:

3. B-д эс харьяалагдах A-ийн эгэл үзэгдлүүдээс тогтох үзэгдлийг A-аас B-г хассан ялгавар үзэгдэл гэж нэрлээд $A - B$ эсвэл $A \backslash B$ гэж тэмдэглэнэ. Өөрөөр хэлбэл, A явагдаж B эс явагдах үзэгдлийг $A - B$ үзэгдэл гэнэ.(Зураг 2.3)

4. Ω-ийг зайлшгүй (гарцаагүй) үзэгдэл гэнэ. Өөрөөр хэлбэл, тухайн туршилтанд заавал явагдах үзэгдлийг зайлшгүй үзэгдэл гэнэ.

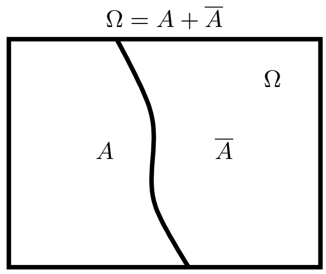
5. Тухайн туршилтанд огт явагдахгүй (илрэхгүй) үзэгдлийг боломжгүй үзэгдэл гэж нэрлээд $\varnothing$ гэж тэмдэглэнэ. ($\overline{\Omega} = \varnothing$)



Зураг 2.3:

6. $A + B = \Omega$, $A \cdot B = \varnothing$ нөхцөлийг хангах B үзэгдлийг A-ийн эсрэг үзэгдэл гэж нэрлээд $B = \overline{A}$ гэж тэмдэглэнэ. Өөрөөр хэлбэл, зөвхөн 2 боломжтой бөгөөд туршилтанд аль нэг нь заавал явагдах үзэгдлийг харилцан эсрэг үзэгдлүүд гэнэ. Харилцан эсрэг үзэгдлүүдийн хувьд $\overline{A} = \Omega - A$ байна. (Зураг 2.4)

7. $A{\cdot}B = \varnothing$ нөхцөлийг хангах A ба B үзэгдлүүдийг нийцгүй гэнэ. Өөрөөр хэлбэл, тухайн туршилтанд зэрэг илрэхгүй үзэгдлүүдийг нийцгүй гэнэ. Эсрэг тохиолдолд нийцтэй гэнэ.



Зураг 2.4:

8. A үзэгдэл явагдмагц B үзэгдэл явагдаж байвал A нь B-г дагуулж байна гээд $A \subset B$ гэж тэмдэглэнэ. A-г B-ийн тухайн тохиолдол ч гэж нэрлэнэ. Энэ нь $\forall\omega_i \in A$ бүрийн хувьд $\omega_i \in B$ гэсэн үг юм.

9. Туршилтын дүнд, тухайн үзэгдлүүдийн зөвхөн нэг нь заавал явагддаг бөгөөд хос хосоороо нийцгүй байвал тэдгээрийг үзэгдлүүдийн гүйцэд (бүтэн) бүлгийг үүсгэж байна гэж ярьдаг.

Жишээ-2.1 Тамирчин байг 2 удаа буудах туршилт хийсэн гэе. $A-$эхний буудалтанд байг онох үзэгдэл, $B-$хоёр дахь буудалтанд байг онох үзэгдэл бол $C = A + B$ нь байг онох үзэгдэл болно. (эхний удаад, эсвэл хоёр дахь удаад, эсвэл хоёуланд нь)

Жишээ-2.2 Шоо орхих туршилтанд $A-$тэгш тоогоор буух үзэгдэл, $B-$3-т хуваагдах тоо буух үзэгдэл бол $C = A \cdot B$ нь 6-д хуваагдах тоогоор буух үзэгдэл болно.

Жишээ-2.3 Шоо орхих туршилтанд "шоо аль нэг нүдээрээ буух үзэгдэл" зайлшгүй үзэгдэл байна .

Жишээ-2.4 "Таамгаар сонгон авсан 2 оронтой тооны цифрүүдийн нийлбэр 20 байх" үзэгдэл боломжгүй үзэгдэл юм.

Жишээ-2.5 Байг нэг удаа буудах туршилтанд "байг онох", "үл онох" үзэгдэлүүд; "зоос хаях" туршилтанд "сүлдээрээ буух", "тоогоороо буух" үзэгдлүүд нь тус тус эсрэг үзэгдлүүд болно.

Жишээ-2.6 Өдөр шөнө солигдох байгалийн туршилтанд "нар мандах", "нар жаргах" үзэгдлүүд; бай буудах туршилтанд "байг онох", "үл онох" үзэглүүд; зоос хаях туршилтанд "сүлд буух", "тоо буух" үзэгдлүүд тус тус нийцгүй үзэгдлүүд юм.

Жишээ-2.7 Шоо орхих туршилтанд $A-$шоо 3 нүдээрээ буух үзэгдэл, $B-$3-д хуваагдах тоогоор тусах үзэгдэл бол A явагдмагц B явагдаж буй тул A нь $B-$г дагуулж байна. Өөрөөр хэлбэл A үзэгдэл $B-$ийн тухайн тохиолдол болно.($A \subset B$)

Жишээ-2.8 Байг 3 удаа буудах туршилтанд "нэг онох", "хоёр онох", гурав онох", "үл онох" үзэгдлүүд нь үзэгдлүүдийн гүйцэд (бүтэн) бүлгийг үүсгэнэ.

2.4 Санамсаргүй үзэгдлийн магадлал. Магадлалын чанар

Геометрт хэрчим бүрт түүний "урт" хэмээх сөрөг бус тоо, физикт бие бүрт "масс" хэмээх сөрөг бус тоо харгалзуулдгийн адил магадлалын онолд санамсаргүй үзэгдэл бүрд түүний явагдах боломжийн хэр хэмжээг харуулсан сөрөг биш тоо харгалзуулдаг. Энэ тоогоо уг үзэгдлийн магадлал гэж нэрлэнэ.

Хэрэв эгэл үзэгдлийн огторгуй $\Omega = \{\omega_i \mid i = 1, 2, \dots\}$ бол, эгэл үзэгдэл ω_i-бүрд түүний магадлал хэмээх, $P(\omega_i) = p_i$ гэж тэмдэглэх, сөрөг биш тоог дараах нөхцөл биелэгдэж байхаар онооё.

$$\sum_{\omega_i \in \Omega} P(\omega_i) = \sum_i p_i = 1 \tag{2.3}$$

$A = \{\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_k}, \dots\}$ нь $A \subseteq \Omega$ байх дурын үзэгдэл байг. A-д дараах магадлалыг онооё.

$$P(A) = \sum_{\omega_i \in A} P(\omega_i) = P(\omega_{i_1}) + P(\omega_{i_2}) + \dots + P(\omega_{i_k}) + \dots \tag{2.4}$$

(2.4)-ийн баруун тал нийлэх тоон цуваа учраас ийм тодорхойлолт бүрэн үндэстэй юм. Үзэгдлийн магадлалын чанаруудаас дурдъя.

1. Дурын A үзэгдлийн хувьд $0 \le P(A) \le 1$.

2. $P(\Omega) = 1$. Өөрөөр хэлбэл, зайлшгүй үзэгдлийн магадлал 1.

3. $P(\varnothing) = 0$. Боломжгүй үзэгдлийн магадлал 0.

4. A, B дурын үзэгдлүүд бол:

$$P(A + B) = P(A) + P(B) - P(A \cdot B) \tag{2.5}$$

Хэрэв A, B үзэгдлүүд нийцгүй бол $(A \cdot B = \varnothing)$

$$P(A + B) = P(A) + P(B) \tag{2.6}$$

(2.5) ба (2.6) томьёог 2-оос олон үзэгдлүүдийн хувьд дэлгэрүүлж болно.

5. A ба $\overline{A}$ харилцан эсрэг үзэгдлүүд бол
$P(A) + P(\overline{A}) = 1$ буюу $P(\overline{A}) = 1 - P(A)$.

6. Хэрэв $A \subset B$ бол
$P(A) \le P(B)$ байхын зэрэгцээ $P(B - A) = P(B) - P(A)$.

2.5 Магадлалын сонгодог тодорхойлолт

Эгэл үзэгдлийн огторгуй Ω-д дараах хоёр шаардлага тавья:

1). Ω-төгсгөлөг олонлог. Өөрөөр хэлбэл

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_n\} \qquad (|\Omega| = n)$$

2). ω_i эгэл үзэгдлүүд бүгдээрээ ижил боломжтой илэрдэг байг. Одоо Ω-ийн эгэл үзэгдэл бүрд, өмнөх зүйлийн (2.3) нөхцөл биелэгдэж байхаар нэгэн ижил $P(\omega_i) = p$ магадлал онооё.

$$\sum_{i=1}^{n} P(\omega_i) = \sum_{i=1}^{n} p = n \cdot p = 1 \Rightarrow p = \frac{1}{n}.$$

Ийнхүү эгэл үзэгдэл бүрд $p = \dfrac{1}{n}$ магадлал оноолоо.
$A \subset \Omega$, $A = \{\omega_{i1}, \omega_{i_2}, \ldots, \omega_{i_k}\}$-үзэгдлийг авч үзье. Өөрөөр хэлбэл, Ω-ийн эгэл үзэгдлүүдийн дотроос бидний сонирхсон шинж чанар бүхий эгэл үзэгдэл яг k ширхэг байг. (2.4) томъёо ёсоор:

$$P(A) = \sum_{\omega_i \in \Omega} p(\omega_i) = \sum_{i=1}^{k} p = k \cdot p = k \cdot \frac{1}{n} = \frac{k}{n}$$

$$P(A) = \frac{k}{n} \tag{2.7}$$

(2.7) томьёог магадлалын сонгодог тодорхойлолт гэнэ.
$\omega_{i_1}, \omega_{i_2}, \ldots, \omega_{i_k}$-г A үзэгдлийн ивээлт цэгүүд гэж нэрлэж заншжээ. Өөрөөр хэлбэл, бидний сонирхсон шинж тэмдэг бүхий эгэл цэгүүдийг уг үзэгдлийн ивээлт цэгүүд гэнэ. Цаашид Ω-ийн эгэл үзэгдлүүдийг товчоор эгэл цэгүүд гэж нэрлэе.

Ийнхүү, магадлалын сонгодог тодорхойлолт ёсоор A үзэгдлийн магадлал нь түүний ивээлт цэгүүдийн тоог бүх эгэл цэгүүдийн тоонд харьцуулсан харьцаатай тэнцүү байна.

Одоо магадлалын сонгодог тодорхойлолт ашиглан бодох зарим хялбар жишээ авч үзье.

Жишээ-2.9 Зоос хаях туршилтанд сүлдээрээ буух үзэгдлийг A гэвэл, A үзэгдлийн магадлалыг ол.

Эгэл үзэгдлийн огторгуй Ω=$\{$с, т$\}$ гэсэн 2 элементтэй. Зоосны 2 тал тэгш хэмтэй учраас сүлд буух, тоо буух үзэгдлүүд ижил боломжтой. Иймд нийт эгэл үзэгдлийн тоо $n = 2$, ивээлт цэгийн тоо $k = 1$ тул (2.7) томъёо ёсоор $P(A) = \dfrac{1}{2}$; Үүнчлэн зоос тоогоороо буух үзэгдлийн магадлал мөн $\dfrac{1}{2}$ байна.

Жишээ-2.10 Зоосыг 3 удаа хаях туршилтанд дэс дараалан сүлдээрээ буух үзэгдлийн магадлалыг ол.

Сонирхсон үзэгдлийг B гэе. Эхлээд эгэл үзэгдлийн огторгуйн элементийн тоог (эгэл цэгүүдийн тоог) олъё. $\{$с, т$\}$ гэсэн $n = 2$ элементтэй олонлогоос $k = 3$ хэмжээт буцаалттай түүвэр зохионо гэсэн үг учраас $n = \widetilde{A}_2^3 = 2^3 = 8$. Дэлгэрэнгүй бичвэл: Ω =$\{$тст, ттт, ттс, тсс, ссс, стс, сст, стт$\}$. Эгэл цэг бүр нь ижил боломжтой, ивээлт цэгийн тоо $k = 1$ тул

$P(B) = \frac{1}{8}$.

Жишээ-2.11 Шоо орхих туршилтанд "5"- нүд буух үзэгдлийн магадлалыг ол.

Шоо "5" нүдээрээ буух үзэгдлийг C гэе. Шоо нэгэн төрлийн, тэгш хэмтэй учраас эгэл цэг бүр нь ижил боломжтой. $n = 6, \quad k = 1$ учраас $P(C) = \frac{1}{6}$.

Дээрх 3 жишээнд авч үзсэн A, B, C үзэгдлүүдийн магадлалыг харьцуулан үзвэл A үзэгдэл нь B ба C үзэгдлийг бодвол явагдах боломжоор илүү, C үзэгдэл B-г бодвол илүү явагдах боломжтой гэж харагдаж байна.

Жишээ-2.12 Театрын үзвэрийн билет худалдан авахад Болд Оюун нарын суудал ангийнх нь 3 оюутны хамт нэг эгнээний дэс дараалсан суудлуудад таарчээ. Эдгээр суудалд Болд Оюун хоёр зэрэгцэж суух үзэгдлийн магадлалыг ол.

Сонирхсон үзэгдлээ D гэж тэмдэглэе. Эхлээд эгэл үзэгдлүүдийн тоо n-ийг олъё. 5 оюутан 5 суудалд ямар нэгэн байдлаар суух байрлалыг нэг эгэл үзэгдэл гэж тооцох бөгөөд хэн нь аль ч суудалд сууж болох тул эдгээр эгэл үзэгдлүүд ижил боломжтой. 5 оюутан 5 суудалд сууж болох нийт боломжийн тоо $n = P_5 = 5!$

Одоо ивээлт цэгүүдийн тоог олъё. Дээрх байрлалуудын дотроос зөвхөн Болд Оюун хоёр нэг дор суусан байрлал нь ивээлт үзэгдэл болно. Эгнээ 5 суудалд зэрэгцээ суудлын тоо 4, нэг зэрэгцээ суудалд Болд Оюун нар суухад бусад 3 нь $P_3 = 3!$ янзаар сууж болох бөгөөд нэг зэрэгцээ суудалд Болд Оюун нар 2 янзаар сууж болно. Эдгээрийг анхааран үржвэрийн дүрмийг хэрэглэвэл $k = 2 \cdot 4 \cdot 3!$. Иймд $P(D) = \frac{2 \cdot 4 \cdot 3!}{5!} = \frac{2}{5}$

Жишээ-2.13 8 эрэгтэй 6 эмэгтэй сурагчидтай ангиас 4 хүнтэй бригад зохион байгуулахад 1 бригадад 2 эрэгтэй 2 эмэгтэй оролцсон байх үзэгдлийн магадлалыг ол.

Нийт 14 сурагчдаас 4 хүнтэй бригад зохион байгуулах боломжийн тоо $n = C_{14}^4$. Харин 8 эрэгтэйгээс 2-ыг сонгон авах боломж C_8^2, 6 эмэгтэйгээс 2-ыг сонгох боломж C_6^2 тул үржвэрийн дүрэм ёсоор 2 эрэгтэй 2 эмэгтэй оролцсон байх 4 хүнтэй бригадын тоо $k = C_8^2 \cdot C_6^2$ болно. Сонирхсон үзэгдлээ E гэвэл $P(E) = \frac{C_8^2 \cdot C_6^2}{C_{14}^4} = \frac{60}{143}$.

(2.7) томьёог гаргахад тавьсан "ω_i-бүр ижил боломжтой" гэсэн шаардлага асуудлыг ихэд явцууруулж буй хэрэг юм. Нөгөө талаас, практик дээр эгэл үзэгдлийн огторгуй Ω-төгслөг байх нь албагүй бөгөөд эгэл үзэгдэл бүр нь ижил боломжтой эсэхийг тогтоох аргагүй үзэгдлүүд бишгүй олон тааралдана. Өөрөөр хэлбэл, аливаа үзэгдлийн магадлалыг ямагт сонгодог тодорхойлолтоор бодох боломжгүй. Иймд учраас өөр хэлбэрийн магадлал тухайлбал, магадлалын статистик тодорхойлолт, геометр магадлалыг авч үзэх шаардлага гардаг.

2.6 Магадлалын статистик тодорхойлолт

Бид, ижил нөхцөлд аливаа туршилтыг дахин давтан хийхэд үзэгдэл явагдах харьцангуй давтамж тогтворждог тухай 2.1 дэд бүлэгт дурдсан билээ.

Ямар нэг туршилтанд A үзэгдэл явагддаг байг.
Туршилтыг n удаа хийхэд A үзэгдэл m удаа явагдсан гэвэл $\dfrac{m}{n}$ нь A үзэгдэл явагдах харьцангуй давтамж болно. n удаагийн туршилтыг дахин давтан хийхэд A үзэгдлийн харьцангуй давтамж өмнөхөөс бага зэрэг ялгаатай гарч болно. Гэх мэтчилэн олон удаагийн туршилтанд $\dfrac{m}{n}$ давтамж тодорхой нэг сөрөг биш тооны орчимд хэлбэлзэнэ. Энэхүү сөрөг биш тоог A үзэгдлийн магадлал болгон сонгон авч, магадлалын статистик тодорхойлолт гэнэ. Энэ тодорхойлолтыг математик томьёоллоор илэрхийлбэл:

$\lim\limits_{n\to\infty} \dfrac{m}{n} = p$. Энэ томъёололд, $\dfrac{m}{n}$ харьцаа $n \to \infty$ үед p тоо руу нийлэхдээ ердийн хязгаарын утгаар биш магадлалаараа нийлэх нийлэлтийн тухай ярьж буй юм.
Магадлалын статистик тодорхойлолт нь бүхэлдээ, магадлалын онолын арга болохынхоо хувьд дурын санамсаргүй үзэгдлийн хувьд хэрэглэгдэхгүй.
Харин тодорхой нөхцлийн үед тухайлбал, дараах чанаруудыг хангасан санамсаргүй үзэгдлийн хувьд хэрэглэх боломжтой:
1.) Авч үзэж буй үзэгдэл нь нэгэн ижил иж бүрдэл нөхцлийн үед хүрэлцээтэй олон удаа хийсэн туршилтын үр дүнгүүд байх
2.) Үзэгдэл нь " статистик тогтвортой" байх: Олон удаагийн туршилт хийхийн өмнө харьцангуй давтамж m/n нь урьдчилан хэлэх боломжгүй санамсаргүй тоо байна. Гэвч, санамсаргүй үзэгдлийн ерөнхий зүй тогтол ёсоор туршилтын тоог олшруулах тутам харьцангуй давтамжийн санамсаргүй шинж чанар багасч ямар нэг тооны орчимд тогтворжно.
3.) Туршилтын тоо n хүрэлцээтэй их байх. Чухам ийм үед л уг үзэгдлийн харьцангуй давтамжийг магадлалтайгаа ойролцоогоор тэнцүү гэж үзнэ. Жишээлбэл, шинээр төрж буй 1000 хүүхэд тутмаас дунджаар 515 эрэгтэй хүүхэд төрж буйг олон жилийн ажиглалт, судалгаагаар тогтоожээ. Тэгвэл энэхүү ажиглалтанд эрэгтэй хүүхэд төрөх харьцангуй давтамж 0.515 байна. Иймд эрэгтэй хүүхэд төрөх үзэгдлийн магадлалыг $p = 0.515$-аар сонгон авч болно гэсэн үг юм.
Мөн 2.1-д авч үзсэн Бюффен, К. Пирсон нарын зоос хаях туршилтын үр дүнгээс үзвэл "зоос сүлдээрээ буух үзэгдлийн магадлал"-ыг $p = 0.5$ гэж үзэх үндэстэй байна.

2.7 Геометр магадлал

Эгэл үзэгдлийн огторгуй тасралтгүй (туршилтын боломжит үр дүн төгсгөлгүй олон) тохиолдолд сонгодог тодорхойлолтыг өргөтгөснөөр геометр магадлал гэсэн ойлголтонд хүрнэ.

R радиустай дугуйд багтсөн зөв гурвалжин өгөгдсөн гэж саная. "Дугуй руу таамгаар шидсэн цэг гурвалжин дотор унах үзэгдлийн магадлалыг ол" гэсэн бодлого практикт байж болно. Энэ бодлогыг магалалын сонгодог тодорхойлолтоор шийдэж болохгүй. Учир нь дугуйн цэг бүрийг эгэл цэг гэж үзэж болох бөгөөд боломж нь ижилдүү боловч бүх эгэл цэгийн тоо төгсгөлгүй олон. Гурвалжин дотор орших цэгүүд нь ивээлт цэгүүд болох ба мөн ивээлт цэгүүдийн тоо төгсгөлгүй. Энэ бодлого нь геометр магадлал гэсэн ойлголтонд хүргэдэг бодлогын жишээ юм.

Хавтгай дээр төгслөг талбайтай G гэсэн муж авья. G муж дотор g гэсэн муж өгөгдсөн гэе. Тэгвэл G мужид таамгаар шидсэн цэг g мужид унах үзэгдлийн магадлалыг ольё. Бидний олох магадлал нь G мужийн хаана унасан, g нь G мужийн хаана байрласан, ямар хэлбэртэй зэргээс хамаарахгүй. Харин G ба g мужийн талбайгаас, ялангуяа g мужийн талбайгаас хамаарна. Иймд сонирхож буй үзэгдлийн магадлал p нь

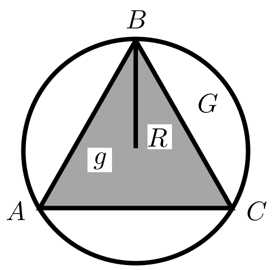
$$p = \frac{S_g}{S_G} = \frac{g \text{ мужийн талбай}}{G \text{ мужийн талбай}} \tag{2.8}$$

томъёогоор тодорхойлогдоно.

Санамсаргүй цэг ямар нэг мужид унах магадлал уг мужийн хэмжээнээс (урт, талбай, эзлэхүүн гэх мэт) хамаарах тохиолдолд геометр магадлалын тодорхойлолтыг хэрэглэх бөгөөд ерөнхий тохиолдолд

$$p = \frac{mes(g)}{mes(G)} \tag{2.9}$$

томъёогоор тодорхойлогдоно. Үүнд: G, g нь Лебегийн төгсгөлөг хэмжээ бүхий хэмжигдэх олонлогууд, $mes(g)$, $mes(G)$ нь харгалзан g, G олонлогийн Лебегийн хэмжээ (measure-хэмжээ). (2.9) томъёог геометр магадлалын томъёо гэнэ. Одоо дээр томъёолсон бодлогыг бодъё.



$$AB = BC = AC = R\sqrt{3},\ S_G = \pi R^2,$$
$$S_g = \frac{3}{4}\sqrt{3}R^2,\ p = \frac{S_g}{S_G} =$$
$$\frac{3\sqrt{3}/4}{\pi R^2}R^2 = \frac{3\sqrt{3}}{4\pi} = 0.4137$$

Зураг 2.5:

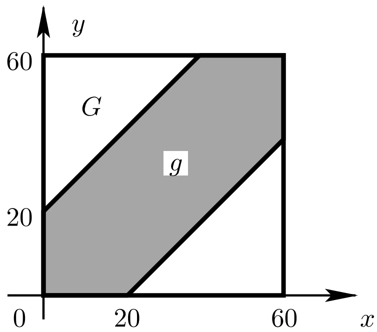
Жишээ-2.14 (Уулзах тухай бодлого)

A ба B хоёр оюутан k-аас $k+1$ цагийн хооронд тодорхой газар уулзахаар болзжээ. Хэн түрүүлж ирсэн нь 20 минут хүлээгээд явахаар тохиролцжээ. Болзсон оюутнууд k, $k+1$ цагийн хооронд аль ч агшинд ирэх боломжтой бөгөөд нэгний нь ирэх агшин нөгөөдөө огт нөлөөлөхгүй бол A ба B оюутны уулзах магадлалыг ол.

A оюутны ирэх агшныг $-x$, B оюутны ирэх агшныг $-y$ гэж тус тус тэмдэглэвэл (x,y) координаттай цэг нь нэг эгэл цэг болно.
$0 \le x \le 60, \;\; 0 \le y \le 60$ учир G муж нь

$$G = \{(x,y) \mid \;\; 0 \le x \le 60, \;\; 0 \le y \le 60\}$$

гэсэн квадрат болно. 2 оюутан уулзах зайлшгүй бөгөөд хүрэлцээтэй нөхцөл нь тэдний ирэх агшнуудын ялгавар 20-иос бага байх явдал юм.
Өөрөөр хэлбэл, $|x-y| \le 20$. Тэгвэл ивээлт цэгүүдийн олонлог
$g = \{(x,y) \mid \;\; |x-y| \le 20\}$ гэсэн муж байна.

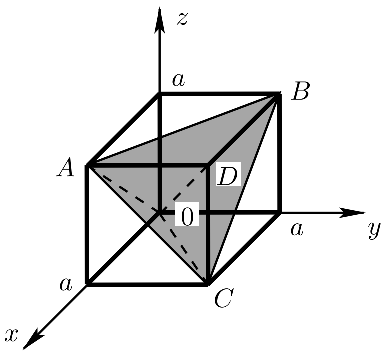


Зураг 2.6:

Мужуудыг Зураг 2.6-д дүрслэв.
$S_G = 60^2, S_g = 60^2 - 2 \cdot \dfrac{40^2}{2} = 60^2 - 40^2$.
Иймд уулзах магадлал:
$p = \dfrac{60^2 - 40^2}{60^2} = \dfrac{5}{9}$, $p = \dfrac{5}{9} > \dfrac{1}{2}$ учир уулзах боломж нь илүү байна.

Жишээ-2.15 a-аас хэтрэхгүй урттай хэрчмүүдээс таамгаар сонгон авсан 3 хэрчмээр гурвалжин байгуулж болох үзэгдлийн магадлалыг ол.

Таамгаар сонгон авсан 3 хэрчмийн уртыг x, y, z гэвэл (x,y,z) координаттай цэг нь нэг эгэл цэг болох ба G муж нь



Зураг 2.7:

$G = \{(x,y,z) \mid 0 \le x \le a, \; 0 \le y \le a, \; 0 \le z \le a\}$ гэсэн 3 хэмжээст огторгуйн куб болно. Эдгээр 3 хэрчмээр гурвалжин байгуулж болохын тулд $x+y \le z, \;\; x+z \le y, \;\; z+y \le x$ нөхцөлүүд зэрэг биелэгдэх ёстой. Өөрөөр хэлбэл,

$$g = \{(x,y,z) \mid \;\; x+y \le z, \;\; x+z \le y, \;\; z+y \le x\}.$$

g муж нь Зураг 2.7-д үзүүлсэн O, A, B, C, D цэгүүдэд оройтой олон талст болно. Бидний сонирхсон үзэгдлийн магадлал эдгээр мужуудын эзлэхүүнээс хамаарах тул $p = \dfrac{V_g}{V_G}$ хэлбэртэй болно. $V_g = a^3, \;\; V_G = a^3 - 3 \cdot \dfrac{a^3}{6} = \dfrac{a^3}{2}$.
Иймд $p = \dfrac{V_g}{V_g} = \dfrac{1}{2}$.

Бүлэг 3

Үзэгдлүүдийн үл хамарах чанар, хялбар томъёонууд

Магадлалын онолын чухал ойлголтуудын нэг нь үзэгдлүүдийн үл хамаарах чанар юм. Энэ чанарыг олж тогтоосноор үзэгдлүүдийн магадлалыг бодоход нилээд хялбар болдог.

3.1 Нөхцөлт магадлал. Үзэгдлүүдийн үл хамаарах чанар

Хэрэв A үзэгдэл явагдах магадлал B үзэгдлийн явагдах ба эс явагдахаас хамаарахгүй байвал A, B үзэгдлүүдийг хамааралгүй гэнэ. (A нь B-ээс хамаарахгүй) эсрэг тохиолдолд хамааралтай гэнэ. B үзэгдэл явагдсан нөхцөлд A үзэгдлийн магадлалыг авч үзэх шаардлага гарч болно. Ийм үед A-ийн магадлалыг, A үзэгдлийн B нөхцөл дахь магадлал гэж нэрлээд $P_B(A)$ буюу $P(A/B)$ гэж тэмдэглэнэ.
Жишээлбэл, хайрцагт 3 цагаан 2 хар бөмбөг байжээ.
Хайрцгаас 2 удаа нэг нэг бөмбөг таамгаар авах туршилт хийжээ гэж бодьё.
A- анхны авалтанд цагаан бөмбөг байх үзэгдэл.
B- дараагийн авалтанд цагаан бөмбөг байх үзэгдэл бол
B үзэгдлийн явагдах магадлал нь A үзэгдлийн явагдах ба эс явагдахаас хамаарна. Үнэхээр, $P(A) = \dfrac{3}{5}$ бөгөөд хэрэв A үзэгдэл явагдсан бол $P(B) = \dfrac{2}{4}$, A үзэгдэл явагдаагүй бол $P(B) = \dfrac{3}{4}$ юм. Иймд $P(B) = \dfrac{2}{4}$ гэсэн магадлалыг $P(B/A) = \dfrac{2}{4}$ гэж тэмдэглэх болж байна.
A үзэгдэлийн B нөхцөл дахь магадлал нь A ба B үзэгдлүүдийн үржвэрийн магадлалыг B үзэгдлийн магадлалд харьцуулсан харьцаатай тэнцүү байна.

Өөрөөр хэлбэл :

$$P(A/B) = \frac{P(AB)}{P(B)} \quad \text{буюу} \quad P(AB) = P(B) \cdot P(A/B) \tag{3.1}$$

Үүний адилаар

$$P(AB) = P(A) \cdot P(B/A) \tag{3.1$'$}$$

(3.1) ба (3.1′) томъёог нэгтгэвэл

$$P(AB) = P(A) \cdot P(A/B) = P(A) \cdot P(B/A) \tag{3.2}$$

(3.2)томъёог хамааралтай үзэгдлүүдийн магадлалыг үржүүлэх теорем гэнэ. Ийнхүү A, B үзэгдлүүдийн үржвэрийн магадлал нь аль нэг үзэгдлийн магадлалыг нөгөө үзэгдлийн энэхүү нөхцөл дахь магадлалаар үржүүлсэнтэй тэнцүү. (3.2) томъёог хоёроос дээш тооны үзэгдлүүд дээр дэлгэрүүлж болно:

$$\begin{aligned} &P(A_1A_2\dots A_k) = P(A_1) \cdot P(A_2A_3\dots A_k/A_1) = \\ &P(A_1) \cdot P(A_2/A_1) \cdot P(A_3A_4\dots A_k/A_1A_2) = \dots = \\ &P(A_1) \cdot P(A_2/A_1) \cdot P(A_3/A_1A_2) \cdots P(A_k/A_1A_2\dots A_{k-1}) \end{aligned} \tag{3.3}$$

Хэрэв

$$P(A/B) = P(A) \tag{3.4}$$

нөхцөл биелэгдэж байвал A-г B-ээс үл хамаарах үзэгдэл гэнэ.
Энэ тохиолдолд (3.2) томъёо

$$P(AB) = P(A) \cdot P(B) \tag{3.5}$$

болох ба (3.5) томъёог хамааралгүй үзэгдлүүдийн магадлалыг үржүүлэх теорем гэнэ. (3.5)-аас үзвэл, хамааралгүй үзэгдлүүдийн үржвэрийн магадлал нь тус бүрийн магадлалуудын үржвэртэй тэнцүү байна. Өөрөөр хэлбэл, A_i-үзэгдлүүд хамаралгүй бол $(i = \overline{1,n})$

$$P(A_1A_2\dots A_n) = P(A_1)P(A_2)\dots P(A_n) \tag{3.6}$$

Хэрэв A нь B-ээс үл хамаарах бол B нь A-аас үл хамаарна.Түүнчлэн A ба B үзэгдлүүд хамааралгүй бол A ба $\overline{B}$, $\overline{A}$ ба B, $\overline{A}$ ба $\overline{B}$ үзэгдлүүд бас хамааралгүй байна.

Жишээ-3.1 Оюутан магадлалын онолын шалгалтын 25 асуултын 5-ыг бэлдэж амжсангүй. Шалгалтын билетийн 3 асуултанд бүгдэд нь амжилттай хариулж чадах үзэгдлийн магадлалыг ол.
A- бүх 3 асуултанд зөв хариулах үзэгдэл.
A_i i-р асуултанд зөв хариулах үзэгдэл. $(i = \overline{1,3})$
Тэгвэл A нь $A = A_1 \cdot A_2 \cdot A_3$ болох ба
$P(A) = P(A_1A_2A_3) = P(A_1) \cdot P(A_2/A_1) \cdot P(A_3/A_1A_2)$.

1-р асуултанд зөв хариулах магадлал $P(A_1) = \dfrac{20}{25}$,
1-р асуултыг зөв хариулсан нөхцөлд 2-р асуултанд зөв хариулах магадлал $P(A_2/A_1) = \dfrac{19}{24}$, 1 ба 2-р асуултаа зөв хариулсан нөхцөлд 3-р асуултаа мөн зөв хариулах магадлал $P(A_3/A_1A_2) = \dfrac{18}{23}$.
Иинхүү $P(A) = \dfrac{20 \cdot 19 \cdot 18}{25 \cdot 24 \cdot 23} = \dfrac{57}{115}$ болно. Энэ бодлогыг магадлалын сонгодог тодорхойлолтоор бодвол $P(A) = \dfrac{C_{20}^3}{C_{25}^3} = \dfrac{57}{115}$ болж, ижил үр дүнд хүрнэ.

Жишээ-3.2 1891 онд Англид (Уэльс) эцэг хүүхэд хоёрын нүдний өнгөнд судалгаа хийхэд дараах баримтууд илэрчээ. Үүнд: эцэг нь хар нүдтэй хүү нь хар нүдтэй хүмүүс 5%, эцэг нь хар нүдтэй хүү нь цэнхэр нүдтэй хүмүүс 7.9%; эцэг нь цэнхэр нүдтэй хүү нь хар нүдтэй хүмүүс 8.9%; эцэг хүү хоёр хоёулаа цэнхэр нүдтэй хүмүүс 78.2% байв . Эцэг хүү хоёрын нүдний өнгөний холбоог тогтоо.
A-эцэг нь хар нүдтэй байх үзэгдэл
$\overline{A}$-эцэг нь цэнхэр нүдтэй байх үзэгдэл
B-хүү нь хар нүдтэй байх үзэгдэл
$\overline{B}$-хүү нь цэнхэр нүдтэй байх үзэгдлийг тус тус тэмдэглэвэл өгөгдсөн ёсоор $P(AB) = 0.05$; $P(A\overline{B}) = 0.079$; $P(\overline{A}B) = 0.089$; $P(\overline{A} \cdot \overline{B}) = 0.782$
а) Эцэг нь хар нүдтэй нөхцөлд хүү нь хар нүдтэй байх магадлалыг олъё.

$$P(B/A) = \frac{P(AB)}{P(A)} = \frac{P(AB)}{P(AB) + P(A\overline{B})} = \frac{0.05}{0.05 + 0.079} = 0.39$$

$$(P(A) = P(A(B + \overline{B})) = P(AB) + P(A\overline{B})).$$

б) Эцэг нь хар нүдтэй байхад хүү нь цэнхэр нүдтэй байх магадлалыг олъё.

$$P(\overline{B}/A) = 1 - P(B/A) = 1 - 0.39 = 0.61$$

в) Эцэг нь цэнхэр нүдтэй нөхцөлд хүү нь хар нүдтэй байх үзэгдлийн магадлал
$P(B/\overline{A}) = \dfrac{P(\overline{A}B)}{P(\bar{A})} = \dfrac{P(\overline{A}B)}{P(\overline{A}B) + P(\overline{A} \cdot \overline{B})} = \dfrac{0.089}{0.089 + 0.0782} = 0.102$
г) Эцэг нь цэнхэр нүдтэй байхад хүү нь цэнхэр нүдтэй байх магадлал

$$P(\overline{A}/\overline{B}) = 1 - 0.02 = 0,898.$$

Жишээ-3.3 Хоёр хайрцгийн 1-р хайрцагт 1 цагаан, 3 хар, 4 улаан харандаа; 2-р хайрцагт 3 цагаан, 2 хар, 3 улаан харандаа тус тус байжээ. Хайрцаг бүрээс нэг нэг харандаа таамгаар авахад өнгө нь ижил байх үзэгдлийн магадлалыг ол.

1-р хайрцгаас цагаан, хар, улаан өнгийн харандаа авах үзэгдлийг харгалзан B_1, C_1, D_1, 2-р хайрцгаас дээрх өнгийн харандаа авах үзэгдлийг харгалзан B_2, C_2, D_2 гэж тэмдэглэвэл бидний сонирхсон үзэгдэл $A = B_1B_2 + C_1C_2 + D_1D_2$ болно. Энд B_1B_2 ба C_1C_2, D_1D_2 үзэгдлүүд нийцгүй учир $P(A) = P(B_1B_2 + C_1C_2 + D_1D_2) = P(B_1B_2) + P(C_2C_2) + P(D_1D_2)$. Мөн B_1 ба B_2, C_1 ба C_2, D_1 ба D_2 үзэгдлүүд хамааралгүй учраас $P(A)=P(B_1)P(B_2)+P(C_1)P(C_2)+P(D_1)P(D_2)=\frac{1}{8}\cdot\frac{3}{8}+\frac{3}{8}\cdot\frac{2}{8}+\frac{4}{8}\cdot\frac{3}{8}=\frac{21}{64}$.

3.2 Ядаж нэг үзэгдэл явагдах магадлал

Хэд хэдэн үзэгдлүүдийн алиных нь ч явагдах магадлал бусдынхаа явагдах ба эс явагдахаас хамаарахгүй байвал тэдгээрийг хамааралгүй үзэгдлүүд гэнэ. Туршилтын дүнд $A_1, A_2, \ldots, A_n$ гэсэн хоорондоо хамааралгүй n үзэгдэл явагдах боломжтой бөгөөд үзэгдэл тус бүрийн явагдах магадлалууд $P(A_1) = p_1$, $P(A_2) = p_2, \ldots, P(A_n) = p_n$ болог. Туршилтаар бүх үзэгдлүүд, эсвэл аль нь ч явагдахгүй байж болно. Дээрх үзэгдлүүдийн эсрэг үзэгдлийг харгалзан

$$P(\overline{A}_1) = q_1, P(\overline{A}_2) = q_2, \ldots, P(\overline{A}_n) = q_n$$

гэе. Хэрэв A нь A_i үзэгдлүүдийн $(i = \overline{1,n})$ ядаж нэг нь явагдах үзэгдэл бол

$$P(A) = 1 - q_1 \cdot q_2 \cdot q_3 \cdots q_n \tag{3.7}$$

байна. Өөрөөр хэлбэл, A_1, $A_2, \ldots, A_n$ гэсэн хамааралгүй үзэгдлүүдийн ядаж нэг нь явагдах магадлал тэдгээрийн эсрэг үзэгдлийн магадлалуудын үржвэрийг нэгжээс хассантай тэнцүү байна. Тухайн тохиолдолд A_i үзэгдлүүд бүгд ижил боломжтой бол $(P(A_i) = p,\ i = \overline{1,n}))$ (3.7) томъёо

$$P(A) = 1 - q^n \tag{3.8}$$

хэлбэртэй болно $(q = 1 - p)$.

Жишээ-3.4 Буудлагын тамирчин бүр байг онох магадлал харгалзан $p_1 = 0,8$; $p_2 = 0,7$; $p_3 = 0,9$ байв. Тамирчин бүр бай руу 1 удаа зэрэг буудахад ядаж нэг нь оносон байх үзэгдлийн магадлалыг ол.

i-р таимрчин онох үзэгдэлийг $A_i, (i = \overline{1,3})$, Ядаж нэг нь онох үзэгдлийг-A гэе. Тамирчин тус бүрийн байг онох магадлал бусдынхаа онох, эс онохоос хамаарахгүй тул хоорондоо хамааралгүй үзэгдлүүд болно. $P(A_1) = q_1 = 1-0.8 = 0.2$; $P(A_2) = q_2 = 1-0.7 = 0.3$; $P(A_3) = q_3 = 1-0.9 = 0.1$ учир (3.7) томъёо ёсоор $P(A)=1-0.2\cdot 0.3\cdot 0.1=0.994$.

3.3 Гүйцэд магадлалын томъёо

$H_1, H_2, \ldots, H_n$ нь бүтэн бүлэг үүсгэх хос хосоороо нийцгүй үзэгдлүүд байг. A үзэгдэл H_i $(i = l, n)$ үзэгдлүүдийн аль нэг нь явагдсан нөхцөлд явагддаг гэж саная. Өөрөөр хэлбэл, $A = AH_1 + AH_2 + \cdots + AH_n$. Энэ тохиолдолд $H_i \ \ (i = \overline{1, n})$ үзэгдлийг таамаглалууд гэнэ. A үзэгдлийн магадлал нь эдгээр таамаглалуудын магадлалыг А үзэгдлийн харгалзах нөхцөлт магадлалаар үржүүлж нэмсэн нийлбэртэй тэнцүү байна.

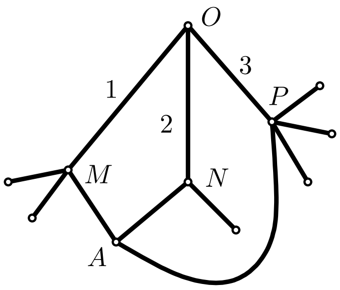
$$P(A) = \sum_{i=1}^{n} P(H_i)P(A/H_i) \tag{3.9}$$

Дэлгэрэнгүй бичвэл:

$$P(A) = P(H_1)P(A/H_1) + P(H_2)P(A/H_2) + \cdots + P(H_n)P(A/H_n).$$

(3.9) томъёог гүйцэд магадлалын томьёо гэнэ. (3.9) нь таамаглал тус бүрийн магадлалыг энэ таамаглалын нөхцөл дахь A үзэгдлийн магадлалаар үржүүлсэн үржвэрүүдийн нийлбэр юм.

Жишээ-3.5 Аялагч O хотоос гарахад очих газрын зам 3 салсан учир аль нэгийг нь таамгаар сонгон авах шаардлагатай болжээ. Сонгосон зам тус бүр дахин хэд хэд салаалсан байсан бөгөөд аль ч замаар нь явсан очих газраа хүрч болох байлаа. Хэрвээ аялагч аль ч замаар явах боломжтой (дөт замаар явах албагүй) гэвэл А хотод очих үзэгдлийн магадлалыг ол. Замын схемийг Зураг 3.1-д үзүүлэв. Энд дараах таамаглалууд дэвшүүлж болно.



Зураг 3.1:

H_1- эхний удаа 1-р замыг сонгон авах үзэгдэл
H_2- 2-р замыг сонгон авах үзэгдэл
H_3- 3-р замыг сонгон авах үзэгдэл

Таамгаар сонгон авч буй учир эдгээр замуудын аль нэгийг сонгон авах боломж ижил. Өөрөөр хэлбэл,
$P(H_1) = P(H_2) = P(H_3) = \dfrac{1}{3}$. Аялагч A хотод очих үзэгдлийг B-ээр тэмдэглэе. Тэгвэл аялагч 1-р замыг сонгон авсан нөхцөлд A хотод очих үзэгдлийн магадлал $\dfrac{1}{3}$ болно. Өөрөөр хэлбэл, $P(B/H_1) = \dfrac{1}{3}$,
Үүнчлэн $P(B/H_2) = \dfrac{1}{2}$, $P(B/H_3) = \dfrac{1}{4}$. Иймд (3.9) томъёо ёсоор:
$P(B) = P(H_1)P(B/H_1) + P(H_2)P(B/H_2) + P(H_3)P(B/H_3) = \dfrac{1}{3} \cdot \dfrac{1}{3} + \dfrac{1}{3} \cdot \dfrac{1}{2} + \dfrac{1}{3} \cdot \dfrac{1}{4} = \dfrac{13}{36}$.

Жишээ-3.6 Таны баруун халаасанд 3ш 20-тын мөнгө, 4ш 15-тын мөнгө, зүүн халаасанд 6ш 20-тын, 3ш 15-тын мөнгө байсан гэе. Та баруун халааснаасаа таамгаар 5 мөнгө авч зүүн халаасандаа хийгээд дараа нь зүүн халааснаасаа таамгаар 1 мөнгө авсан байг. Таны зүүн халааснаасаа авсан мөнгө 20-тын мөнгө байх үзэгдлийн магадлалыг ол.

Энэ тохиолдолд дараах 3 таамаглал байж болно.
H_1- Баруун халааснаас 1ш 20-тын , 4ш 15-тын мөнгө авах үзэгдэл,

H_2- 2ш 20-тын, 3ш 15-тын мөнгө авах үзэгдэл,
H_3- 3ш 20-тын, 2ш 15-тын мөнгө авах үзэгдэл.
Тэгвэл $P(H_1)=\frac{C_3^1C_4^4}{C_7^5}=\frac{1}{7},\ P(H_2)=\frac{C_3^2C_4^3}{C_7^5}=\frac{4}{7},\ P(H_3)=\frac{C_3^3C_4^2}{C_7^5}=\frac{2}{7}.$
Сонирхсон үзэгдлээ A гэвэл:
$P(A/H_1)=\frac{7}{14},\ P(A/H_2)=\frac{8}{14},\ P(A/H_3)=\frac{9}{14}.$ Иймд.

$$P(A)=\sum_{i=1}^{3}P(H_i)P(A/H_i)=\frac{1}{7}\cdot\frac{7}{14}+\frac{4}{7}\cdot\frac{8}{14}+\frac{2}{7}\cdot\frac{9}{14}=0,54.$$

3.4 Байесийн томъёо

$H_1, H_2, \ldots, H_n$-нь бүтэн бүлэг үүсгэх, хос хосоороо нийцгүй таамаглалууд ба туршилт хийхийн өмнө эдгээр таамаглалын магадлал $P(H_1),\ P(H_2),\ldots,P(H_n)$ мэдэгдэж байг. Туршилт хийгдэж санамсаргүй үзэгдэл A явагдсан гэе. Тэгвэл A үзэгдэл явагдсаны дараа анхны таамаглалуудын магадлал яаж өөрчлөгдөх вэ? гэсэн бодлого авч үзье. Өөрөөр хэлбэл $P(H_i)$, $P(A/H_i)$ магадлалуудын тусламжтайгаар $P(H_i/A)$ магадлалуудыг олно гэсэн үг.
$P(A\cdot H_i)=P(A)P(H_i/A)=P(H_i)(A/H_i)\quad (i=\overline{1,n})$ ба (3.9) томьёог анхаарвал

$$P(H_i/A)=\frac{P(H_i)P(A/H_i)}{\sum\limits_{i=1}^{n}P(H_i)P(A/H_i)} \tag{3.10}$$

(3.10) томъёог таамаглалуудын теорем буюу Байесийн томьёо гэнэ.

Жишээ-3.7 Ямар нэг салбарын нийт бүтээгдэхүүний 30%-ийг I үйлдвэр, 25%-ийг II үйлдвэр, үлдсэн хувийг III үйлдвэр хийдэг байжээ. I үйлдвэрийн хийсэн бүтээгдхүүний 1%, II-ын 1.5%, III-ын 2% тус тус гологдол байдаг байв. Хэрэглэгчийн худалдан авсан бүтээгдэхүүн гологдол байв. Энэхүү гологдол I үйлдвэрийн бүтээгдэхүүн байх үзэгдлийн магадлалыг ол.
H_1-бүтээгдэхүүн I-р үйлдвэрийнх байх үзэгдэл,
H_2-бүтээгдэхүүн II-р үйлдвэрийнх байх үзэгдэл,
H_3-бүтээгдэхүүн III-р үйлдвэрийнх байх үзэгдэл,
A- худалдан авсан бүтээгдхүүн гологдол байх үзэгдэл бол
$P(H_1)=0.30,\ P(H_2)=0.25,\ P(H_3)=0.45,$
$P(A/H_1)=0.01,\ P(A/H_2)=0.015,\ P(A/H_3)=0.02$ гэж өгөгдсөн тул
$P(A)=P(H_1)P(A/H_1)+P(H_2)P(A/H_2)+P(H_3)P(A/H_3)=$
$0.30\cdot 0.01+0.25\cdot 0.015+0.45\cdot 0.02=0.015,\ \ P(H_1/A)=\frac{0.3\cdot 0.01}{0.015}=0,2.$
Иймд нийт гологдол бүтээгдэхүүний 20 орчим хувийг I үйлдвэр үйлдвэрлэжээ. Жишээнээс үзвэл туршилтын дараа H_1 таамаглалын магадлал туршилтын өмнөх магадлалаасаа багассан байна.

Жишээ-3.8 Хоёр тамирчин бие биеэсээ үл хамааран бай руу нэг нэг удаа бууджээ. I тамирчин байг онох магадлал 0.8, II тамирчин байг онох магадлал 0.4. Бай руу буудсаны дараа хэн нэг нь оножээ. I буудагч байг оносон байх үзэгдлийн магадлалыг ол.

Туршилт явагдахын өмнө дараах таамаглалууд байж болно:
H_1—I ба II тамирчны хэн нь ч онохгүй, H_2—I ба II тамирчин хоёулаа онох, H_3—I нь оножи, II нь онохгүй, H_4—I нь онохгүй, II нь онох.
Таамаглал тус бүрийн магадлал:

$$P(H_1) = 0.2 \cdot 0.6 = 0.12, \qquad P(H_2) = 0.8 \cdot 0.4 = 0.32,$$
$$P(H_3) = 0.8 \cdot 0.6 = 048, \qquad P(H_4) = 0.2 \cdot 0.4 = 0.08.$$

Байг 1 онох үзэгдлийг A-аар тэмдэглэж, дээрх тамаглал тус бүрийн үе дэх A үзэгдлийн магадлалыг олбол
$P(A/H_1) = 0,\ P(A/H_2) = 0,\ P(A/H_3) = 1,\ P(A/H_4) = 1.$ Харин туршилт явагдсаны дараа (байг 1 оносны дараа) H_1, H_2-таамаглалууд боломжгүй болох бөгөөд H_3, H_4 таамаглал тус бүрийн магадлал нь

$$P(H_3/A)=\frac{P(H_3)P(A/H_3)}{P(H_3)P(A/H_3)+P(H_4)P(A/H_4)}=\frac{0.048 \cdot 1}{0.48 \cdot 1+0.08 \cdot 1}=\frac{6}{7}.$$

$$P(H_4/A)=\frac{P(H_4)P(A/H_4)}{P(H_3)P(A/H_3)+P(H_4)P(A/H_4)}=\frac{0.08 \cdot 1}{0.48 \cdot 1+0.08 \cdot 1}=\frac{1}{7}.$$

Ийнхүү I тамирчин байг оносон байх үзэгдлийн магадлал $\frac{6}{7}$ байна.

Бүлэг 4

Бернуллийн схем

4.1 Үл хамаарах туршилтуудын дараалал. Бернуллийн томъёо.

Туршилтын дүнд A ба $\overline{A}$ үзэгдлийн зөвхөн нэг нь явагдах бөгөөд A үзэгдэл p магадлалтай илэрдэг гэе. Өөрөөр хэлбэл, $P(A)=p$, $P(\overline{A})=q$ $(p+q=1)$. Дээрх туршилтыг нэгэн ижил нөхцөлд олон дахин давтан хийе. Ийм туршилтуудын дараалллыг үл хамаарах, санамсаргүй туршилтын дараалал буюу Бернуллийн схем гэнэ. Ийм санамсаргүй туршилтын жишээ амьдрал практикт олон тааралддаг. Тухайлбал, технологийн нэг ижил нөхцөлд үйлдвэрлэж буй бүтээгдэхүүн стандартын шаардлаганд тохирох, эс тохирох; t хугацааны туршид радио идэвхт бодис задрах, эс задрах; мөнгө болон шоог олон дахин хаяхад сонирхсон үзэгдэл явагдах, эс явагдах гэх мэт.

Туршилтыг n удаа давтан хийхэд A үзэгдэл яг k удаа явагдах магадлалыг олъё. Сонирхож буй үзэгдлээ $A_{n,k}$, түүний магадлалыг $P_n(k)$ гэе.
A_i нь i-р $(i=\overline{1,n})$ туршилтанд A явагдах үзэгдэл болог.
Тэгвэл $A_{n,k}$ үзэгдлийг дараах байдлаар илэрхийлж болно.

$$A_{n,k}=A_1A_2\dots A_k\overline{A}_{k+1}\dots\overline{A}_n+\dots+\overline{A}_1\overline{A}_2\dots\overline{A}_{n-k}A_{n-(k-1)}\dots A_n$$

n удаа туршилт хийхэд A үзэгдэл k удаа явагдах нийт тоо C_n^k ширхэг байна.
Бүх нэмэгдэхүүнүүд хос хосоороо нийцгүй учир

$$P_n(k)=P(A_{n,k})=\underbrace{p^kq^{n-k}+\dots+p^kq^{n-k}}_{C_n^k}=C_n^kp^kq^{n-k},\quad \text{буюу}$$

$$P_n(k)=C_n^kp^kq^{n-k},\quad (k=\overline{0,n}) \tag{4.1}$$

(4.1) томъёог Бернуллийн томъёо гэнэ.

Жишээ-4.1 A, B хоёр тоглогч зоос хаяж тоглохоор болжээ. Хэрэв зоосыг 3 хаяхад сүлдээрээ 2 удаа буувал A тоглогч хожихоор, 4 удаа хаяхад

3 удаа сүлд буувал B тоглогч хожихоор тохирчээ. Аль тоглогчийн хожих боломж илүү вэ?

Зоосыг ижил нөхцөлд дахин давтан хаяж буй тул Бернуллийн схем болно. Сүлдээрээ буух үзэгдлийг C гэвэл: $P(C)=P(\overline{C})=\frac{1}{2}(p=q=\frac{1}{2})$. A тоглогчийн хожих магадлал (4.1) томъёо ёсоор $P_3(2)=C_3^2p^2q^1=3\left(\frac{1}{2}\right)^2\cdot\left(\frac{1}{2}\right)=\frac{3}{8}$,
B тоглогчийн хожих магадлал: $P_4(3)=C_4^3p^3q^1=4\left(\frac{1}{2}\right)^3\cdot\left(\frac{1}{2}\right)^1=\frac{1}{4}$,
$P_3(2) > P_4(3)$ учир A тоглогчийн хожих магадлал давуу байна.

Жишээ-4.2 Нэг хоногийн турш цахилгаан энергийн зарцуулалт өгөгдсөн нормоос давахгүй байх магадлал $p = 0.75$. Ойрын 6 хоногийн дөрөвт нь цахилгаан энергийн зарцуулалт нормоосоо давахгүй байх үзэгдлийн магадлалыг ол.

6 хоног тутамд цахилгаан энерги хэвийн зарцуулах үзэгдлийн магадлал тогтмол $p = 0.75$ учир хоног тутамд энерги илүү зарцуулах үзэгдлийн магадлал $q = 1 - 0.75 = 0.25$. Иймд сонирхсон магадлал Бернуллийн томъёо ёсоор

$$P_6(4) = C_6^4p^4q^2 = \frac{6\cdot 5}{1\cdot 2}(0.75)^4(0.25)^2 = 0,3.$$

4.2 Хамгийн их магадлалтай тоо

$P_n(k) = C_n^kp^kq^{n-k}$ томъёонд n-ийг бэхэлбэл $P_n(k)$ нь k-аас хамаарсан функц болно. Одоо k-ийн ямар утганд $P_n(k)$ функц максимум утгаа авах вэ?
Өөрөөр хэлбэл, $P_n(0),\ P_n(1),\dots,P_n(n)$ тоонуудын аль нь хамгийн их вэ? гэсэн асуудал тавья.
Тэгвэл np−д ойрхон байх k_0 тоо нь хамгийн их магадлалтай гэж харуулж болно. Тодорхой хэлбэл, Хэрэв $np - q$ бүхэл тоо биш бол

$$np - q < k_0 < np + p \tag{4.2}$$

нөхцлийг хангах k_0 тоо хамгийн их магадлалтай тоо байна.
Хэрэв $np - q$ нь бүхэл тоо бол $k_{01} = np - q$, $k_{02} = np + p$ гэсэн хоёр утганд $P_n(k)$ функц хамгийн их утгаа авна.

Жишээ-4.3 Ажилчин ижил төрлийн 12 машинд үйлчилнэ. Машин 1 цагийн дотор ажилчны үйлчилгээ шаардах магадлал $\frac{1}{3}$ бол 1 цагийн дотор ажилчнаас үйлчилгээ шаардах машины хамгийн их магадлалтай тоо ба хамгийн их магадлалыг ол.

Манай тохиолдолд $n = 12,\ p = \frac{1}{3}$ учир $np-q = 12\cdot\frac{1}{3}-\frac{2}{3} = 3\frac{1}{3}$ бүхэл биш. Имйд (4.2) ёсоор хамгийн их магадлалтай тоо $k_0 = 4$ байна. Бернуллийн томъёо ёсоор

$$P_{12}(4) = C_{12}^4 \cdot p^4 \cdot q^8 = 495 \cdot \left(\frac{1}{3}\right)^4 \cdot \left(\frac{2}{3}\right)^8 = 0,238.$$

4.3 Муавр-Лапласын локаль, интеграл томъёо

Бернуллийн томъёог хэрэглэх явцад n ба k-ийн хүрэлцээтэй их утганд $P_n(k)$ магадлалыг олох шаардлага олонтаа гардаг. Тухайлбал, нэг ижил детал үйлдвэрлэдэг үйлдвэрийн бүтээгдэхүүн гологдол байх магадлал $0,02$. 500 бэлэн бүтээгдэхүүний дотор 10 гологдол бүтээгдэхүүн байх магадлалыг ол гэсэн бодлогонд уг магадлал $P_{500}(10) = C_{500}^{10}(0.02)^{10}(0.98)^{490}$ гэж олдох ба C_{500}^{10}-тоог бодоход түвэгтэй байна. Цаашилбал, 500 бүтээгдэхүүний дотор байх гологдлын тоо 10-аас 20-ын дотор байх үзэгдлийн магадлалыг олъё гэвэл $P(10 \le k \le 20) = P_{500}(10) + P_{500}(11) + \cdots + P_{500}(20)$ гэж олдоно. Өмнөх бодлогоос хэд дахин төвөгтэй байна. Ийм учраас туршилтын тоо $n \to \infty$ үед $P_n(k)$ магадлалыг хялбар бөгөөд бага алдаатайгаар ойролцоо бодох аргыг английн математикч Муавр, францын математикч Лаплас нар гаргажээ.
n-хүрэлцээтэй их үед дараах томъёо хүчинтэй байна.

$$P_n(k) \approx \frac{1}{\sqrt{npq}} \cdot \varphi(x) \tag{4.3}$$

Үүнд: $x = \dfrac{k - np}{\sqrt{npq}}$, $\varphi(x) = \dfrac{1}{\sqrt{2\pi}} \cdot e^{-x^2/2}$
(4.3) томъёо нь Бернуллийн томъёоны $n \to \infty$ үеийн үнэлгээ бөгөөд Муавр-Лапласын локаль томъёо гэж нэрлэгдэнэ.
$\varphi(x)$ функцийн аргументийн эерэг утгуудын таблицыг зохиосон байдаг. $\varphi(x)$ функц $x > 0$ үед монотон буурах бөгөөд $\varphi(-x) = \varphi(x)$ буюу тэгш функц учир сөрөг утгуудын таблиц нь дээрх таблицтай адил байна. $\varphi(5) = 0,0000015$ ба $x > 5$ үед $\varphi(x) \approx 0$ гэж тооцно. Иймд $\varphi(x)$ функцийн таблиц $x > 5$ үед эс зохиогдоно.(Таблиц N1-ийг үз) Одоо дээр томъёолсон бодлогыг бодъё.
$n = 500,\ p = 0.02,\quad np = 500 \cdot 0.02 = 10,\ \ npq = 10 \cdot 0.98 = 9,8$ үед (4.3) томъёо ёсоор

$$P_{500}(10) = \frac{1}{\sqrt{9.8}} \cdot \varphi\left(\frac{10 - 10}{\sqrt{9.8}}\right) = \frac{1}{\sqrt{9.8}} \cdot \varphi(0)$$

Таблицаас харвал $\varphi(0) = 0.397$ учир $P_{500}(10) \approx 0,124$.

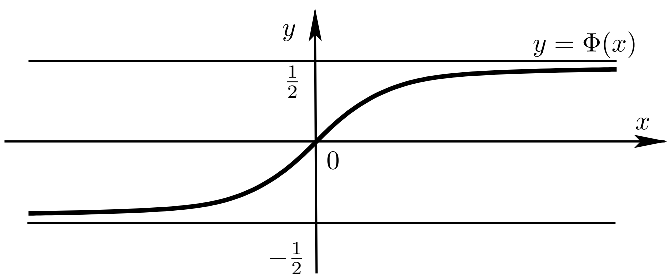
Зарим тохиолдолд n-ийн хүрэлцээтэй их утганд $P_n(k)$ магадлалыг биш, $P_n(k_1 \le k \le k_2)$ магадлалыг сонирхох шаардлага гардаг. Тэгвэл бидний сонирхсон үзэгдэл k_1-ээс цөөнгүй, k_2-оос олонгүй удаа явагдах үзэгдлийн магадлал Муавр-Лапласын интеграл томъёогоор үнэлэгдэнэ.
n-ийн хүрэлцээтэй их утганд

$$P_n(k_1 \le k \le k_2) \approx \Phi\left(\frac{k_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{k_1 - np}{\sqrt{npq}}\right) \tag{4.4}$$

томъёо хүчинтэй байна. Үүнд:

$$\Phi(x) = \int\limits_0^x \varphi(t)dt = \frac{1}{\sqrt{2\pi}} \int\limits_0^x e^{-t^2/2}dt - \text{Лапласын функц.}$$

$\Phi(x)$ функц дараах чанаруудыг хангана.
1. $\Phi(x)$-сондгой функц: $\Phi(-x) = -\Phi(x)$
2. x нь $[0;\infty[$ завсарт өсөхөд $\Phi(x)$ функц $[0;\frac{1}{2}[$ завсарт өснө. Функцийн графикийг Зураг 4.1-д үзүүлэв.



Зураг 4.1:

3. $x\!=\!5$ үед $\Phi(5)$ нь $\frac{1}{2}$-ээс $3\cdot 10^{-8}$-ээр ялгагдана. Өөрөөр хэлбэл, $x > 5$ үед $\Phi(x) \approx \frac{1}{2}$. Иймд $\Phi(x)$ функцийн таблиц $[0;4]$ завсарт зохиогдсон байна. (Хавсралтын таблиц N2-ыг үз). Өмнөх бодлогын $P(10 \le k \le 20)$ магадлалыг олъё. (4.4) томъёо ёсоор

$$P_n(10 \le k \le 20) \approx \Phi\left(\frac{20-10}{\sqrt{9.8}}\right) - \Phi\left(\frac{10-10}{\sqrt{9.8}}\right) \approx \Phi(3.19) - \Phi(0).$$

$\Phi(x)$ функцийн утгыг таблицаас харвал $\Phi(3.19) = 0.49929$,
$\Phi(0) = 0$. Иймд $P_n(10 \le k \le 20) \approx \Phi(3.19) - \Phi(0) = 0,49929$.
Санамж: (4.3) ба (4.4) томъёоны нарийвчлал npq үржвэр өсөх тутам сайжрах учраас Муавр-Лапласын локаль ба интеграл томъёог ихэвчлэн $npq \ge 10$ тохиолдолд хэрэглэнэ. p, q магадлалын аль нэг нь "0"-д хэдий чинээн ойрхон байвал n-ийг төдий чинээн ихээр сонгон авахад хүрнэ. Иймд $p \to 0$ (эсвэл $q \to 0$) үед (4.3) ба (4.4) томъёог үл хэрэглэнэ.

Жишээ-4.4 Заводын үйлдвэрлэсэн гэрлийн шил гологдол байх магадлал 0.02. 1000 гэрлийн шилийг шалгахаар сонгон авчээ. Түүвэр дэх гологдол шилний харьцангуй давтамж нь магадлалаасаа 0.01-ээс багаар ялгагдах магадлалыг үнэл.
Түүвэр дэх гологдол шилний тоог k гэвэл $\left|\frac{k}{1000} - 0.02\right| < 0.01$ байх магадлалыг үнэлнэ гэсэн үг. Энэ тэнцэтгэл биш нь $11 \le k \le 29$ тэнцэтгэл биштэй

тэнцүү чанартай. Иймд $P\left(\left|\frac{k}{1000}-0.02\right|<0.01\right)=P_{1000}(11\le k\le 29)$.
$npq=1000\cdot 0.02\cdot 0.98=19.6>10$ учир Лапласын интеграл томъёог хэрэглэе.
Манай тохиолдолд
$x_1=\frac{11-1000\cdot 0.02}{\sqrt{100\cdot 0.02\cdot 0.98}}\approx -2.03,\ x_2=\frac{29-20}{4.43}\approx 2,03$. $\Phi(-2.03)\approx -0.4788$,
$\Phi(2.03)\approx 0.4788$ Иймд (4.4) томъёо ёсоор

$$P_{100}(11\le k\le 29)\approx 0.4788+0.4788=0,9576.$$

4.4 Пуассоны томъёо

Өмнөх зүйлийн санамжид дурдсан ёсоор, тухайн туршилт бүрд A үзэгдлийн явагдах магадлал p харьцангүй бага үед (ийм үзэгдлийг ховор үзэгдэл гэнэ.) Муавр-Лапласын томъёоны нарийвчлал буурна. Ийм учраас энэ тохиолдолд A үзэгдэл k удаа явагдах магадлал Пуассоны хязгаарын томъёогоор үнэлэгдэнэ.
Бернуллийн схемийн нөхцөлд $n\to\infty,\ \ p\to 0$ бөгөөд $\lambda=np$ хэмжигдэхүүн тогтмол ($\lambda=const$) бол дараах томъёо хүчинтэй байна.

$$P_n(k)\approx\frac{\lambda^k}{k!}e^{-\lambda} \tag{4.5}$$

(4.5) томъёоны нарийвчлал $\left|P(k)-\frac{\lambda^k}{k!}e^{-\lambda}\right|<np^2$ тэнцэтгэл бишээр үнэлэгдэнэ.
Пуассоны томъёо (4.5)-ын гайхамшигтай чанар нь: тодорхой тооны туршилтын магадлалыг олохын тулд n ба p тоог мэдэх шаардлагагүй бөгөөд харин нийт туршилтанд A үзэгдлийн явагдах дундаж тоо болох $\lambda=n\cdot p$ тоог мэдэхэд хангалттай байдагт оршино.
Пуассоны томъёоны $P_n(k)=\frac{\lambda^k}{k!}\cdot e^{-\lambda}$ функцийн таблицыг k ба λ-ийн утгуудад зохиосон байдаг. (Хавсралтын таблиц N3-ыг үз.)

Жишээ-4.5 Зуурсан гуриланд тодорхой тооны үзэм хийж хольсны дараа тэнцүү хэсгүүдэд хэрчиж үзэмтэй талх хийнэ. Нийт талхны тоо N, бүх үзэмний тоо n бол таамгаар сонгон авсан нэг талханд яг k үзэм байх магадлалыг үнэл.

n ширхэг үзэм гуриланд холихыг, үзэм тус бүрийг гуриланд хаях замаар туршилтыг n удаа хийсэн гэж үзэж болно. Бидний сонгон авсан талханд үзэм орсон байх үзэгдлийг A гэе. Бүх талхны тоо N, үзмээ гуриланд хийгээд хольж байгаа учир үзэм бүр нэгэн ижил магадлалтайгаар аль ч талханд орж болно. Иймд A үзэгдлийн магадлал $p=\frac{1}{N}$, Үйлдвэрлэх талхны тоо N олон учраас p-магадлал бага тоо байна. Харин $\lambda=np$ нь $\frac{n}{N}$-тэй тэнцүү бөгөөд 1 талханд орох үзэмний дундаж тоо болно. Сонирхсон магадлал $P_n(k) \approx \dfrac{\lambda^k}{k!}e^{-\lambda}$ томъёогоор бодогдох бөгөөд $\lambda = 8$ тохиолдолд зарим утгуудыг олбол

$P_n(0) \approx 0.000 \quad P_n(1) \approx 0.003 \quad P_n(2) \approx 0.011$
$P_n(3) \approx 0.029 \quad P_n(4) \approx 0.057 \quad P_n(5) \approx 0.092$
$P_n(6) \approx 0.012 \quad P_n(7) \approx 0.139 \quad P_n(8) \approx 0.139$

гэх мэтчилэн бодоход $k > 14$ үед $P_n(k) < 0.001$ болно.
Эндээс үзвэл нийт N талхны дунджаар 0.31% нь 1 үзэм, 1.1% нь 2 үзэм, 2.9% нь 3 үзэм гэх мэт тус тус агуулж байна.

Жишээ-4.6 Утасны станц 400 хэрэглэгчид үйлчилнэ. Хэрэглэгч бүр 1 цагийн туршид станц уруу утасдах магадлал 0.01. цагийн дотор 5 хэрэглэгч уг станц уруу утасдах магадлалыг ол.

Манай тохиолдолд $p = 0.01 \quad n = 400$ учир Пуассоны томъёог хэрэглэж болно. $\lambda = 400 \cdot 0.01 = 4$. (4.5) томъёо ёсоор $P_{400}(5) \approx \dfrac{4^5}{5!}e^{-4} \approx 0.156293$ Хэрэв 1 цагийн дотор 4-өөс олонгүй хэрэглэгч утасдах магадлалыг олъё гэвэл $P_{400}(0 \le k \le 4) = P_{400}(0) + P_{400}(1) + P_{400}(2) + P_{400}(3) + P_{400}(4) \approx 0.018316 + 0.073263 + 0.146525 + 0.195367 + 0.195367 = 0.628838$ болно.

Бүлэг 5

Санамсаргүй хэмжигдэхүүн, түүний тархалтын хууль

5.1 Санамсаргүй хэмжигдэхүүн

Туршилтын үр дүнд ямар нэг тоон утга авах боловч тэрхүү үр дүнг урьдчилан хэлэх боломжгүй хэмжигдэхүүнийг санамсаргүй гэнэ. Санамсаргүй хэмжигдэхүүнийг дискрет ба тасралтгүй гэж хоёр ангилна. Санамсаргүй хэмжигдэхүүний авч болох утгуудыг түүний боломжит утгууд гэнэ.

Хэрэв санамсаргүй хэмжигдэхүүний боломжит утгууд нь төгсгөлөг, эсвэл төгсгөлгүй тоон дарааллын хэлбэрээр бичигдэж байвал дискрет санамсаргүй хэмжигдэхүүн гэнэ. Хэрэв санамсаргүй хэмжигдэхүүний боломжит утгууд нь тоон тэнхлэгийн ямар нэг завсраас бүрдэж байвал түүнийг тасралтгүй гэнэ.

Дискрет санамсаргүй хэмжигдэхүүн ξ-ийн боломжит утгууд $x_1, x_2, \dots, x_n$ ба эдгээр утгуудыг хүлээн авах магадлалууд $p_1, p_2, \dots, p_n$-ийн хоорондох холбоог уг хэмжигдэхүүний тархалтын хууль гэнэ.
Тэгвэл тодорхойлолт ёсоор $P(\xi = x_i) = p_i, \quad (i = \overline{1,n})$ болно. Санамсаргүй хэмжигдэхүүний тархалтын хууль нь таблиц, график, томъёогоор өгөгдөх ба дискрет санамсаргүй хэмжигдэхүүн гол төлөв таблицаар өгөгдөнө.

$\xi:$	x_1	x_2	$\dots$	x_n	$\dots$
$p:$	p_1	p_2	$\dots$	p_n	$\dots$

Санамсаргүй хэмжигдэхүүн$-\xi$, туршилтын үр дүнд боломжит утгынхаа аль нэгийг авах үзэгдлийн магадлал нь $P(\xi = x_1) + P(\xi = x_2) + \ldots + P(\xi = x_n)$ болох бөгөөд энэ нь зайлшгүй үзэгдэл тул $\sum\limits_{i=1}^{n} P(\xi = x_i) = \sum\limits_{i=1}^{n} p_i = 1$ байна.

5.2 Тархалтын функц түүний чанар

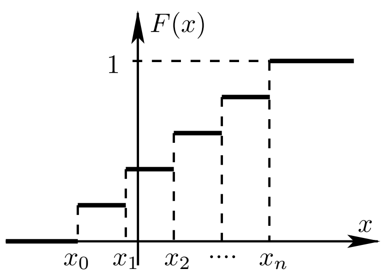
Тоон тэнхлэг дээр $F(x)$ гэсэн нэг хувьсагчийн функц тодорхойлогдсон гэж үзье. $\forall x-$ийн хувьд "ξ нь x-ээс бага утгаа авах магадлал"-ыг илэрхийлсэн функцийг санамсаргүй хэмжигдэхүүний тархалтын функц гэнэ. Өөрөөр хэлбэл:

$$F(x) = P(\xi < x) \tag{5.1}$$

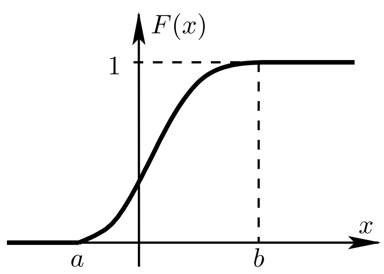
$F(x)$ функцийг ξ-санамсаргүй хэмжигдэхүүний тархалтын функц гэнэ.
Тархалтын интеграл функц, интеграл хууль гэх нь бий.
Тархалтын функцийн чанаруудыг авч үзье.
1. $F(x)$ бүх тоон тэнхлэг дээр тодорхойлогдох ба $0 \leq F(x) \leq 1$.
2. $F(x)$ үл буурна, өөрөөр хэлбэл: $x_1 \leq x_2$ үед $F(x_1) \leq F(x_2)$.
3. Хэрэв ξ-ийн боломжит утгууд $]-\infty, +\infty[$ завсарт байвал
$F(-\infty) = \lim\limits_{x \to -\infty} F(x) = 0, \quad F(+\infty) = \lim\limits_{x \to +\infty} F(x) = 1$ биелнэ.
4. $P(x_1 \leq \xi \leq x_2) = F(x_2) - F(x_1)$.
Дискрет санамсаргүй хэмжигдэхүүний тархалтын функц нь дараах томъёогоор илэрхийлэгдэнэ.

$$F(x) = \sum_{x_i \leq x} P(\xi = x_i)$$

Үүнд, $x_i < x$ байх бүх i-дугааруудын хувьд нийлбэр авагдана.
Хэрэв дискрет ба тасралтгүй санамсаргүй хэмжигдэхүүний тархалтын функцүүдийг графикаар дүрсэлбэл:



Зураг 5.1:



Зураг 5.2:

Тасралтгүй санамсаргүй хэмжигдэхүүний тархалтын функцийг сөрөг биш утгатай, хэсэг хэсэг тасралтгүй функц $f(x)$-ийн тусламжтай дараах байдлаар илэрхийлж болно.

$$F(x) = \int_{-\infty}^{x} f(t)dt \tag{5.2}$$

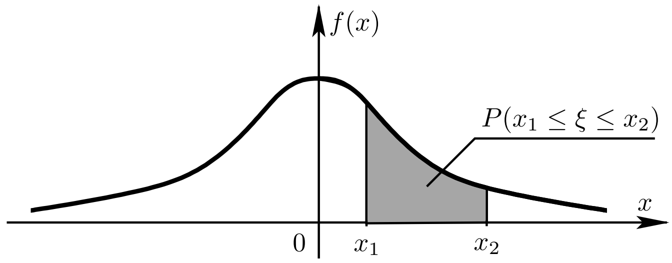
$f(x)$-ийг тархалтын нягтын функц буюу тархалтын дифференциал функц гэнэ. Нягтын функцийн тодорхойлолтоос дараах чанарууд мөрдөн гарна.
1. $f(x) \geq 0$.

2. $P(x_1 \le \xi \le x_2) = \int\limits_{x_1}^{x_2} f(x)dx.$

3. $\int\limits_{-\infty}^{+\infty} f(x)dx = 1, \quad \left(\xi \in [a,b] \text{ бол } \int\limits_a^b f(x)dx = 1\right)$

4. $F'(x) = f(x).$

Нягтын функцийг графикаар дүрсэлбэл:



Зураг 5.3:

Санамсаргүй хэмжигдэхүүн нягтын функцээрээ өгөгдсөн үед түүний ямар нэг завсраас утгаа авах магадлал болон тархалтын функцийг байгуулж болно.

Жишээ-5.1 Санамсаргүй хэмжигдэхүүн нягтын функцээрээ өгөгдөв:

$$f(x) = \begin{cases} 0 & x \le 0 \\ \dfrac{x}{8} & 0 < x \le 4 \\ 0 & x > 4 \end{cases}$$

1.) $P(-1 \le \xi \le 2) = ?$

2.) $F(x)$ функцийг ол

3.) $F(x), f(x)$ функцийн график байгуул.

Нягтын функцийн чанар ёсоор

1.) $P(-1 \le \xi \le 2) = \int\limits_{-1}^{2} f(x)dx = \int\limits_{-1}^{0} f(x)dx + \int\limits_{0}^{2} f(x)dx =$

$$\int\limits_{-1}^{0} 0dx + \int\limits_{0}^{2} \frac{x}{8}dx = \frac{x^2}{16}\Big|_0^2 = \frac{1}{4}.$$

2.) $F(x)$-функцийг дараах муж тус бүр дээр байгуулбал:

$x \le 0$ үед $\quad F(x) = P(\xi \le x) = \int\limits_{-\infty}^{x} f(x)dx = 0,$

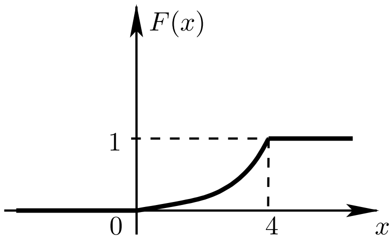
$0 < x \le 4$ үед $\quad F(x) = \int\limits_{-\infty}^{x} f(x)dx = \int\limits_{-\infty}^{0} 0dx + \int\limits_{0}^{x} f(x)dx = \int\limits_{0}^{x} \frac{x}{8}dx = \frac{x^2}{16}.$

$x > 4$ үед $\quad F(x) = \int\limits_{-\infty}^{x} f(x)dx = \int\limits_{-\infty}^{0} 0dx + \int\limits_{0}^{4} f(x)dx + \int\limits_{4}^{x} 0dx = \int\limits_{0}^{4} \frac{x}{8}dx = 1.$

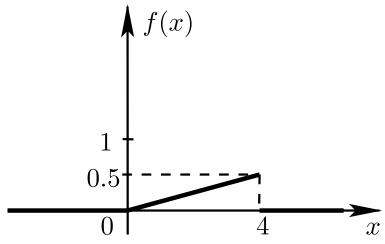
Одоо эмхэтгэж бичвэл $F(x)$ дараах хэлбэртэй болно.

$$F(x) = \begin{cases} 0 & x \le 0 \\ \dfrac{x^2}{16} & 0 < x \le 4 \\ 1 & x > 4 \end{cases}$$

3.) Тархалтын болон нягтын функцийн графикийг Зураг 5.4,5.5-д үзүүлэв.



Зураг 5.4:



Зураг 5.5:

5.3 Зарим тархалтууд

Түгээмэл хэрэглэгдэх зарим тархалтын хуультай танилцъя.
1. **Бином тархалт.** Туршилт бүрд A үзэгдэл тогтмол p магадлалтай илэрдэг гэж үзье. Тэгвэл n дараалсан туршилтанд A үзэгдлийн явагдах тоо k-нь $0, 1, 2 \ldots, n$ гэсэн боломжит утгууд бүхий дискрет санамсаргүй хэмжигдэхүүн ξ болох бөгөөд тархалтын хууль нь дараах хэлбэртэй байна.

$$P(\xi = k) = C_n^k p^k q^{n-k}, \quad \text{үүнд} \quad q = 1 - p, \quad k = 0, 1 \ldots, n.$$

2. **Паусcоны тархалт.** Туршилтын тоог хязгааргүй ихэсгэхэд $np \approx npq$ бол бином тархалт нь Пуассоны тархалт болж хувирах бөгөөд дараах хуулиар илэрхийлэгдэнэ.

$$P(\xi = k) = \frac{e^{-\lambda} \cdot \lambda^k}{k!} \quad \text{үүнд} \quad \lambda = n \cdot p, \quad k = 0, 1, 2 \ldots$$

Өөрөөр хэлбэл, туршилтын тоог хүрэлцээтэй ихэсгэхэд A үзэгдлийн явагдах магадлал p багасах боловч үзэгдэл явагдах дундаж тоо λ-нь тогтмол байдаг байна. Эндээс үзэхэд Пуассоны тархалт λ гэсэн ганцхан параметрээр бүрэн тодорхойлогдож байна. Пуассоны тархалт нийтийн үйлчилгээний онол болон үйлдвэрлэлийн технологийн тасралтгүй ажиллагааг загварчлахад өргөн хэрэглэгдэнэ.

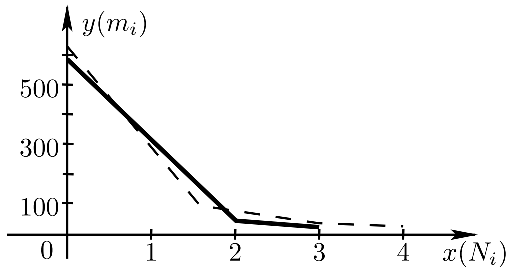
Жишээ-5.2 Ээрмэлийн үйлдвэрт 1000 ээрүүлээс 1 минутад утас тасарсан тоогоор зохиогдсон дараах хүснэгтийг авч үзье.

N_i	m_i	$\sum m_i$	$\sum N_i$
0	600	0	606
1	320	320	303
2	70	140	76
3	10	30	13
4	0	0	2
$\sum$	1000	490	1000

Үүнд N_i-тасралтын тоо
$m_i - N_i$-тасралттай ээрүүлийн тоо
$\sum m_i - m_i$-ээрүүлийн нийт тасралтын тоо
$\sum N_i$-онолын хувьд N_i тасралттай байж болох ээрүүлийн тоо

Хүснэгтээс ажиглахад 1 минутад тасрах ээрүүлийн тоо нь санамсаргүй хэмжигдэхүүн бөгөөд ээрүүлийн тоо ихэсэхэд тасралтын тоо багасаж байгаа нь харагдаж байна. Иймд тасралтын тоог Пуассоны тархалттай гэж үзэх

үндэстэй. Гэхдээ үүнийг дараах байдлаар шалгая. 1000 ээрүүлийн дундаж тасралтын тоо λ-г олбол $\lambda = \dfrac{490}{1000} = 0,49$.
0 тасралттай байх (огт тасралтгүй) үзэгдлийн магадлал нь
$P_{1000}(0) = \dfrac{e^{-0.49} \cdot 0.49^0}{0!} = 0,606$. Тэгвэл 0-тасралттай байх нийт ээрүүлийн тоо нь $1000 \cdot P_{1000}(0) = 606$. Мөн 1-тасралттай байх үзэгдлийн магадлалыг олбол $P_{1000}(1) = \dfrac{e^{-0.49} \cdot 0.49}{1!} = 0,303$. Яг адилханаар онолын хувьд 1-тасралттай байх нийт ээрүүлийн тоо нь $1000 \cdot P_{1000}(1) = 303$ гэж олдоно. Нийт N_i тасралттай ээрүүлийн тоо ба онолын хувьд тасралттай байж болох ээрүүлийн тоог тасралтын тооноос хамааруулан графикаар дүрсэлж харьцуулан үзүүлбэл:



Зураг 5.6:

Иймд ээрүүлийн утас тасрах тоог Пуассоны тархалттай гэж үзэж болно.

Жишээ-5.3 Хэсэг хугацааны туршид ээрүүлийн утас тасрах тоо нь Пуассоны хуулиар тархана. Нэг ажилчин 1000 ээрүүлэнд үйлчилдэг бөгөөд 1 минутад дунджаар 4 тасралт болдог. Тэгвэл 1 минутанд 5 удаа утас тасрах үзэгдлийн магадлалыг ол.

Нөхцөл ёсоор $n = 1000,\ p = 0.004,\ k = 5 \quad \lambda = n \cdot p = 4$. Иймд

$$P_{1000}(5) = \frac{4^5 \cdot e^{-4}}{5!} = 0,156.$$

Дээрх жишээнд Пуассоны тархалтыг өөр хэлбэрээр ашиглаж болно.
Тухайлбал, ажилчин 1000 ээрүүлэнд үйлчилдэг ба 1 минутанд дунджаар 2-тасралтыг засдаг бол 1 минутанд 3 тасралтыг арилгах магадлалыг олъё.
Үүнд $\lambda = 2,\ k = 3$. Иймд $P_{1000}(3) = \dfrac{2^1 \cdot e^{-2}}{3!} = 0.18$

Нийтийн үйлчилгээний онолд ажиглалтын тоо тогтмол байдаг бөгөөд t хугацааны турш k-удаа үйлчилгээ хийдэг гэж үзнэ. Хэрэв нэгж хугацаанд хийгдэх дундаж үйлчилгээний тоо m-мэдэгдэж байна гэж үзвэл $\lambda = m \cdot t$ болох ба t-хугацааны турш k удаа үйлчилгээ хийгдэх магадлал нь Пуассоны хуулиар илэрхийлэгдэнэ.

$$P_k(t) = \frac{(m \cdot t)^k \cdot e^{-mt}}{k!}$$

Хэрэв $k = 0$ бол энэ нь тоног төхөөрөмж t-хугацааны турш тасралтгүй ажиллах магадлал болно.

$$P_0(t) = e^{-mt}$$

Мөн тасралтгүй ажиллах хугацааны тархалтын функцийг дараах байдлаар олж болно.

$$F(t) = P(T \leq t)$$

Үүний тулд эсрэг үзэгдлийн магадлалыг олъё:

$$P(T \geq t) = 1 - P(T \leq t) = 1 - F(t).$$

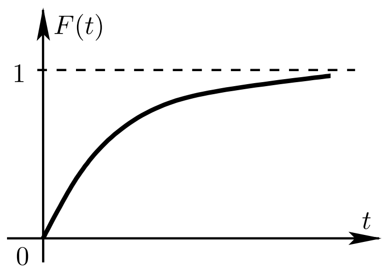
$P(T \geq t) = e^{-mt}$ гэдгийг анхаарвал $e^{-mt} = 1 - F(t)$ болох ба эндээс
$F(x) = 1 - e^{-mt}$.
3. **Илтгэгч тархалт.** Тархалтын функц нь

$$F(x) = \begin{cases} 1 - e^{-mt}, & t \geq 0, \quad \text{үед} \\ 0 & t < 0 \quad \text{үед} \end{cases}$$

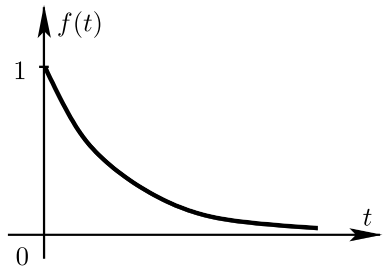
томъёогоор тодорхойлогдох санамсаргүй хэмжигдэхүүнийг илтгэгч тархалттай гэнэ. Илтгэгч тархалтын нягтын функц

$$f(x) = \begin{cases} m \cdot e^{-mt} & t \geq 0, \quad \text{үед} \\ 0 & t < 0 \quad \text{үед} \end{cases}$$

$(m > 0)$ хэлбэртэй болох ба $F(t)$, $f(t)$ функцүүдийн графикийг дараах зурагт дүрслэв.



Зураг 5.7:



Зураг 5.8:

4. **Жигд тархалт.** Тасралтгүй санамсаргүй хэмжигдэхүүн ξ-ийн бүх боломжит утгууд нь $[a, b]$ хэрчим болог...

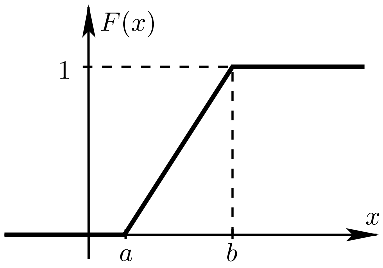
Хэрэв ξ-санамсаргүй хэмжигдэхүүний тархалтын функц

$$F(x) = \begin{cases} 0 & x < a \quad \text{үед} \\ \dfrac{x-a}{b-a} & a \leq x \leq b \quad \text{үед} \\ 1 & x > b \quad \text{үед} \end{cases}$$

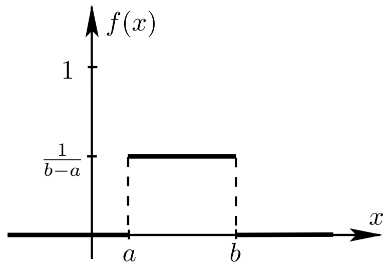
хуулиар өгөгдсөн бол түүнийг жигд тархалттай гэнэ. Тархалтын функцийг ашиглан нягтын функцийг тодорхойлбол:

$$f(x) = \begin{cases} 0 & x < a \quad \text{үед} \\ \dfrac{1}{b-a} & a \le x \le b \quad \text{үед} \\ 0 & x > b \quad \text{үед} \end{cases}$$

$F(x)$, $f(x)$ функцүүдийн графикийг дүрсэлбэл:



Зураг 5.9:

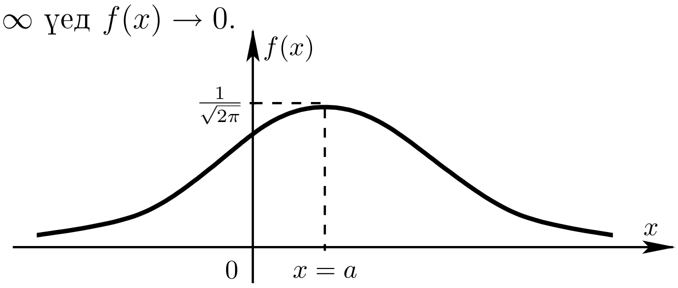


Зураг 5.10:

5. **Хэвийн (нормаль) тархалт.** Хэрэв тасралтгүй санамсаргүй хэмжигдэхүүний нягтын функц

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-a)^2/2\sigma^2}$$

хэлбэртэй бол түүнийг хэвийн тархалттай гэнэ. $f(x)$-функцийн график нь $x = a$ босоо шулууны хувьд тэгш хэмтэй, $x = a$ цэг дээр хамгийн их утгаа авах ба $x \to \pm\infty$ үед $f(x) \to 0$.



Зураг 5.11:

Хэвийн тархалттай хэмжигдэхүүний тархалтын функц дараах хэлбэртэй байна.

$$F(x) = \int_{-\infty}^{x} f(x)dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-(x-a)^2}{2\sigma^2}} dx = 0.5 + \Phi\left(\frac{x-a}{\sigma}\right).$$

Санамсаргүй хэмжигдэхүүн $[x_1, x_2]$ завсраас утгаа авах магадлалыг $\Phi(x) = \dfrac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-t^2/2} dt$ гэсэн Лапласын функц ашиглан олбол:

$$P(x_1 \le \xi \le x_2) = F(x_2) - F(x_1) = \Phi\left(\frac{x_2-a}{\sigma}\right) - \Phi\left(\frac{x_1-a}{\sigma}\right).$$

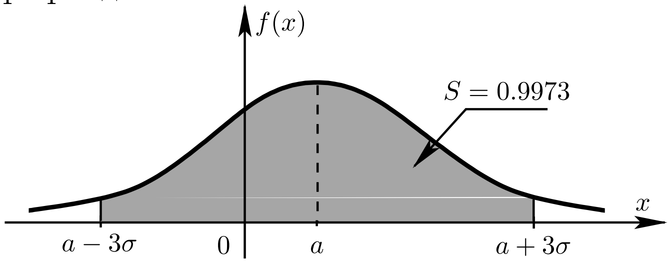
Энэ үр дүнг ашиглан "гурван σ-ийн дүрмийг" гаргая.

$$P(|\xi - a| < \sigma) = P(a - \sigma < \xi < a + \sigma) = 2\Phi(1) = 0.6826$$

$$P(|\xi - a| < 2\sigma) = P(a - 2\sigma < \xi < a + 2\sigma) = 2\Phi(2) = 0.9544$$

$$P(|\xi - a| < 3\sigma) = P(a - 3\sigma < \sigma < a + 3\sigma) = 2\Phi(3) = 0.9973$$

Эндээс харахад хэвийн тархалттай санамсаргүй хэмжигдэхүүн математик дунджаасаа хазайх хазайлтын абсолют хэмжигдэхүүн 3σ-аас их байх нь бараг боломжгүй үзэгдэл байна.



Зураг 5.12:

Жишээ-5.4 Нэгэн бүс нутгийн эрчүүдийн өндөр нь $a = 170.2$ см ба $\sigma = 5.6$ см параметрүүдтэй хэвийн тархсан санамсаргүй хэмжигдэхүүн бол 166-172 см өндөртэй хүмүүс нийт эрчүүдийн хэдэн хувийг эзлэх вэ?

Бодлогын нөхцөл ёсоор $x_2 = 172, \quad x_1 = 166$ тул
$P(166 < \xi < 172)$ магадлалыг олох ёстой. Муавр-Лапласын интеграл томъёог ашиглавал:
$P(166 < \xi < 172) = \Phi\left(\dfrac{172 - 170.2}{5.6}\right) - \Phi\left(\dfrac{166 - 170.2}{5.6}\right) =$
$\Phi(0.32) - \Phi(-0.75) = 0.3999$ (буюу $\approx 40\%$) Эндээс 166-172 см өндөртэй эрчүүд нь дунджаар 40%-ийг эзлэж байна.

Жишээ-5.5 Нэгэн бүс нутгийн үндэстний эмэгтэйчүүдийн цээжний тойргийн урт нь $a = 96.8$ см ба $\sigma = 4.5$ см параметрүүд бүхий хэвийн тархсан санамсаргүй хэмжигдэхүүн байв. Хэрвээ оёдлын үйлдвэр эмэгтэйчүүдийн пальто үйлдвэрлэдэг бол пальто үйлдвэрлэлийн размерүүдийн харьцааг тогтоо.

Тухайлбал, 42 размерийн пальтог хэдэн хувь үйлдвэрлэхийг олохын тулд дараах магадлалыг олох шаардлагатай.
42-размерийн хувьд $W(42) = P(82 < \xi < 86)$,
44-размерийн хувьд $W(44) = P(86 < \xi < 90)$ г.м.
$P(82 < \xi < 86) = \Phi\left(\dfrac{86 - 96.8}{4.5}\right) - \Phi\left(\dfrac{82 - 96.8}{4.5}\right) =$
$\Phi(-2.40) - \Phi(-3.29) = 0.008$ (буюу 0.8%),
$P(86 < \xi < 90) = \Phi\left(\dfrac{90 - 96.8}{4.5}\right) - \Phi\left(\dfrac{86 - 96.8}{4.5}\right) = 0.057$
буюу 0.57%. Бусад бүх размерүүдийн хувьд дээрхийн адилаар олж үр дүнг

дараах хүснэгтэнд харуулав.

размер	$[x_1, x_2]$	t_1	t_2	$\Phi(t_2)-\Phi(t_1)$	үйлд-вэрлэх хувь (%)
42	[82,86]	-3.29	-2.40	0.008	0.8%
44	[86,90]	-2.40	-1.51	0.057	5.7%
46	[90,94]	-1.51	-0.62	0.202	20.2%
48	[94,98]	-0.62	0.27	0.339	33.9%
50	[98,102]	0.27	1.16	0.271	27.1%
52	[102,106]	1.16	2.04	0.102	10.2%
54	[106,110]	2.04	2.94	0.019	1.9%

Бүлэг 6

Санамсаргүй хэмжигдэхүүний тоон үзүүлэлтүүд

Санамсаргүй хэмжигдэхүүний тархалтын хууль нь уг хэмжигдэхүүнийг бүрэн тодорхойлох боловч практикт санамсаргүй хэмжигдэхүүний утгууд ба тэдгээр утгаа хүлээн авах магадлалыг нэг бүрчлэн мэдэхэд хүндрэлтэй төдийгүй, заримдаа боломжгүй байдаг. Харин санамсаргүй хэмжигдэхүүний тархалтын онцлогийг нилээд товч, ерөнхий хэлбэрээр илэрхийлсэн тоонуудыг мэдэх шаардлага гардаг. Ийм тоонуудыг санамсаргүй хэмжигдэхүүний тоон үзүүлэлтүүд гэж нэрлэнэ. Эдгээр нь санамсаргүй хэмжигдэхүүний математик дундаж, дисперс, янз бүрийн эрэмбийн моментууд, моод, медиан, асимметрийн болон эксцессийн коэффициентууд юм.

6.1 Математик дундаж түүний чанарууд

Дискрет санамсаргүй хэнмжигдэхүүн ξ-ийн бүх боломжит утгуудыг харгалзах магадлалаар нь үржүүлж нэмсэн нийлбэрийг түүний математик дундаж гэнэ. $M(\xi)$, $M[\xi]$, m_ξ, $E(\xi)$ гэх мэтчилэн тэмдэглэнэ.

$\xi_i:$	x_1	x_2	$\cdots$	x_n
$p_i:$	p_1	p_2	$\cdots$	p_n

$$M(\xi) = \sum_{i=1}^{n} x_i p_i \tag{6.1}$$

Хэрэв ξ нь $f(x)$ гэсэн нягттай тасралтгүй санамсаргүй хэмжигдэхүүн бол математик дунджийг дараах томъёогоор олно.

$$M(\xi) = \int_{-\infty}^{+\infty} x \cdot f(x) dx \tag{6.2}$$

Тухайн тохиолдолд ξ нь боломжит утгуудаа $[a, b]$ завсраас авдаг бол

$$M(\xi) = \int_a^b x \cdot f(x)dx \tag{6.3}$$

Математик дунджийн үндсэн чанаруудыг тэмдэглэе.

1. Тогтмол хэмжигдэхүүний мат.дундаж өөртэйгөө тэнцүү: $M(C) = C$.

2. Тогтмол тоог мат.дунджийн тэмдгийн өмнө гаргаж болно: $M(C \cdot \xi) = C \cdot M(\xi)$.

3. Санамсаргүй хэмжигдэхүүнүүдийн нийлбэрийн мат.дундаж тэдгээрийн мат.дунджуудын нийлбэртэй тэнцүү: $M\left(\sum\limits_{i=1}^{n} \xi_i\right) = \sum\limits_{i=1}^{n} M(\xi_i)$

4. Хэрэв $\eta = \sum\limits_{i=1}^{n} a_i\xi_i + b$ бол $M(\eta) = \sum\limits_{i=1}^{n} a_i M(\xi_i) + b$.

5. Хамааралгүй санамсаргүй хэмжигдэхүүнүүдийн үржвэрийн мат.дундаж нь тэдгээрийн мат.дунджуудын үржвэртэй тэнцүү:

$$M(\xi_1 \cdot \xi_2 \cdots \xi_n) = M(\xi_1) \cdot M(\xi_2) \cdots M(\xi_n).$$

Жишээ-6.1 10 м даавууны "утас тасрах тоо"-ны тархалтын хууль дараах хүснэгтээр өгөгдөв.

ξ :	0	1	2	3
p_i :	0.2	0.4	0.3	0.1

Энэ санамсаргүй хэмжигдэхүүний математик дунджийг ол.

(6.1) томъёо ёсоор $M(\xi) = 0 \cdot 0.2 + 1 \cdot 0.4 + 2 \cdot 0.3 + 3 \cdot 3 \cdot 0.1 = 1.3$

Энэ нь 10 м даавууны утсанд дунджаар 1,3 тасралт ноогдоно гэсэн үг юм.

Жишээ-6.2 Тасралтгүй санамсаргүй хэмжигдэхүүн тархалтын функцээрээ өгөгдөв.

$$F(x) = \begin{cases} 0 & x \le 2 \\ \frac{1}{8}(x-1)^2 - \frac{1}{8} & 2 < x \le 4 \\ 1 & x > 4 \end{cases}$$

Санамсаргүй хэмжигдэхүүний математик дунджийг ол.

Эхлээд нягтын функцийг олъё ($f(x) = F'(x)$).

$$f(x) = \begin{cases} 0 & x \le 2 \\ \frac{1}{4}(x-1) & 2 < x \le 4 \\ 0 & x > 4 \end{cases}$$

$$M\xi = \int_2^4 x f(x)dx = \int_2^4 x\frac{1}{4}(x-1)dx = \frac{1}{4}\left(\frac{x^3}{3} - \frac{x^2}{2}\right)\Bigg|_2^4 = \frac{19}{6}.$$

6.2 Дисперс, дисперсийн чанарууд

Санамсаргүй хэмжигдэхүүн математик дунджаасаа хазайх хазайлтын квадратын математик дунджийг дисперс гэж нэрлэнэ. $D(\xi)$, $D[\xi]$, σ_{ξ}^{2}, $V(\xi)$ гэх мэтчилэн тэмдэглэнэ.

$$D(\xi) = M[\xi - M(\xi)]^2 \tag{6.4}$$

$[\xi - M(\xi)]^2$ санамсаргүй хэмжигдэхүүний боломжит утгууд нь $[x_1 - M(\xi)]^2$, $[x_2 - M(\xi)]^2, \ldots, [x_n - M(\xi)]^2$ бөгөөд эдгээр утгуудыг авах магадлалууд харгалзан $p_1, p_2, \ldots, p_n$ бол дисперс нь $D(\xi) = \sum_{i=1}^{n} [x_i - M(\xi)]^2 p_i$ гэж бодогдоно.
(6.4) томъёотой тэнцүү чанартай дараах томъёог хэрэглэх нь тохиромжтой байдаг.

$$D(\xi) = M(\xi^2) - M^2(\xi) \tag{6.5}$$

Тасралтгүй санамсаргүй хэмжигдэхүүний дисперсийг:

$$D(\xi)=\int_{-\infty}^{+\infty} [x-M(\xi)]^2 f(x)dx \quad \left(D(\xi)=\int_{-\infty}^{+\infty} x^2 f(x)dx-(M\xi)^2 \right) \tag{6.6}$$

томъёогоор олно.
Санамсаргүй хэмжигдэхүүний дисперсээс язгуур гаргасныг уг санамсаргүй хэмжигдэхүүний дундаж квадрат хазайлт буюу стандарт хазайлт ($\sigma(\xi)$, σ_{ξ}) гэж нэрлэнэ.

$$\sigma(\xi) = \sqrt{D(\xi)} \tag{6.7}$$

Одоо дисперсийн чанаруудаас дурдья.
1. Тогтмол хэмжигдэхүүний дисперс тэгтэй тэнцүү: $D(C) = 0$.
2. Тогтмол тоог дисперсийн тэмдгийн өмнө квадрат зэрэг дэвшүүлж гаргана: $D(C \cdot \xi) = C^2 \cdot D(\xi)$
3. $K(\xi;\eta) = M[(\xi - M\xi) \cdot M(\eta - M\eta)] = M(\xi \cdot \eta) - M\xi \cdot M\eta$ хэмжигдэхүүнийг ξ, η санамсаргүй хэмжигдэхүүнүүдийн корреляцийн момент (ковариац) гэнэ.

$$\forall\ \xi, \eta - \text{ийн хувьд} \quad D(\xi+\eta)=D\xi+D\eta+2\cdot K(\xi;\eta) \quad \text{байна.}$$

4. $\xi_1,\ \xi_2, \ldots, \xi_n$−хамааралгүй бол

$$D(\xi_1+\xi_2+\ldots+\xi_n)=D\xi_1+D\xi_2+\ldots+D\xi_n.$$

Жишээ-6.3 Жишээ-6.1-д авч үзсэн дараах

$\xi:$	0	1	2	3
$p_i:$	0.2	0.4	0.3	0.1

санамсаргүй хэмжигдэхүүний дисперсийг ол.
Дисперсийг 2 янзаар олж үзүүлье.
I **арга:** Дисперсийн тодорхойлолт ашиглан олъё.

Энэ тохиолдолд $(\xi - M\xi)^2$ хэмжигдэхүүний тархалтын хуулийг байгуулах шаардлагатай:

$(\xi - M\xi)^2$:	$(0-1.3)^2$	$(1-1.3)^2$	$(2-1.3)^2$	$(3-1.3)^2$
p_i:	0.2	0.4	0.3	0.1

буюу

$(\xi - M\xi)^2$:	1.69	0.09	0.49	2.89
p_i:	0.2	0.4	0.3	0.1

Иймд (6.4) томъёо ёсоор:
$D(\xi) = M[(\xi - M\xi)^2] = 1.69\cdot0.2 + 0.09\cdot0.4 + 0.49\cdot0.3 + 2.89\cdot0.1 = 0.81,\ \sigma(\xi) = 0,9.$
II **арга:** (6.5) томъёог ашиглая. Энэ тохиолдолд ξ^2 хэмжигдэхүүний тархалтын хуулийг байгуулах шаардлагатай:

ξ^2 :	0^2	1^2	2^2	3^2
p_i:	0.2	0.4	0.3	0.1

Иймд $M(\xi^2) = 0\cdot0.2 + 1\cdot0.4 + 4\cdot0.3 + 9\cdot0.1 = 2,5.$
(6.5) томъёо ёсоор $D(\xi) = M(\xi^2) - (M\xi)^2 = 2.5^2 - 1.3^2 = 0.81,\quad \sigma(\xi) = 0,9.$
Одоо Жишээ-6.2-д авч үзсэн тасралтгүй санамсаргүй хэмжигдэхүүний дисперс ба стандарт хазайлтыг бодъё:

$$D(\xi) = \int_a^b x^2 f(x)dx - (M\xi)^2 = \int_2^4 x^2 \frac{1}{4}(x-1)dx - \left(\frac{19}{6}\right)^2 =$$

$$= \frac{1}{4}\left(\frac{x^4}{4} - \frac{x^3}{3}\right)\Big|_2^4 - \left(\frac{19}{6}\right)^2 = \frac{11}{36},\quad \sigma(\xi) = \frac{\sqrt{11}}{6}.$$

6.3 Зарим тархалтын тоон үзүүлэлтүүд

Одоо V бүлгийн § 3-д авч үзсэн тархалтын хуулиудын тоон үзүүлэлтийг олъё.

1. ξ нь Бином тархалттай санамсаргүй хэмжигдэхүүн байг.

$$\xi_i = \begin{cases} 1 & \text{Хэрэв}\quad i-\text{р туршилтанд } A \text{ үзэгдэл явагдвал} \\ 0 & \text{Хэрэв}\quad i-\text{р туршилтанд } A \text{ үзэгдэл эс явагдвал} \end{cases}$$

гэсэн санамсаргүй хэмжигдэхүүнийг авъя. Тэгвэл
$\xi = \xi_1 + \xi_2 + \cdots + \xi_n = \sum\limits_{i=1}^{n} \xi_i$ болно. Бернуллийн схем учир ξ_i бүр хамааралгүй.
Иймд
$M(\xi) = M\left(\sum\limits_{i=1}^{n} \xi_i\right) = \sum\limits_{i=1}^{n} M\xi_i = \sum(1\cdot p + 0\cdot q) = np,$
$D(\xi) = D\left(\sum\limits_{i=1}^{n} \xi_i\right) = \sum\limits_{i=1}^{n} D(\xi_i),\quad D(\xi_i) = M(\xi_i^2) - (M\xi_i)^2 =$

$1^2 \cdot p + 0^2 \cdot q - p^2 = p - p^2 = p(1-p) = pq, \quad D(\xi) = \sum\limits_{i=1}^{n} D(\xi_i) = \sum\limits_{i=1}^{n} pq = npq.$
Дүгнэвэл, бином тархалттай санамсаргүй хэмжигдэхүүний хувьд

$$M(\xi) = np, \quad D(\xi) = npq, \quad \sigma(\xi) = \sqrt{npq} \tag{6.8}$$

2. Хэрэв ξ нь Пуассоны тархалттай санамсаргүй хэмжигдэхүүн бол түүний математик дундаж

$$M(\xi) = \sum_{k=0}^{\infty} k \cdot \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \cdot \lambda \cdot \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = e^{-\lambda} \lambda \cdot e^{\lambda} = \lambda.$$

Одоо дисперсийг бодъё. Өмнөх үр дүнгээ ашиглавал:
$\lambda = \sum\limits_{k=0}^{\infty} k \dfrac{\lambda^k}{k!} e^{-\lambda} \implies \lambda e^{\lambda} = \sum\limits_{k=0}^{\infty} k \dfrac{\lambda^k}{k!}$ Тэнцэтгэлийн 2 талыг λ-аар дифференциалчилбал:

$$e^{\lambda} + \lambda e^{\lambda} = \sum_{k=0}^{\infty} k \frac{k \lambda^{k-1}}{k!} = \sum_{k=0}^{\infty} k^2 \frac{\lambda^{k-1}}{k!}.$$

$\lambda e^{-\lambda}$—аар үржүүлсний дараа

$$\lambda + \lambda^2 = \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} = M\xi^2, \quad D\xi = M\xi^2 - \lambda^2 = \lambda + \lambda^2 - \lambda^2 = \lambda$$

Иймд $D(\xi) = \lambda$. Дүгнэвэл

$$M(\xi) = \lambda, \; D(\xi) = \lambda, \; \sigma(\xi) = \sqrt{\lambda}. \tag{6.9}$$

Пуассоны тархалт нийтийн үйлчилгээний онолд чухал үүрэг гүйцэтгэх ба **үзэгдлүүдийн** нэгж хугацаан дахь **энгийн урсгал**ын тоо нь Пуассоны тархалттай байна. Хугацааны ямар нэг санамсаргүй агшинд хойно хойноосоо цувран явагдах ижил төрлийн үзэгдлүүдийн дараалалыг үзэгдлүүдийн урсгал гэх ба дараах 3 нөхцлийг хангах урсгалыг **энгийн** гэнэ. Үүнд:
1. урсгалын магадлалт үзүүлэлтүүд нь хугацаанаас үл хамаарах (стационар),
2. хугацааны тухайн агшинд үзэгдлүүд хэсэг бүлгээрээ биш зөвхөн нэг нэгээрээ дараалан явагдах (ординар),
3. хоорондоо үл огтлолцох хугацааны дурын 2 интервалын хувьд нэгд нь илрэх үзэгдлийн тоо нөгөөд нь илрэх үзэгдлийн тооноос үл хамаарах (дараа гийнхдаа үл нөлөөлөх)

Үзэгдлүүдийн энгийн урсгалыг Пуассоны урсгал буюу **Пуассоны процесс** гэж нэрлэдэг.
t хугацаанд Пуассоны процессийн k үзэгдэл илрэх магадлал дараах томъёогоор тодорхойлогдоно:

$$P_t(k) = \frac{(\lambda t)^k}{k!} \cdot e^{-\lambda t} \tag{6.10}$$

Үүнд, λ нь урсгалын эрчим буюу нэгж хугацаанд илрэх үзэгдлүүдийн дундаж тоо.
3. $\xi-$Геометр тархалттай санамсаргүй хэмжигдэхүүн байг. Боломжит утгуудын магадлал нь

$$P(\xi=k)=q^{k-1}\cdot p=(1-p)^{k-1}\cdot\ p,\ \ k=1,2,\ldots$$

томъёогоор олдох $\xi-$санамсаргүй хэмжигдэхүүнийг геометр тархалттай гэнэ. Геометр тархалтын тоон үзүүлэлтүүд:

$$M(\xi)=1/p\ ,\quad D(\xi)=q/p^2=(1-p)/p^2. \tag{6.11}$$

Бернуллийн схемийн нөхцөлд A үзэгдлийг явагдах хүртэл хийсэн туршилтын тоо геометр тархалттай байна.
4. $\xi-$Гипргеометр тархалттай санамсаргүй хэмжигдэхүүн байг. Боломжит утгуудын магадлал нь

$$P(\xi=m)=\frac{C_M^m\cdot C_{N-M}^{n-m}}{C_N^n}$$

Үүнд, $m=0,1,2,\ldots,min(m,M),\ m\leq N,\ n\leq N$ томъёогоор тодорхойлогдох $\xi-$санамсаргүй хэмжигдэхүүнийг гипргеометр тархалттай гэнэ.
Гипргеометр тархалтын тоон үзүүлэлтүүд:

$$M(\xi)=n\cdot\frac{M}{N},\qquad D(\xi)=\frac{nm}{N-1}\left(1-\frac{M}{N}\right)\cdot\left(1-\frac{n}{N}\right). \tag{6.12}$$

Гипргеометр тархалт нь хүлээн авах статистикт өргөн хэрэглэгдэнэ.
5. ξ нь илтгэгч тархалттай санамсаргүй хэмжигдэхүүн бол

$$M(\xi)=\int\limits_{-\infty}^{\infty}x\cdot f(x)dx=\int\limits_{0}^{+\infty}t\cdot me^{-mt}dt=m\cdot\int\limits_{0}^{+\infty}te^{-mt}dt=\frac{1}{m},$$

$$D(\xi)=M(\xi^2)-[M(\xi)]^2=\int\limits_{0}^{+\infty}t^2me^{-mt}dt-\left(\frac{1}{m}\right)^2=$$

$=\dfrac{2}{m^2}-\dfrac{1}{m^2}=\dfrac{1}{m^2},\quad \sigma(\xi)=\sqrt{D(\xi)}=\dfrac{1}{m}.$ Дүгнэвэл:

$$M(\xi)=1/m\ ,\quad D(\xi)=1/m^2,\quad \sigma(\xi)=1/m. \tag{6.13}$$

6. Хэрэв ξ нь $[a,b]$ хэрчим дээр жигд тархсан санамсаргүй хэмжигдэхүүн бол

$$M(\xi)=\int\limits_{-\infty}^{\infty}x\cdot f(x)dx=\int\limits_{a}^{b}x\cdot\frac{dx}{b-a}=\frac{a+b}{2},$$

$$D(\xi)=\int_{-\infty}^{+\infty}(x-M\xi)^2 f(x)dx=\int_a^b\left(x-\frac{a+b}{2}\right)^2\frac{dx}{b-a}=\frac{(b-a)^2}{12},$$

$$\sigma(\xi)=\sqrt{\frac{(b-a)^2}{12}}=\frac{(b-a)}{2\sqrt{3}}=\frac{\sqrt{3}}{6}(b-a).\ \text{Дүгнэвэл:}$$

$$M(\xi)=\frac{a+b}{2}\ ,\quad D(\xi)=\frac{(b-a)^2}{12},\quad \sigma(\xi)=\frac{(b-a)}{2\sqrt{3}}. \tag{6.14}$$

Илтгэгч тархалттай санамсаргүй хэмжигдэхүүн нь Пуассоны процессийн үед ((6.10) томъёог үз!) үзэгдэл анх удаа явагдах хүртлэх хугацааны интервалын уртыг тодорхойлно. Ерөнхий тохиолдолд, Пуассоны процесст яг r үзэгдэл явагдах хүртлэх хугацааны интервалын уртыг илэрхийлсэн хэмжигдэхүүнийг **Эрлангийн тархалт**тай санамсаргүй хэмжигдэхүүн гэнэ. Эрлангийн тархалтын нягт нь

$$f(x)=\frac{\lambda^r\cdot x^{r-1}}{(r-1)!}\cdot e^{-\lambda x},\quad x>0,\quad r=1,2,\ldots \tag{6.15}$$

томъёогоор тодорхойлогдоно. Тухайн тохиолдолд, $r=1$ үед Эрлангийн тархалт нь илтгэгч тархалт болно. Эрлангийн тархалтын тоон үзүүлэлтүүд:

$$M(\xi)=r/\lambda,\ D(\xi)=r/\lambda^2. \tag{6.16}$$

7. Хэрэв ξ нь a, σ^2 параметрүүд бүхий хэвийн тархалттай санамсаргүй хэмжигдэхүүн бол (товчоор $N(a,\sigma^2)$ гэж тэмдэглэнэ)

$$M(\xi)=\int_{-\infty}^{+\infty}x\cdot\frac{1}{\sqrt{2\pi}\sigma}exp\left[-\frac{(x-a)^2}{2\sigma^2}\right]dx=a,$$

$$D(\xi)=\int_{-\infty}^{+\infty}(x-a)^2\cdot\frac{1}{\sqrt{2\pi}\sigma}exp\left[-\frac{(x-a)^2}{2\sigma^2}\right]dx=\sigma^2,\ \sigma(\xi)=\sigma.$$

$$M(\xi)=a,\ D(\xi)=\lambda^2,\ \sigma(\xi)=\sigma. \tag{6.17}$$

Тухайлбал, $f(x)=\dfrac{1}{0.1\cdot\sqrt{2\pi}}\cdot exp\left[-\dfrac{(x-7)^2}{23\cdot(0.1)^2}\right]$ гэсэн тархалтын нягт бүхий хэвийн тархалттай санамсаргүй хэмжигдэхүүний параметрүүд нь
$M(\xi)=7,\ \ D(\xi)=(0.1)^2,\ \ \sigma(\xi)=0.1$ болно.
Хэрэв энэ санамсаргүй хэмжигдэхүүний хувьд 3σ-ийн дүрмийг хэрэглэвэл $P(|\xi-7|<3\cdot 0.1)=P(6,7\le\xi\le 7.3)\approx 0.9973$ буюу ξ санамсаргүй хэмжигдэхүүн $[6.7;7.3]$ хэрчмээс утгаа авах нь практикт бараг гарцаагүй үзэгдэл болно.

6.4 Дээд эрэмбийн моментууд ба бусад тоон үзүүлэлтүүд

k **эрэмбийн анхны момент:** $\nu_k = M(\xi^k)$ хэмжигдэхүүнийг ξ санамсаргүй хэмжигдэхүүний k эрэмбийн анхны момент гэнэ. Дискрет санамсаргүй хэмжигдэхүүний хувьд

$$\nu_k = \sum_{i=1}^{n} x_i^k \cdot p_k, \quad (k = 1, 2, \ldots) \tag{6.18}$$

томъёогоор бодох ба хэрэв ξ тасралтгүй санамсаргүй хэмжигдэхүүн бол

$$\nu_k = \int_{-\infty}^{+\infty} x^k \cdot f(x)dx \tag{6.19}$$

Тухайн тохиолдолд $\nu_1 = M(\xi), \quad \nu_2 = M(\xi^2)$ болох тул

$$D(\xi) = M(\xi^2) - (M\xi)^2 = \nu_2 - \nu_1^2.$$

k **эрэмбийн төвийн момент:** $\mu_k = M[(\xi - M\xi)^k]$ хэмжигдэхүүнийг ξ санамсаргүй хэмжигдэхүүний k эрэмбийн төвийн момент гэнэ. μ_k нь дискрет санамсаргүй хэмжигдэхүүний хувьд:

$$\mu_k = \sum_{i=1}^{n} [x_i - M(\xi)]^k p_i, \quad k = 1, 2, \ldots \tag{6.20}$$

тасралтгүй санамсаргүй хэмжигдэхүүний хувьд:

$$\mu_k = \int_{\infty}^{\infty} [x - M(\xi)]^k \cdot f(x)dx \tag{6.21}$$

томъёогоор тус тус олдоно.

Тухайн тохиолдолд $\mu_1 = M(\xi - M\xi) = 0, \ \mu_2 = M(\xi - M\xi)^2 = D(\xi)$,
$\mu_3 = \nu_3 - 3\nu_1 - \nu_2 + 2\nu_1^3, \quad \mu_4 = \nu_4 - 4\nu_3\nu_1 + 6\nu_2\nu_1 - 3\nu_1^4$ байхыг харуулж болно.

Санамсаргүй хэмжигдэхүүний моод: Дискрет санамсаргүй хэмжигдэхүүний хамгийн их магадлалтай утгыг моод гэнэ. Тасралтгүй санамсаргүй хэмжигдэхүүний хувьд нягтын функц нь хамгийн их утгаа авах цэгийг түүний моод гэнэ. Өөрөөр хэлбэл, геометр утгаараа тархалтын муруйн максимумын абсцисс юм. 2 ба олон моодтой тархалт байхаас гадна максимумгүй мөртлөө минимумтай тархалт оршин байна.

Санамсаргүй хэмжигдэхүүний медиан:
$P(\xi < m_e) = P(\xi > m_e)$ нөхцлийг хангах m_e утгыг ξ санамсаргүй хэмжигдэхүүний медиан гэнэ. Геометр утгаараа, тархалтын муруйгаар хязгаарлагдсан дүрсийн талбайг хагаслан хуваах шулуун дээр орших цэгийн абсцисс юм.

Дээрх талбайнууд тэнцүү учраас $P(\xi < m_e) = P(\xi > m_e) = F(m_e) = 0.5$. Нэг моодтой бөгөөд тэгш хэмтэй тархалтын хувьд моод ба медиан хоёр давхцана. Тухайлбал, хэвийн тархалтын моод ба медиан нь a байна.

Асимметрийн болон эксцессийн коэффициент:

Хэрэв тархалтын муруй $x = M(\xi)$ шулууны хувьд тэгш хэмтэй байвал уг тархалтыг тэгш хэмтэй гэнэ. Тэгш хэмтэй биш тархалтын хувьд түүний тэгш хэмт тархалтаас ялгагдах ялгааг дараах 2 үзүүлэлтээр тодорхойлно.

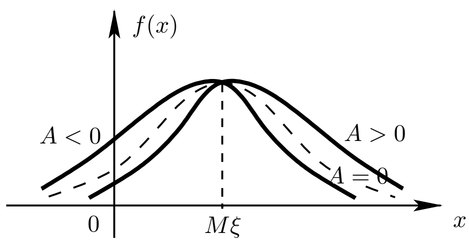
$$A = \frac{\mu_3}{\sigma^3} \tag{6.22}$$

хэмжигдэхүүнийг тархалтын асимметрийн (тэгш хэмийн) коэффициент гэнэ. Тархалтын муруйн ихэнх хэсэг дунджийнхаа аль талд оршиж буйг энэ хэмжигдэхүүн илэрхийлнэ.

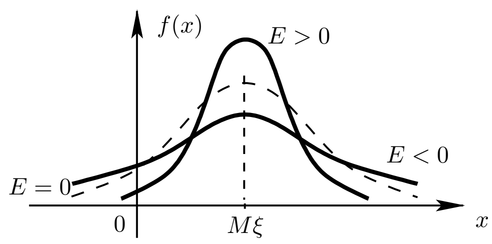
$$E = \frac{\mu_4}{\sigma^4} - 3 \tag{6.23}$$

хэмжигдэхүүнийг эксцессийн коэффициент гэнэ. Эксцессийн коэффициент тархалтын муруйн огцомыг илэрхийлнэ.

Хэвийн тархалтын хувьд $A = E = 0$ байна. Ижилхэн матеаматик дундаж, дисперстэй боловч тэгш хэмийн ба эксцессийн коэффициентууд нь өөр өөр байх санамсаргүй хэмжигдэхүүний нягтын функцийг дараах хоёр зургаар харуулав.(Зураг 6.1, Зураг 6.2)



Зураг 6.1:



Зураг 6.2:

Бүлэг 7

Хоёр хэмжээст санамсаргүй хэмжигдэхүүн

7.1 Хоёр хэмжээст санамсаргүй хэмжигдэхүүн түүний тархалтын функц

Туршилтын дүнд 2 болон түүнээс олон санамсаргүй хэмжигдэхүүн илэрч болно. Иймд хоёр буюу хэд хэдэн санамсаргүй хэмжигдэхүүнийг зэрэг авч үзэх шаардлага гардаг.

Хэрэв ξ_1 ба ξ_2 санамсаргүй хэмжигдэхүүний хувьд $\xi_1 < x, \;\; \xi_2 < y$ байх үзэгдлүүд зэрэг явагдах магадлал $P(\xi_1 < x, \;\; \xi_2 < y)$ тодорхойлогдсон бол ξ_1, ξ_2 санамсаргүй хэмжигдэхүүний эрэмбэлэгдсэн хосыг хоёр хэмжээст санамсаргүй хэмжигдэхүүн гэнэ. Дурын x ба y-ийн хувьд тодорхойлогдсон

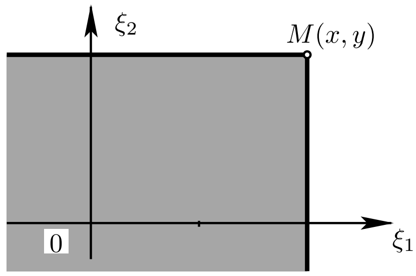
$$F(x, y) = P(\xi_1 < x, \; \xi_2 < y) \tag{7.1}$$

функцийг хоёр хэмжээст санамсаргүй хэмжигдэхүүний тархалтын функц буюу хамтын тархалтын функц гэнэ.

Хоёр хэмжээст санамсаргүй хэмжигдэхүүний тархалтын функцийн геометр утга нь: санамсаргүй цэг, $M(x, y)$ цэгт оройтой түүнээс зүүн доод хэсэгт орших төгсгөлгүй мужид (Зураг 7.1) унах магадлалыг илэрхийлнэ. Тархалтын функцийн тодорхойлолтоос

1. $0 \le F(x, y) \le 1$.
2. $F(x, y)$-хувьсагч бүрээрээ үл буурах
3. $\lim\limits_{x \to -\infty} F(x, y) = F(-\infty; y) = 0$

   $\lim\limits_{y \to -\infty} F(x, y) = F(x, -\infty) = 0.$



Зураг 7.1:

$$\lim_{\substack{x \to +\infty \\ y \to +\infty}} F(x,y) = 1, \text{ өөрөөр хэлбэл: } F(+\infty; +\infty) = 1.$$

$$4.)\ P(a \le \xi_1 \le b;\quad c \le \xi_2 \le d) = F(b;d) - F(a;d) - F(b;c) + F(a;c). \tag{7.2}$$

чанаруудтай болох нь мөрднө.

Хэрэв ξ_1 ба ξ_2 тус бүр нь дискрет санамсаргүй хэмжигдэхүүнүүд бол (ξ_1, ξ_2) нь дискрет санамсаргүй хэмжигдэхүүний систем болно. Тухайлбал ξ_1-ийн боломжит утгууд нь $x_1, x_2, \dots x_n$; ξ_2-ийн боломжит утгууд нь $y_1, y_2, \dots y_m$ бол хоёр хэмжээст санамсаргүй хэмжигдэхүүн $(\xi_1; \xi_2)$-ийн боломжит утгууд нь $(x_i; y_j)$ $(i = \overline{1,n},\ j = \overline{1,m})$ хосууд болно. Хоёр хэмжээст санамсаргүй хэмжигдэхүүн (x_i, y_j) хос утгаа авах магадлалыг p_{ij}-гэж тэмдэглэе. Өөрөөр хэлбэл,

$$p_{ij} = P(\xi_1 = x_i,\ \xi_2 = y_j), \quad (i = \overline{1,n},\ j = \overline{1,m}) \tag{7.3}$$

Дискрет санамсаргүй хэмжигдэхүүний системийн тархалтын хуулийг дараах хэлбэрийн хүснэгтээр илэрхийлнэ.

$\xi_1 \setminus \xi_2$	y_1	y_2	$\dots$	y_j	$\dots$	y_m	p_i
x_1	p_{11}	p_{12}	$\dots$	p_{1j}	$\dots$	p_{1m}	p_1
x_2	p_{21}	p_{22}	$\dots$	p_{2j}	$\dots$	p_{2m}	p_2
$\vdots$							
x_i	p_{i1}	p_{i2}	$\dots$	p_{ij}	$\dots$	p_{im}	p_i
$\vdots$			$\dots$		$\dots$		
x_n	p_{n1}	p_{n2}	$\dots$	p_{nj}	$\dots$	p_{nm}	p_n
q_j	q_1	q_2	$\dots$	q_j	$\dots$	q_m	$\sum p_i = \sum q_j = 1$

Үүнд

$$\sum_{j=1}^{m} p_{ij} = p_i \quad (i = \overline{1,n}) \tag{7.4}$$

$$\sum_{i=1}^{n} p_{ij} = q_j \quad (j = \overline{1,m}) \tag{7.5}$$

$$\sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} = \sum_{i=1}^{n} p_i = 1 \tag{7.6}$$

$$\sum_{j=1}^{m} \sum_{i=1}^{n} p_{ij} = \sum_{j=1}^{m} q_j = 1 \tag{7.7}$$

Хоёр хэмжээст санамсаргүй хэмжигдэхүүн (ξ_1, ξ_2)-ийн утга ямар нэг (D) мужид харъяалагдах үзэгдлийн магадлал нь, $\varphi(x,y)$ гэсэн сөрөг биш утгатай функцээс (D) мужаар авсан хоёрлосон интегралтай тэнцүү бол (ξ_1, ξ_2)

эрэмбэлэгдсэн хосыг хоёр хэмжээст тасралтгүй санамсаргүй хэмжигдэхүүн гэнэ.

$$P((\xi_1,\xi_2)\in D)=\iint\limits_{(D)}\varphi(x,y)dxdy \tag{7.8}$$

$\varphi(x,y)$-ийг хоёр хэмжээст санамсаргүй хэмжигдэхүүний нягтын функц буюу хамтын тархалтын нягт гэж нэрлэнэ.
Хоёр хэмжээст санамсаргүй хэмжигдэхүүний тархалтын нягт нь
1) $\varphi(x,y)\geq 0, \quad \forall(x,y)\in R^2$.
2) $\iint\limits_{R^2}\varphi(x,y)dxdy=1$.
3) $\varphi(x,y)=\dfrac{\partial^2 F(x,y)}{\partial x\partial y}$ чанаруудыг хангана.
Хэрэв (D) мужийг $(D)=\{(\xi_1,\xi_2)|-\infty<\xi_1<x,\ -\infty<\xi_2<y\}$ гэж авбал $P((\xi_1,\xi_2)\in D)=P(\xi_1<x,\ \xi_2<y)=\iint\limits_{(D)}\varphi(x,y)dxdy=\int\limits_{-\infty}^{x}\int\limits_{-\infty}^{y}\varphi(x,y)dxdy$
буюу тархалтын функцийг ашиглавал:

$$F(x,y)=\int\limits_{-\infty}^{x}\int\limits_{-\infty}^{y}\varphi(x,y)dxdy \tag{7.9}$$

Хоёр хэмжээст тасралтгүй санамсаргүй хэмжигдэхүүний тархалтын функц $F(x,y)$-ийг мэдсэнээр ξ_1 ба ξ_2 санамсаргүй хэмжигдэхүүн тус бүрийн тархалтын ба нягтын функцийг олж болно.
Тухайлбал, $F_1(x)$-нь ξ_1-ийн тархалтын функц бол

$$F_1(x)=P(\xi_1<x)=P(\xi_1<x,-\infty<\xi_2<+\infty)=$$

$$=\int\limits_{-\infty}^{x}\int\limits_{-\infty}^{+\infty}\varphi(x,y)dxdy=\int\limits_{-\infty}^{x}\left(\int\limits_{-\infty}^{\infty}\varphi(x,y)dy\right)dx$$

Эндээс нягтын функц $f_1(x)$

$$f_1(x)=\int\limits_{-\infty}^{+\infty}\varphi(x,y)dy \tag{7.10}$$

Үүнчлэн ξ_2-ийн хувьд $F_2(y)=\int\limits_{-\infty}^{y}\left(\int\limits_{-\infty}^{+\infty}\varphi(x,y)dx\right)dy$ буюу

$$f_2(y)=\int\limits_{-\infty}^{+\infty}\varphi(x,y)dx \tag{7.11}$$

ξ_1, ξ_2 санамсаргүй хэмжигдэхүүний хувьд

$$P(\xi_1 < x, \quad \xi_2 < y) = P(\xi_1 < x) \cdot P(\xi_2 < y) \tag{7.12}$$

нөхцөл биелэгдэж байвал хамааралгүй хэмжигдэхүүнүүд гэнэ. Өөрөөр хэлбэл, хамтын тархалтын функц нь санамсаргүй хэмжигдэхүүн тус бүрийн тархалтын функцүүдийн үржвэртэй тэнцүү байвал тэдгээрийг хамааралгүй гэнэ.

$\varphi(x, y)$-хамтын тархалтын нягт бүхий тасралтгүй санамсаргүй хэмжигдэхүүний системийн хувьд (7.12) нөхцөлийг бичвэл

$$\varphi(x, y) = f_1(x) \cdot f_2(y) \tag{7.13}$$

Үүнд, $f_1(x) - \xi_1$ санамсаргүй хэмжигдэхүүний нягтын функц,
$f_2(y) - \xi_2$ санамсаргүй хэмжигдэхүүний нягтын функц.
$M\xi_1$, $M\xi_2$-математик дунджууд нь системийн сарнилын төвийн координатыг илэрхийлэх бөгөөд дараах томъёогоор тодорхойлогдоно.
Дискрет санамсаргүй хэмжигдэхүүний системийн хувьд:

$$\begin{aligned} M(\xi_1) &= \sum_{i=1}^{n}\sum_{j=1}^{m} x_i p_{ij} = \sum_{i=1}^{n} x_i p_i \\ M(\xi_2) &= \sum_{j=1}^{m}\sum_{i=1}^{n} y_j p_{ij} = \sum_{j=1}^{m} y_j q_j \end{aligned} \tag{7.14}$$

Тасралтгүй санамсаргүй хэмжигдэхүүний системийн хувьд:

$$\begin{aligned} M(\xi_1) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x \cdot \varphi(x, y) dxdy = \int_{-\infty}^{+\infty} x \cdot f_1(x) dx \\ M(\xi_2) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y \cdot \varphi(x, y) dxdy = \int_{-\infty}^{+\infty} y \cdot f_2(y) dy \end{aligned} \tag{7.15}$$

Үүнд $\varphi(x, y)$-хамтын тархалтын нягт.
$D(\xi_1)$, $D(\xi_2)$-диспресүүд нь санамсаргүй цэгийн OX ба OY тэнхлэгийн дагуух сарнилыг илэрхийлэх ба дараах томъёогоор олно.
Дискрет санамсаргүй хэмжигдэхүүний системийн хувьд::

$$\begin{aligned} D(\xi_1) &= \sum_{i=1}^{n}\sum_{j=1}^{m} (x_i - M\xi_1)^2 p_{ij} = \sum_{i=1}^{n} (x_i - M\xi_1)^2 p_i \\ D(\xi_2) &= \sum_{j=1}^{m}\sum_{i=1}^{n} (y_j - M\xi_2)^2 p_{ij} = \sum_{j=1}^{m} (y_j - M\xi_2)^2 q_j \end{aligned} \tag{7.16}$$

Тасралтгүй санамсаргүй хэмжигдэхүүний системийн хувьд:

$$\begin{aligned} D(\xi_1) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (x - M\xi_1)^2 \varphi(x, y) dxdy = \int_{-\infty}^{+\infty} (x - M\xi_1)^2 f_1(x) dx \\ D(\xi_2) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (y - M\xi_2)^2 \varphi(x, y) dxdy = \int_{-\infty}^{+\infty} (y - M\xi_2)^2 f_2(y) dy \end{aligned} \tag{7.17}$$

Жишээ-7.1 Нягтын функц $\varphi(x,y) = \dfrac{1}{\pi^2(1+x^2)(1+y^2)}$,
$(D) = \{(x,y) \mid 0 \le x \le 1, \;\; 0 \le y \le 1\}$ бол
1.) $P((\xi_1, \xi_2) \in D) = ?$
2.) Системийн тархалтын функц $F(x,y)$-ийг тус тус ол.

1.) Тодорхойлолт ёсоор

$$P(0 \le \xi_1 \le 1, \;\; 0 \le \xi_2 \le 1) = \frac{1}{\pi^2} \int\limits_0^1 \int\limits_0^1 \frac{dxdy}{(1+x^2)(1+y^2)} =$$

$$= \frac{1}{\pi^2} \int\limits_0^1 \frac{dy}{1+y^2} \int\limits_0^1 \frac{dx}{1+x^2} = \frac{1}{\pi^2} (\mathrm{arctg} y|_0^1)(\mathrm{arctg} x|_0^1) = \frac{1}{16}.$$

$$\textbf{2.)} \;\; F(x,y) = P(\xi_1 < x, \; \xi_2 < y) = \frac{1}{\pi^2} \int\limits_{-\infty}^{x} \frac{dx}{1+x^2} \int\limits_{-\infty}^{y} \frac{dy}{1+y^2} =$$

$$= \frac{1}{\pi^2} \left(\mathrm{arctg} x + \frac{\pi}{2} \right) \cdot \left(\mathrm{arctg} y + \frac{\pi}{2} \right)$$

Хэрэв (X, Y) сиситемийн нягтын функц нь

$$\varphi(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \cdot exp \left[-\frac{1}{1-\rho^2} \left(\frac{(x-a)^2}{2\sigma_x^2} - \rho \frac{(x-a)(y-b)}{\sigma_x\sigma_y} + \frac{(y-b)^2}{2\sigma_y^2} \right) \right] \tag{7.18}$$

($\sigma_x > 0, \;\; \sigma_y > 0, \;\; \rho$-тогтмол) томъёогоор илэрхийлэгдэх бол $(X, Y)-$ийг хоёр хэмжээст хэвийн тархалттай гэнэ. Үүнд: $\sigma_x, \;\; \sigma_y$ нь дундаж квадрат хазайлтууд, (a, b) нь системийн сарнилын төв, ρ-корреляцийн коэффициент.

Хэрэв хэвийн тархалттай системийн X, Y-хэмжигдэхүүнүүд корреляц хамааралгүй бол ($\rho = 0$) системийн тархалтын нягт дараах хэлбэртэй болно.

$$\varphi(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} exp \left[-\frac{(x-a)^2}{2\sigma_x^2} - \frac{(y-b)^2}{2\sigma_y^2} \right] \tag{7.19}$$

(7.18) ба (7.19) томъёогоор тодорхойлогдох гадаргууг нормаль тархалтын гадаргуу гэнэ. (7.19) гадаргуугийн түвшний шугам нь

$$\frac{(x-a)^2}{\sigma_x^2} + \frac{(y-b)^2}{\sigma_y^2} = c^2, \quad (c = const) \tag{7.20}$$

хэлбэртэй болно. Эдгээр нь $c\sigma_x, \;\; c\sigma_y$, хагас тэнхлэгүүд бүхий эллипсийг дүрслэх ба сарнилын эллипсүүд гэж нэрлэгдэнэ. $\sigma_x = \sigma_y = \sigma$ тохиолдолд сарнилын эллипсүүд

$$(x-a)^2 + (y-b)^2 = (c\sigma)^2 \tag{7.21}$$

гэсэн тойргууд болох тул өгөгдсэн хэвийн тархалтыг дугуй нормаль тархалт гэнэ. Дугуй нормаль тархалтын нягт дараах хэлбэртэй болно.

$$\varphi(x,y) = \frac{1}{2\pi\sigma^2} exp\left[-\frac{(x-a)^2}{2\sigma^2} - \frac{(y-b)^2}{2\sigma^2}\right]. \tag{7.22}$$

7.2 Хоёр хэмжээст тасралтгүй тархалтын жишээ

Жишээ-7.2 (Хавтгайн муж дээрх жигд тархалт.)
Ω-төгсгөлөг талбайтай хавтгайн муж байг. Ω-муж дээр санамсаргүй цэгүүдийн нягт нь тогтмол, Ω-ийн гадна тэгтэй тэнцүү байх тархалтыг хоёр хэмжээст жигд тархалт гэнэ. Өөрөөр хэлбэл,

$$\varphi(x,y) = \begin{cases} c & \text{хэрэв} \quad (x,y) \in \Omega \\ 0 & \text{хэрэв} \quad (x,y) \notin \Omega \end{cases}$$

$c-$тогтмолыг $\iint\limits_{\Omega} \varphi(x,y)dxdy = 1$ нөхцөлөөс олно.

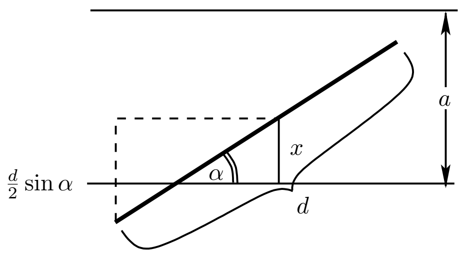
$$\iint\limits_{\Omega} \varphi(x,y)dxdy = \iint\limits_{\Omega} cdxdy = c \cdot S_\Omega = 1 \to c = \frac{1}{S_\Omega}.$$

$S_\Omega - \Omega$ мужийн талбай. Хэрэв G-нь Ω-ийн ямар нэгэн хэсэг бол санамсаргүй цэг (x,y) G мужид унах магадлал $P((x,y) \in G) = \iint\limits_{(G)} \frac{1}{S_\Omega} dxdy = \frac{1}{S_\Omega} \cdot S_G$, S_G-G мужийн талбай. Иймхүү $P((x,y) \in G) = \frac{S_G}{S_\Omega}$.

Жишээ-7.3 (Бюффены бодлого)

Хоорондоо a зайтай параллель шулуунууд татагдсан хавтгай руу d-урттай зүү хаяжээ. Хэрэв $d < a$ бол зүү ямар нэгэн шулууныг огтолж унах магадлалыг ол.

Дараах хоёр санамсаргүй хэмжигдэхүүнийг оруулъя. x-зүүний голоос хамгийн ойр орших шулуун хүртлэх зай. α-зүү ба уг шулууны хоорондох өнцөг (Зураг 7.2).

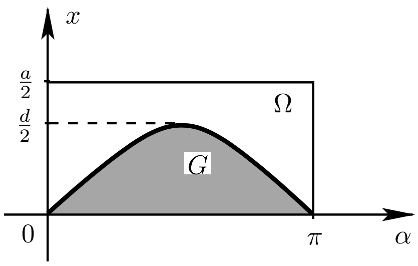


Зураг 7.2:

x, α-хэмжигдэхүүнүүд хамааралгүй бөгөөд жигд тархалттай болох нь илэрхий. x-нь $[0, \frac{a}{2}]$ завсарт, α нь $[0, \pi]$ завсарт жигд тархалттай. Иймд санамсаргүй цэг (x, α)-ийн боломжит байрлалын муж нь $\Omega = \{(x,\alpha) \mid 0 \le x \le \frac{a}{2}, \quad 0 \le \alpha \le \pi\}$ гэсэн тэгш өнцөгт болно. (x,α)-гэсэн системийн тархалтын нягтын нь $\varphi(x,\alpha) = \varphi_1(x) \cdot \varphi_2(\alpha)$ тул

$$\varphi(x,\alpha) = \begin{cases} \frac{2}{a} \cdot \frac{1}{\pi}, & \text{хэрэв} \quad 0 \le x \le \frac{a}{2}, \quad 0 \le \alpha \le \pi \\ 0, & \text{бусад} \quad \text{тохиолдолд} \end{cases}$$

Зүү ямар нэг шугамыг огтолж унах байрлалын муж нь G болно. Зөвхөн $x \leq \dfrac{d}{2}\sin\alpha$ үед дээрх нөхцөл биелнэ (Зураг 7.3). Өөрөөр хэлбэл,



Зураг 7.3:

$$G = \left\{ (x, y) \in \Omega \mid x \leq \frac{d}{2}\sin\alpha \right\}.$$
$$S_\Omega = \frac{a}{2} \cdot \pi = \frac{a\pi}{2},$$
$$S_G = \int_0^\pi \frac{d}{2}\sin\alpha d\alpha = -\frac{d}{2}\cos\alpha|_0^\pi = d,$$
$$p = \frac{S_G}{S_\Omega} = \frac{2d}{a\pi}.$$

Бүлэг 8

Түгээмэл хэрэглэгдэх зарим тархалтын хуулиуд

8.1 Вейбуллийн тархалт

Тархалтын функц нь $F(x)=1-exp\left\{-\left(\frac{x-a}{b}\right)^{c}\right\}$ хуулиар илэрхийлэгдэх тархалтыг Вейбуллийн тархалт гэнэ. $a,\ b,\ c-$ параметрүүд. Нягтын функцийг бичвэл

$$f(x)=\frac{c}{b}\Big(\frac{x-a}{b}\Big)^{c-1}\cdot exp\left\{-\left(\frac{x-a}{b}\right)^{c}\right\}$$

Үүнд, $x\geq a,\quad b>0,\quad c>0.$

$\xi-$Вейбуллийн тархалттай бол математик дундаж нь:

$$M(\xi)=a+b\cdot\Gamma\Big(1+\frac{1}{c}\Big).$$

Үүнд: $\Gamma(x)=\int\limits_{0}^{+\infty}e^{-t}\cdot t^{x-1}dt$ буюу Гамма функц гэж нэрлэнэ.

Тухайн тохиолдолд $\Gamma(n+1)=n!,\quad n\in N.$

8.2 Гамма тархалт

$$f(x)=\frac{1}{\Gamma(\alpha+1)\cdot\beta^{\alpha+1}}\cdot x^{\alpha}\cdot e^{-x/\beta},\ (x>0)$$

нягт бүхий тархалтыг Гамма тархалт гэнэ. Үүнд, $\beta>0,\ \alpha>-1$ нөхцлийг хангах параметрүүд. Хэрэв ξ нь гамма тархалттай бол түүний математик дундаж ба диспрес дараах томъёогоор тодорхойлогдоно.

$$M(\xi)=\beta(\alpha+1),\quad D(\xi)=\beta^{2}(\alpha+1).$$

8.3 Бета тархалт

ξ санамсаргүй хэмжигдэхүүний тархалтын нягт

$$f(x) = \frac{\Gamma(\alpha+\beta+2)}{\Gamma(\alpha+1)\cdot\Gamma(\beta+1)}\cdot x^{\alpha}\cdot(1-x)^{\beta}, \quad (0<x<1)$$

томъёогоор тодорхойлогдох бол түүнийг Бета тархалт гэнэ.
Үүнд: $\alpha > -1,\ \beta > -1$. Мат.дундаж ба дисперс нь:

$$M(\xi) = \frac{\alpha+1}{\alpha+\beta+2}, \qquad D(\xi) = \frac{(\alpha+1)(\beta+1)}{(\alpha+\beta+2)^2(\alpha+\beta+3)}.$$

8.4 Хи-квадрат (χ^2) тархалт

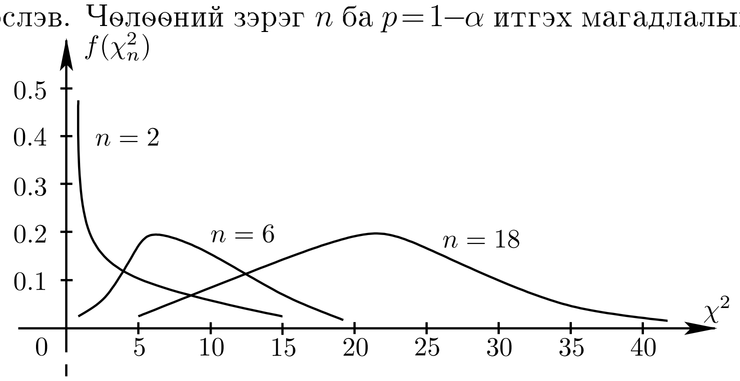
$\xi_1, \xi_2, \ldots, \xi_n$ хамааралгүй санамсаргүй хэмжигдэхүүнүүд бөгөөд $\xi_i (i = \overline{1,n})$ бүр нь $a = 0,\ \ \sigma^2 = 1$ параметрүүд бүхий хэвийн тархалттай байг.
Тэгвэл: $\chi^2(n) = \xi_1^2+\xi_2^2+\cdots+\xi_n^2$ хэмжигдэхүүнийг n чөлөөний зэрэг бүхий χ^2 тархалттай санамсаргүй хэмжигдэхүүн гэнэ. $\chi^2(n)$- санамсаргүй хэмжигдэхүүний тархалтын нягт

$$f(x) = \frac{1}{2^{n/2}\cdot\Gamma\left(\frac{n}{2}\right)}\cdot x^{(n/2)-1}\cdot e^{-x/2}, \quad (x \geq 0)$$

хэлбэртэй байна. Энэ илэрхийлэл $\alpha = \frac{n}{2} - 1, \quad \beta = 2$ байх үеийн Гамма тархалтын тухайн тохиолдол юм.
$\chi^2(n)$ тархалтын математик дундаж ба дисперс нь:

$$M(\chi^2(n)) = n, \quad D(\chi^2(n)) = 2n.$$

Чөлөөний зэрэг n-ийн янз бүрийн утганд тархалтын нягтын функцийг дараах графикт дүрслэв. Чөлөөний зэрэг n ба $p=1-\alpha$ итгэх магадлалын янз бүрийн



Зураг 8.1:

утгуудад харгалзах $\chi^2(n)$-ын таблицыг зохиосон байдаг (хавсралтын табл N4 -ийг үз).

Чөлөөний зэргийн их утгуудад ($n > 30$) $\chi^2(n)$ тархалт $N(n, 2n)$ тархалт уруу нийлнэ. Иймд практик тооцоонд $n > 30$ үед $\chi^2(n)$ тархалтыг $a = n$, $\sigma^2 = 2n$ параметрүүд бүхий хэвийн тархалтаар сольдог. Тархалтын функцийн график дээр орших p ординаттай цэгийн абсцисс x_p-г p-эрэмбийн квантиль гэнэ. Хи-квадрат тархалтын p-эрэмбийн квантилийг $\chi_p^2(n)$ гэвэл $n > 30$ үед эдгээр квантилуудыг $N(0, 1)$ тархалтын квантилаар олж болно:

$$\chi_p^2(n) \approx \frac{1}{2}\left(U_p + \sqrt{2n-1}\right)^2$$

Үүнд, U_p-нь $N(0, 1)$ тархалтын квантиль.
Бага эрэмбийн квантилиудын хувьд илүү нарийвчлалтай дараах томъёог хэрэглэнэ.

$$\chi_p^2(n) \approx n\left(1 - \frac{2}{9n} + U_p\sqrt{\frac{2}{9n}}\right)^3.$$

8.5 Логнормаль тархалт

Хэрэв эерэг тодорхойлогдсон санамсаргүй хэмжигдэхүүн ξ-ийн хувьд $\ln\xi$-санамсаргүй хэмжигдэхүүн a, σ^2 параметрүүд бүхий хэвийн тархалттай бол түүнийг логнормаль тархалттай гэнэ.
Өөрөөр хэлбэл, $\xi \sim N(a, \sigma^2): \ a = M(\ln\xi), \ \ \sigma^2 = D(\ln\xi)$.
Логнормаль тархалтын нягт дараах хэлбэртэй болно.

$$f(x) = \begin{cases} \dfrac{1}{\sqrt{2\pi}\sigma x} \cdot exp\Big[-\dfrac{(\ln x - a)^2}{2\sigma^2}\Big], & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Логнормаль тархалтын параметрүүд:

$$M(\xi) = e^{a+\sigma^2/2}, \quad D(\xi) = e^{2a+\sigma^2} \cdot (e^{\sigma^2} - 1).$$

8.6 Стьюдентийн тархалт

$\xi, \xi_1, \xi_2, \ldots, \xi_k$ санамсаргүй хэмжигдэхүүн тус бүр $a = 0$ ба дурын σ^2 параметрүүдтэй хэвийн тархалттай байг. ξ нь бусдаасаа (ξ_i-бүрээс) хамааралгүй. Харин ξ_i санамсаргүй хэмжигдэхүүнүүдийн дотор шугаман хамааралгүй n-хэмжигдэхүүн байдаг гэе. Тэгвэл ($n \leq k$)

$$T = \frac{\xi}{\sqrt{\dfrac{1}{n}(\xi_1^2 + \xi_2^2 + \cdots + \xi_k^2)}}$$

хэмжигдэхүүнийг n-чөлөөний зэрэг бүхий Стьюдентийн тархалттай гэнэ. $t-$тархалт гэж нэрлэх нь бий. Стьюдентийн тархалтын нягт ба параметрүүд

$$S_n(x) = \frac{1}{\sqrt{n\pi}} \cdot \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-(n+1)/2}, \quad -\infty < x < +\infty$$

$$M[S_n(x)] = 0, \quad D[S_n(x)] = \frac{n}{n-2}$$

томъёогоор тодорхойлогдоно. Чөлөөний зэрэг n ба $p = 1 - a$ магадлалын янз бүрийн утгуудад Стьюдентийн тархалтын функцийн таблицыг зохиосон байдаг. (N7 таблицыг үз.)
$n \to \infty$ үед Стьюдентийн тархалт нь но'рмчлогдсон нормаль тархалт уруу нийлнэ. Стьюдентийн тархалтын нягт $x = 0$ шулууны хувьд тэгш хэмтэй учир энэ тархалтын квантилиуд $t_{1-p}(n) = -t_p(n)$ нөхцлөөр холбогдоно. $n > 30$ үед $t_p(n) \approx U_p$ томъёогоор олно. Үүнд U_p нь $N(0,1)$ тархалтын квантиль.

8.7 Фишерийн тархалт (F-тархалт)

$\dfrac{\chi^2(n_1)}{n_1}$ ба $\dfrac{\chi^2(n_2)}{n_2}$ гэсэн хамааралгүй санамсаргүй хэмжигдэхүүний хувьд $F(n_1, n_2) = \dfrac{\chi^2(n_1)/n_1}{\chi^2(n_2)/n_2}$ хэмжигдэхүүний тархалтыг n_1 ба n_2 чөлөөний зэрэгтэй Фишерийн тархалт гэнэ. Фишерийн тархалтын нягт ба тоон үзүүлэлтүүд:

$$f_F(x) = \begin{cases} 0, & x \le 0 \\ \dfrac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma\left(\frac{n_1}{2}\right) \cdot \Gamma\left(\frac{n_2}{2}\right)} \left(\dfrac{n_1}{n_2}\right)^{n_1/2} \dfrac{x^{(n_1/2)-1}}{\left(1 + \frac{n_1}{n_2}x\right)^{(n_1+n_2)/2}}, & x > 0 \end{cases}$$

$$M[F(n_1, n_2)] = \frac{n_2}{n_2 - 2}, \ (n_2 > 2).$$

$$D[F(n_1, n_2)] = \frac{2n_2^2(n_1 + n_2 - 2)}{n_1(n_2 - 2)^2(n_2 - 4)}, \ (n_2 > 4).$$

Фишерийн тархалтын p-эрэмбийн квантилийг $F_p(n_1, n_2)$ гэвэл тэр нь $(1-p)$ эрэмбийн квантильтай дараах харьцаагаар холбогдоно.

$$F_p(n_1, n_2) = \frac{1}{F_{1-p}(n_2, n_1)}$$

$N(0,1), \ \chi^2(n), \ S_n(x), \ F(n_1, n_2)$ тархалтууд дараах холбоотой болохыг тэмдэглэе.

$$S_n^2(x) = F(1, n), \quad F(n, \infty) = \frac{\chi^2(n)}{n}, \quad \chi^2(1) = N(0;1).$$

Бүлэг 9

Их тооны хуулиуд

Бид үүний өмнө зөвхөн нэг санамсаргүй хэмжигдэхүүний тархалтын хууль түүний параметрүүдийн тухай судалсан билээ. Тухайн нэг санамсаргүй хэмжигдэхүүний хувьд чухам ямар утгаа авахыг хэлж болохгүй учраас олон санамсаргүй хэмжигдэхүүний нийлбэрийн хувьд тодорхой зүйл хэлж болохгүй мэт. Гэвч олон санамсаргүй хэмжигдэхүүний нийлбэр, нэмэгдэхүүний тоог олшроход тодорхой хууль зүй тогтолд захирагдаж ирдэг байна. Иймд их тооны хууль нь, олон тооны туршилтын дундаж үзүүлэлтүүд тодорхой нөхцлийн үед ямар нэг тоон утганд ойртох тухай өгүүлсэн хэд хэдэн теоремуудаас тогтоно.

9.1 Чебышевын тэнцэтгэл биш

Чебышевын тэнцэтгэл биш нь санамсаргүй хэмжигдэхүүний утга өөрийнхөө дундаж утгаас хазайх хазайлтын абсалют хэмжигдэхүүн ямар нэг эерэг тооноос хэтрэхгүй байх магадлалыг үнэлж болох уу? гэсэн асуудалд хариу өгнө.

Хэрэв санамсаргүй хэмжигдэхүүн ξ-ийн дисперс төгсгөлөг бол дурын эерэг тоо ε-гийн хувьд

$$P(|\xi - M(\xi)| \leq \varepsilon) \geq 1 - \frac{D(\xi)}{\varepsilon^2} \tag{9.1}$$

тэнцэтгэл биш биелнэ.

9.2 Чебышевын теорем

Хэрэв $\xi_1, \xi_2, \ldots, \xi_n$ гэсэн хос хосоороо хамааралгүй санамсаргүй хэмжигдэхүүний дисперсүүд нь жигд зааглагдсан $(D\xi_i < C,\ i=\overline{1,n},\ C-const)$ бол $\forall\varepsilon > 0$ хувьд

$$\lim_{n\to\infty} P\left(\left|\frac{\xi_1+\xi_2+\ldots+\xi_n}{n} - \frac{M(\xi_1)+M(\xi_2)+\ldots+M(\xi_n)}{n}\right| < \varepsilon\right) = 1$$

тэнцэтгэл хүчинтэй. Товчоор бичвэл

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\sum_{i=1}^{n}\xi_i - \frac{1}{n}\sum_{i=1}^{n}M(\xi_i)\right| < \varepsilon\right) = 1 \tag{9.2}$$

Эндээс харахад санамсаргүй хэмжигдэхүүн тус бүр өөрсдийнхөө математик дунджаасаа их ялгагдаж болох боловч санамсаргүй хэмжигдэхүүний тоо хязгааргүй олшроход тэдгээрийн арифметик дундаж нь математик дунджуудын арифметик дунджаас бараг ялгагдахгүй байна.
Практикт Чебышевийн теоремийг дараах үнэлгээ хэлбэрээр ашиглана:

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n}\xi_i - \frac{1}{n}\sum_{i=1}^{n}M(\xi_i)\right| < \varepsilon\right) \geq 1 - \frac{C}{n\varepsilon^2.} \tag{9.3}$$

Тухайн тохиолдолд $M(\xi_i) = a,\ (i=\overline{1,n})$ бол

$$\frac{M(\xi_1) + M(\xi_2) + \cdots + M(\xi_n)}{n} = a.$$

Тэгвэл (9.2) томъёо $\lim\limits_{n\to\infty} P\left(\left|\frac{1}{n}\sum\limits_{i=1}^{n}\xi_i - a\right| < \varepsilon\right) = 1$ хэлбэртэй болно.
Үүнийг тайлбарлавал: Ямар нэгэн физик хэмжигдэхүүний жинхэнэ утга "a"-г олохын тулд бие биенээсээ үл хамаарах туршилтууд хийсэн гэе. Хэмжилт бүрд тодорхой хэмжээний алдаа гарна. Иймд хэмжилт бүрийн үр дүнг ξ_i (i-туршилтын дугаар) гэсэн санамсаргүй хэмжигдэхүүн гэж үзэх тул $M(\xi_i) = a,\ (i=\overline{1,n})$ юм. Энэ хэмжилтүүд ямар нэг нарийвчлалтай хийгдэнэ. Өөрөөр хэлбэл, $D(\xi)=C,\ (i=\overline{1,n})$. Иймд их тооны хуулийн нөхцөл ёсоор туршилтын тоо хүрэлцээтэй их байхад хэмжилтийн дүнгүүдийн арифметик дундаж нь жинхэнэ утга "a"-аас бараг үл ялгагдана.

9.3 Бернуллийн теорем.

Хэрэв үл хамаарах туршилтын дүнд A-үзэгдэл нэг ижил p магадлалтайгаар илэрдэг бол түүний явагдах харьцангүй давтамж ба p магадлалын ялгавар, абсалют хэмжигдэхүүнээрээ дурын эерэг тоо ε-гоос бага байх магадлал туршилтын тоо хүрэлцээтэй өсөхөд нэгд хичнээн ч ойрхон байж болно. Өөрөөр хэлбэл,

$$\lim_{n\to\infty} P\left(\left|\frac{m}{n} - p\right| \leq \varepsilon\right) = 1 \tag{9.4}$$

Бернуллийн теоремийг хязгаарын сонгодог утгаар, бүх n-ийн хувьд $\lim\limits_{n\to\infty}\frac{m}{n}=p$ байна гэж ойлгож үл болно. Өөрөөр хэлбэл, $\left|\frac{m}{n}-p\right|<\varepsilon$ тэнцэтгэл биш зарим n-ийн утганд алдагдахыг Бернулийн теоремд зөвшөөрч болно. Үүнтэй холбогдон магадлалаараа нийлэх нийлэлт, бараг гарцаагүй нийлэх

нийлэлтүүд тодорхойлогддог билээ.
Практикт (9.4) теоремийг дараах үнэлгээ хэлбэрээр ашиглана:

$$P\left(\left|\frac{m}{n}-p\right| \leq \varepsilon\right) \geq 1-\frac{p(1-p)}{n\varepsilon^2}. \tag{9.5}$$

9.4 Пуассоны теорем.

Үл хамаарах туршилтаар i-р туршилт бүрд A үзэгдэл явагдах магадлал p_i $(i=\overline{1,n})$ бол

$$\lim_{n\to\infty} P\left(\left|\frac{m}{n}-\frac{p_1+p_2+\cdots+p_n}{n}\right| \leq \varepsilon\right) = 1. \tag{9.6}$$

Бернуллийн теорем ижил (тогтмол) нөхцлийн үед (туршилтын тоо хангалттай их байхад) үзэгдлийн давтамжийн тогтворжилтыг илэрхийлж байгаа бол Пуассоны теорем ижил биш (тогтвортой биш) өөр өөр нөхцлийн үед давтамжийн тогтворжилтыг илэрхийлж байна.

9.5 Ляпуновын теорем (Хязгаарын гол теорем)

Хэвийн тархалттай, төсгөлөг тооны үл хамаарах санамсаргүй хэмжигдэхүүнүүдийн нийлбэр мөн хэвийн тархалттай болохыг хялбархан баталж болно. Хэдийгээр санамсаргүй хэмжигдэхүүн тус бүр нь хэвийн тархаагүй ч тодорхой нөхцөл биелэгдэх үед тэдгээрийн нийлбэр хэвийн тархалттай байж болно. Үүнд Ляпуновын их тооны хууль хариу өгнө.

Хэрэв $\xi_1, \xi_2, \ldots, \xi_n$ гэсэн хамааралгүй санамсаргүй хэмжигдэхүүний дарааллын хувьд

1.) Санамсаргүй хэмжигдэхүүн бүр төгсгөлөг мат.дундаж, дисперстэй: ($M(\xi_i)=a_i,\ D(\xi_i)=\sigma_i^2,\ i=\overline{1,n}$)

2.) Аль ч санамсаргүй хэмжигдэхүүний утга бусад санамсаргүй хэмжигдэхүүний утгуудаас эрс ялгаатай биш бол

эдгээр санамсаргүй хэмжигдэхүүний нийлбэр $\eta=\xi_1+\xi_2+\ldots+\xi_n$ нь $n-$ийг хүрэлцээтэй өсөхөд $\sum\limits_{i=1}^{n} a_i$ мат.дундаж, $\sum\limits_{i=1}^{n} \sigma_i^2$ дисперс бүхий стандарт хэвийн тархалт руу нийлнэ. Өөрөөр хэлбэл,

$$F_\eta(x)=P\left(\frac{\sum\limits_{i=1}^{n}\xi_i-\sum\limits_{i=1}^{n}a_i}{\sqrt{\sum\limits_{i=1}^{n}\sigma_i^2}}<x\right) \xrightarrow{n\to\infty} \frac{1}{2\pi}\int\limits_{-\infty}^{x} e^{-t^2/2}dt. \tag{9.7}$$

Энэ тохиолдолд, санамсаргүй хэмжигдэхүүний дараалал $\xi_1, \xi_2, \dots, \xi_n$–ийг хязгаартаа хэвийн тархалттай гэнэ.
Тухайн тохиолдолд $\xi_1, \xi_2, \dots, \xi_n$ нь $(M\xi_i) = a, \ D(\xi_i) = \sigma^2$ гэсэн мат.дундаж, дисперс бүхий нэгэн ижил тархалттай, үл хамаарах санамсаргүй хэмжигдэхүүний дараалал бол $\forall x$–ийн хувьд

$$F_\eta(x) = P\left(\frac{\sum\limits_{i=1}^{n}\xi_i - na}{\sigma\sqrt{n}} < x\right) \xrightarrow{n\to\infty} \frac{1}{2\pi}\int\limits_{-\infty}^{x} e^{-t^2/2}dt.$$

Үүнд, $\sum\limits_{i=1}^{n} M\xi_i = na, \ \sum\limits_{i=1}^{n} D\xi_i = n\sigma^2.$
Хязгаарын гол теоремийн мөрдлөг нь IV бүлгийн §3-д авч үзсэн Муавр-Лапласын локаль ба интеграл томъёо болохыг тэмдэглэе.

Бүлэг 10

Математик статистикийн элементүүд

Магадлалын онолын аргууд дээр тулгуурласан математикийн биеэ даасан салбар бол математик статистик юм.
Санамсаргүй үзэгдэл, процессийн ямарч судалгаа шууд ба шууд бишээр статистик туршилт дээр тулгуурладаг. Математик статистик нь онолын ба практик дүгнэлт хийх зорилгоор статистик туршилтын өгөгдөл ба үр дүнг боловсруулан системчилж, судалж буй объектдоо тохирсон арга, математик загварыг сонгон авах асуудлыг судалдаг. Өөрөөр хэлбэл, санамсаргүй үзэгдлийг ажиглах явцад гарган авсан туршилтын статистик өгөгдлүүдийг бүртгэх, боловсруулах, шинжлэх аргуудыг судална гэсэн үг.
Статистик судалгааны үндсэн зорилго нь эх олонлогийг түүврийн аргаар судалж түүний тархалт болон тоон үзүүлэлтүүдийн магадлалт чанаруудыг тайлбарлахад оршино. Тодруулан хэлбэл, түүврийн аргын тусламжтайгаар статистик таамаглалуудыг дэвшүүлэх, тархалтын үл мэдэгдэх праметрүүдэд статистик үнэлгээ өгөх, праметрүүдийн итгэх завсрыг байгуулах, статистик таамаглалуудыг шалгах, сонгон авсан загвартаа корреляц, регресс, дисперсийн шинжилгээ хийх явдал юм.

10.1 Түүврийн арга. Статистик тархалт. Полигон, гистограмм.

Санамсаргүй үзэгдлийг судлах явцад түүнтэй холбоотой ямар нэг санамсаргүй хэмжигдэхүүний бүх боломжит утгуудыг нэг бүрчлэн авч үзэх шаардлагатай боловч практикт энэ нь боломжгүй. Энэ бэрхшээлээс гарах үндсэн арга бол түүврийн арга юм. Түүврийн аргын үндсэн санаа нь, санамсаргүй хэмжигдэхүүний бүх боломжит утгуудыг судлах биш тэдгээрээс тодорхой тооны утгуудыг таамгаар төлөөлүүлэн авч судалгаа явуулаад нийт бүрэлдэхүүний тухай дүгнэлтүүд гаргахад оршино.

Тухайн туршилтад илэрч болох бүх боломжит үр дүнгүүдийн олонлогийг эх олонлог, эх олонлогийн нийт элементийн тоог эх олонлогийн хэмжээ гэнэ. Эх олонлог төгсгөлөг ба төгсгөлгүй хэмжээтэй байж болно.

Эх олонлогоос таамгаар сонгон авсан n-тооны элементүүдийг n-хэмжээт түүвэр гэнэ. Түүвэр нь буцаалтай ба буцаалтгүй 2 янз байна.

ξ-санамсаргүй хэмжигдэхүүнийг судлах зорилгоор үл хамаарах туршилтууд хийсэн гэж саная. Туршилтын дүнд илэрсэн санамсаргүй хэмжигдэхүүний утгуудыг

$$x_{(1)}, x_{(2)}, \dots, x_{(i)}, \dots, x_{(n)} \tag{10.1}$$

гэе. (10.1) нь n-хэмжээт түүвэр бөгөөд түүнийг статистик цуваа гэнэ. Статистик цуваа нь цаашдын боловсруулалтын үндсэн материал болж өгнө. Статистик цувааны утгуудыг эрэмбэлж, өсөх дарааллаар нь байрлуулъя.

$$x_1, x_2, \dots, x_i, \dots, x_n \tag{10.2}$$

(10.2) цуваag дискрет вариацын цуваа гэнэ. Өөрөөр хэлбэл, эрэмблэгдсэн статистик цуваag дискрет вариацын цуваа гэнэ. (10.2) цувааны x_i утгуудыг вариантууд гэж нэрлэнэ. Вариантууд дотор тэнцүү утгууд байж болох ба x_i утга n_i удаа илэрсэн бол n_i тоог x_i утгын давтамж гэнэ. Энэ тохиолдолд (10.2) цуваа дараах хэлбэртэй болно.

x_1	x_2	x_3	$\dots$	x_i	$\dots$	x_k
n_1	n_2	n_3	$\dots$	n_i	$\dots$	n_k

(10.3)

Үүнд, k-вариантын тоо,

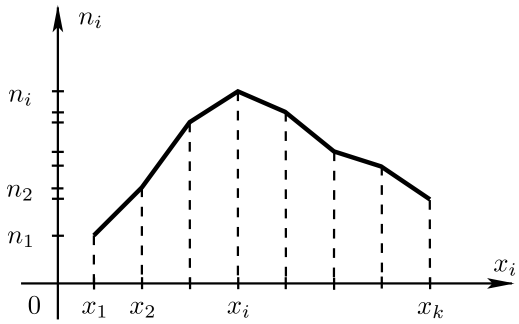
$$n_1 + n_2 + \dots + n_i + \dots + n_k = n \quad \text{буюу} \quad \sum_{i=1}^{k} n_i = n. \tag{10.4}$$

Түүврийн хэмжээ n олон үед, давтамжуудыг тоолохдоо алдаа гаргахгүйн тулд x_i утгын давталт бүрд цэг юмуу зураас харгалзуулах замаар тодорхой дүрсүүдийг ашигладаг. Жишээлбэл, давталтуудыг 5, 5-аар нь багцалж тоолъё гэвэл, /, //, ///, ////, ~~////~~; эсвэл 10, 10-аар нь багцлаж тоолъё гэвэл гэх мэт дүрсүүдийг ашиглаж болно.

$$\omega_i = \frac{n_i}{n} \tag{10.5}$$

тоог x_i утгын илрэх харьцангүй давтамж гэнэ.
Харьцангүй давтамжуудын хувьд

$$\sum_{i=1}^{k} \omega_i = \sum_{i=1}^{k} \frac{n_i}{n} = 1 \quad \text{нөхцөл биелнэ.} \tag{10.6}$$



Зураг 10.1:

Түүврийн вариантууд ба давтамжуудын хамаарлыг түүврийн статистик тархалт гэнэ. Статистик тархалтыг (10.3) хэлбэрийн хүснэгтээр үзүүлэхээс гадна графикаар дүрслэн үзүүлдэг. OX тэнхлэгийн дагуу вариантуудыг байрлуулж OY тэнхлэгийн дагуу давтамжуудыг байрлуулбал $M_1(x_1, n_1), M_2(x_2, n_2), \ldots, M_k(x_k, n_k)$ координатууд бүхий хавтгайн цэгүүд дүрслэгднэ.

Эдгээр цэгүүдийг дэс дараалан хэрчмүүдээр холбоход үүсэх тахир шугамыг давтамжуудын полигон гэнэ.

Үүнчлэн харьцангуй давтамжуудын полигоныг байгуулж болно.

Тасралтгүй санамсаргүй хэмжигдэхүүний хувьд дискрет вариацын цуваа авч үзэх нь тохиромжгүй учраас интервалын вариацын цуваа авч үздэг.

Туршилтын үр дүнд санамсаргүй хэмжигдэхүүн боломжит утгуудаа $[a, b]$ хэрчим дээр авдаг байг. Санамсаргүй хэмжигдэхүүний утгуудыг Δx урттай $[x_0, x_1], [x_1, x_2], \ldots, [x_{k-1}, x_k]$ гэсэн завсруудад хуваая.

Үүнд: $a = x_0 < x_1 < x_2 < \ldots x_{k-1} < x_k = b, \quad \Delta x = x_i - x_{i-1},\ i = (\overline{1, k})$.

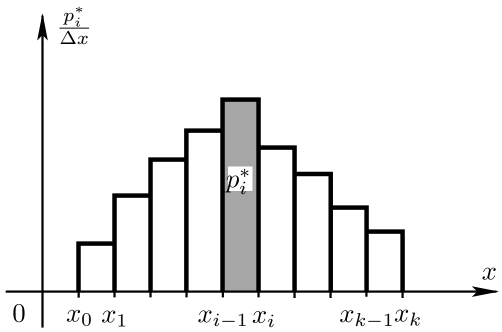
Санамсаргүй хэмжигдэхүүний утга i-р завсарт $([x_{i-1}, x_i])$ n_i удаа илэрсэн байг. Тэгвэл санамсаргүй хэмжигдэхүүний нийт илэрсэн утгуудын тоо $n = \sum\limits_{i=1}^{k} n_i$ буюу $\sum\limits_{i=1}^{k} \dfrac{n_i}{n} = 1$ болно.

$x_1 - x_0, x_2 - x_1, \ldots, x_k - x_{k-1}$ утгуудыг интервалын ялгаврууд, $R = x_k - x_0$-ийг вариацын далайц гэж тус тус нэрлэнэ. i-р завсарт илэрсэн утгуудын харьцангуй давтамжийг $p_i^* = \dfrac{n_i}{n}$ гэж тэмдэглэе. Тэгвэл тасралттай санамсаргүй хэмжигдхүүний хувьд байгуулсан статистик тархалт дараах хэлбэртэй болох ба түүнийг интервалын вариацын цуваа гэнэ.

дугаар i	интервалууд	давтамж n_i	харьцангуй давтамж p_i^*
1	$[x_0, x_1]$	n_1	p_1^*
2	$[x_1, x_2]$	n_2	p_2^*
$\vdots$	$\vdots$	$\vdots$	$\vdots$
k	$[x_{k-1}, x_k]$	n_k	p_k^*

(10.7)

$[x_{i-1}, x_i]$ завсар бүр дээр $\Delta x = x_i - x_{i-1},\ \dfrac{p_i^*}{\Delta x}$ талууд бүхий тэгш өнцөгт байгуулбал зураг 10.2-д үзүүлсэн шаталсан тэгш өнцөгтүүд үүсэх ба түүнийг харьцангуй давтамжуудын гистограмм гэж нэрлэнэ. Энэ тохиолдолд OY тэнхлэгийн дагуу давтамжуудийг авбал гистограммын нийт талбай түүврийн хэмжээтэй тэнцүү байна. Харин OY тэнхлэгийн дагуу харьцангуй давтамжуудыг авбал гистограммын нийт талбай нэгтэй тэнцүү болно.



Зураг 10.2:

Гистограммын нэг тэгш өнцөгтийн талбай p_i^*-тэй тэнцүү учраас гистограммын нийт талбай:

$$S = \sum_{i=1}^{k} p_i^* = \sum_{i=1}^{k} \frac{n_i}{n} = 1. \qquad (10.8)$$

Хэрэв OX тэнхлэгийн дагуух интервалуудын урт нь 1 байхаар масштаб сонгон авсан бол OY тэнхлэгийн дагуу нэгж масштаб нь нэгж ажиглалтанд харгалзана гэж тооцно.

Хэрэв интервалын вариацын цувааг полигоноор дүрслэх шаардлага гарвал интервалуудыг тэдгээрийн дунджуудаар сольж дискрет вариацын цувааны адилаар полигоныг байгуулна.
Интервалын вариацын цувааг байгуулахад интервалын урт Δx-ийг зөвөөр сонгон авах нь судлаж буй хэмжигдэхүүний гол шинж чанарыг харуулахад тустай. Δx интервалын оновчтой хэмжээг тогтооход Стэрджесийн томъёог ашиглаж болно:

$$\Delta x = \frac{x_{max} - x_{min}}{1 + 3.322 \cdot \ln n} \qquad (10.9)$$

Үүнд: x_{max}, x_{min} нь хамгийн их ба хамгийн бага утгатай вариантууд, n-түүврийн хэмжээ. Хэрэв Δx нь бутархай тоо гарвал түүнд ойр орших бүхэл тоог сонгон авна, эсвэл түүнтэй ойролцоо утга бүхий түвэгтэй биш бутархайг сонгон авна. Эхний интервалын эхлэл x_0-ийг $x_0 = x_{min} - \dfrac{\Delta x}{2}$ гэж сонгох ба дараагийн интервалын эхлэл $x_1 = x_0 + h$-ийг эхний интервалын төгсгөлтэй давхацсан байхаар, гэх мэт бусад интервалуудыг байрлуулна. Ажиглалтын утгуудыг тоолох явцад ямар нэг утга интервалын хуваалтын цэгтэй давхцвал түүнийг арын интервалд нь оруулан тоолно.

10.2 Туршилтын тархалтын функц

ξ-ямар нэг санамсаргүй хэмжигдхүүн, $x \in R$ байг.
$(\xi < x)$ үзэгдлийн харьцангуй давтамжийг тодорхойлох функцийг туршилтын тархалтын функц гэнэ. Өөрөөр хэлбэл,

$$F_n(x) \stackrel{def}{=\!=} P^*(\xi < x) \qquad (10.10)$$

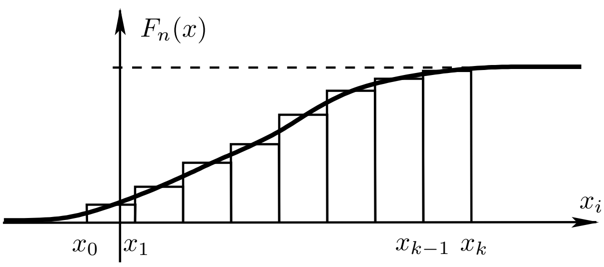
Үүнд, $F_n(x)$-туршилтын тархалтын функц, $P^*(\xi < x)$ нь $(\xi < x)$ үзэгдлийн харьцангуй давтамж. (10.10) томъёог n-хэмжээст түүврийн хувьд товч болон дэлгэрэнгүй бичвэл:

$$F_n(x) = \sum_{x_i < x} \frac{n_i}{n} \qquad (10.11)$$

$$F_n(x) = \begin{cases} 0 & x \le x_0 \\ \dfrac{n_1}{n} & x_0 < x \le x_1 \\ \dfrac{n_1 + n_2}{n} & x_1 < x \le x_2 \\ \dots & \dots\dots\dots\dots \\ \sum\limits_{i=1}^{k-1} \dfrac{n_i}{n} & x_{k-1} < x \le x_k \\ 1 & x > x_k \end{cases} \qquad (10.12)$$

Туршилтын тархалтын функцийн утгуудыг хуримтлагдсан харьцангуй давтамжууд гэнэ. Туршилтын тархалтын функц дараах чанаруудыг хангана.
1. $0 \le F_n(x) \le 1$
2. $F_n(x)$- үл буурах функц байна.
3. a ба b нь харгалзан хамгийн бага ба хамгийн их вариантууд бол
$x \le a$ үед $F_n(x) = 0, \quad x \ge b$ үед $F_n(x) = 1$.
Эх олонлогийн тархалтын функц $F(x)$-ийг туршилтын тархалтын функцээс ялгахын тулд онолын тархалтын функц гэж нэрлэнэ. Бернуллийн их тооны хууль ёсоор туршилтын тархалтын функц $F_n(x)$ нь онолын тархалтын функц $F(x)$-рүү магадлалаараа нийлнэ. Туршилтын тархалтын функцийн график дискрет санамсаргүй хэмжигдэхүүний тархалтын функцийн графикийн адилаар байгуулагдана.

Тасралтгүй санамсаргүй хэмжигдэхүүний хувьд $F_n(x)$ функцийг олохдоо интервалын вариацын цуваа байгуулна. Интервалын дунджуудыг интервал бүрийн төлөөлөгч болгон сонгон авч координатын хавтгай дээр $\left(\dfrac{x_{i-1} + x_i}{2},\ F_n(x)\right)$ цэгүүдийг алгуур муруйгаар холбовол тасралтгүй санамсаргүй хэмжигдэхүүний туршилтын тархалтын функцийн график байгуулагдана (Зураг 10.3).

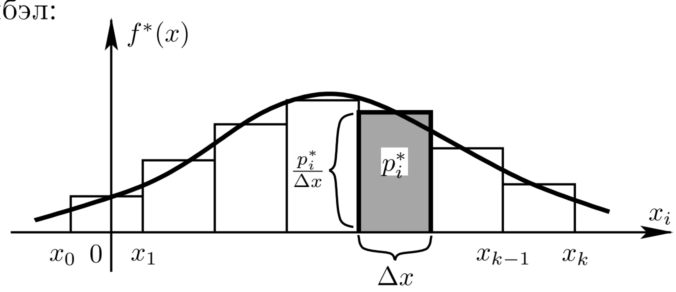


Зураг 10.3:

Харьцангуй давтамжуудын гистограммыг байгуулж, гистограммын тэгш өнцөгтүүдийн өндрүүдтэй тэнцүү утга бүхий дараах функцийг байгуулъя.

$$f^*(x) = \begin{cases} \dfrac{p_1^*}{\Delta x} & \text{хэрэв} \quad x \in [x_0, x_1] \\ \dfrac{p_2^*}{\Delta x} & \text{хэрэв} \quad x \in [x_1, x_2] \\ \dots & \dots \qquad \dots \\ \dfrac{p_k^*}{\Delta x} & \text{хэрэв} \quad x \in [x_{k-1}, x_k] \end{cases} \tag{10.13}$$

Энэ функц нь туршилтын тархалтын нягтын функц болох бөгөөд графикийг дүрсэлж үзүүлбэл:



Зураг 10.4:

Бернуллийн их тооны хууль ёсоор n-ийн хүрэлцээтэй их утганд $F_n(x)$ ба $f^*(x)$ функцийн график нь харгалзан онолын тархалтын болон нягтын функцийн графиктай ойролцоо байна.

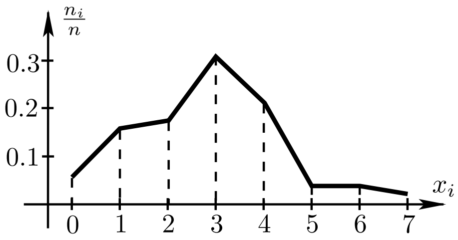
Жишээ-10.1 Өгөгдсөн A-түүврийн утгуудаар вариацын цуваа, харьцангуй ба хуримтлагдсан давтамжууд, харьцангуй давтамжийн полигон (гистограмм), туршилтын тархалтын функц түүний графикийг тус тус байгуул.

A-түүвэр

2	4	2	4	3	3	3	2	0	6	1	2	3	2	2	4	3	3	5	1
0	2	4	3	2	2	3	3	1	3	3	3	1	1	2	3	1	4	3	1
7	4	3	4	2	3	2	3	3	1	4	3	1	4	5	3	4	2	4	5
3	6	4	1	3	2	4	1	3	1	0	0	4	6	4	7	4	1	3	

$n=79$, $x_{min}=0$, $x_{max}=7$, $R=x_{max}-x_{min}=7-0=7$ тул дискрет вариацын цуваа байгуулъя.

x_i	давтамжийн тооцоо	n_i	n_i/n	хуримтлагдсан давтамж
0	////	4	0.0506	0.0506
1	~~////~~ ~~////~~ ///	13	0.1646	0.2152
2	~~////~~ ~~////~~ ////	14	0.1772	0.3924
3	~~////~~ ~~////~~ ~~////~~ ~~////~~ ///	24	0.3038	0.6962
4	~~////~~ ~~////~~~~////~~ /	16	0.2025	0.8987
5	///	3	0.0380	0.9367
6	///	3	0.0380	0.9747
7	//	2	0.0252	1.0000
$\sum$	–	79	1.0000	–

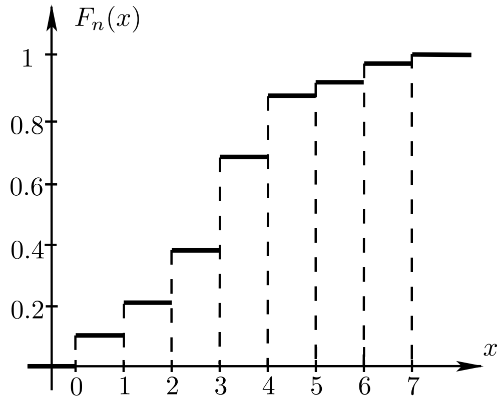


Зураг 10.5:

Гистограммыг байгуулахын тулд $[0, 7]$ хэрчмийг $\Delta x = 1$ урттай интервалуудад хуваая. Харьцангуй давтамжийн полигоныг Зураг10.5-д үзүүлэв.

Хуримтлагдсан харьцангуй давтамжуудаар туршилтын тархалтын функцийг бичиж, график байгуулбал:

$$F_n(x) = \begin{cases} 0 & x \le 0 \\ 0.0506 & 0 < x \le 1 \\ 0.2152 & 1 < x \le 2 \\ 0.3924 & 2 < x \le 3 \\ 0.6962 & 3 < x \le 4 \\ 0.8987 & 4 < x \le 5 \\ 0.9367 & 5 < x \le 6 \\ 0.9747 & 6 < x \le 7 \\ 1 & x > 7 \end{cases}$$



Зураг 10.6:

Жишээ-10.2 Дараах түүврийн хувьд жишээ-10.1-д авч үзсэн бодлогыг бодъё:

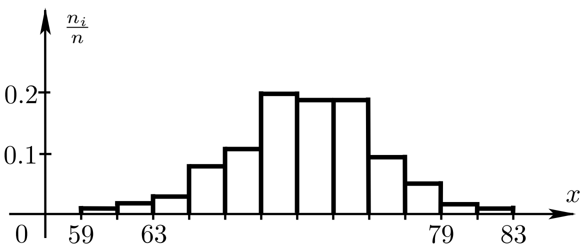
B-түүвэр

65	71	67	73	68	68	72	68	67	70	78	74	79	65	72
65	71	70	69	69	76	71	63	77	75	70	74	65	71	68
74	69	69	66	71	69	73	74	80	69	73	76	69	69	67
67	74	68	74	60	70	66	70	68	64	75	78	71	70	69
73	75	74	72	80	72	69	69	71	70	73	65	66	67	69
71	70	72	76	72	73	64	74	71	76	68	69	75	76	73
74	78	66	75	72	69	68	63	70	70	78	76	73	73	67
71	66	66	72	69	71	71	68	72	69	73	73	66	72	73
70	69	74	72	69	74	70	74	72	76	77	66	62	69	74
76	74	69	64	75	71	76	68	68	78	71	71	68	67	74
68	81	72	68	72	71	71	71	69	61	74	66	70	72	65
67	73	78	73	71	75	73	71	72	68	67	69	69	77	63
71	74	67	68	69	74	69	67	74	66	74	74	69	75	70
73	63	77	74	75										

$n = 200$, $x_{min} = 60$, $x_{max} = 81$, $R = x_{max} - x_{\min} = 21$ учир $\Delta x = 2$ байхаар сонгон авч интервалын вариацын цуваа зохиоё.

интервал x_i	давтамжийн тооцоо	n_i	n_i/n	хуримтлагдсан давтамж
$[59; 61[$		1	0.005	0.005
$[61; 63[$		2	0.010	0.015
$[63; 65[$		7	0.035	0.050
$[65; 67[$		16	0.080	0.130
$[67; 69[$		27	0.135	0.265
$[69; 71[$		40	0.200	0.465
$[71; 73[$		38	0.190	0.655
$[73; 75[$		38	0.190	0.845
$[75; 77[$		18	0.090	0.935
$[77; 79[$		9	0.045	0.980
$[79; 81[$		3	0.015	0.995
$[81; 83[$		1	0.005	1.0000
$\sum$	–	200	1.0000	–

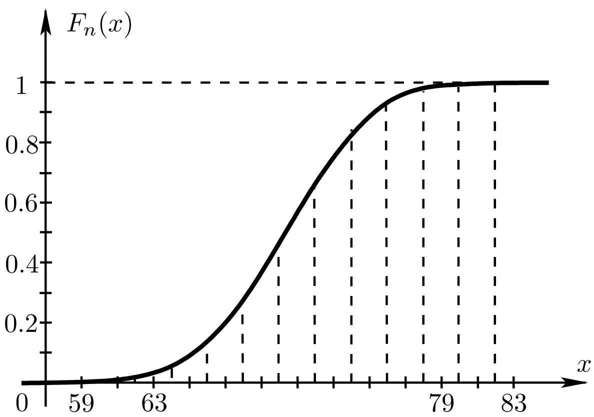
Туршилтын тархалтын функцыг олж график зурвал:



OX-тэнхлэгийн дагуу $59 \div 83$ утгуудыг, OY-тэнхлэгийн дагуу n_i/n давтамжийн утгуудыг байрлуулж интервалын вариацын цуваны харьцангуй давтамжийн гистограммыг Зураг 10.7-д дүрслэв.

Зураг 10.7:

$$F_n(x) = \begin{cases} 0 & x \le 59 \\ 0.005 & 59 < x \le 61 \\ 0.015 & 61 < x \le 63 \\ 0.050 & 63 < x \le 65 \\ 0.130 & 65 < x \le 67 \\ 0.265 & 67 < x \le 69 \\ 0.465 & 69 < x \le 71 \\ 0.655 & 71 < x \le 73 \\ 0.845 & 73 < x \le 75 \\ 0.935 & 75 < x \le 77 \\ 0.980 & 77 < x \le 79 \\ 0.995 & 79 < x \le 81 \\ 1 & x > 81 \end{cases}$$



Зураг 10.8:

10.3 Түүврийн тоон үзүүлэлтүүд

Ямар нэг санамсаргүй хэмжигдэхүүний илэрсэн утгуудаар зохиогдсон n хэмжээт түүвэр $x_1, x_2, \ldots, x_n$ (10.2) өгөгдсөн байг. Бүх ажиглагдсан утгуудын арифметик дундаж:

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad (10.14)$$

(10.14)-ийг түүврийн дундаж гэнэ.
Хэрэв түүвэр вариацын цуваагаар дөгөгдсөн бол (10.14) дараах хэлбэртэй болно.

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{k} x_i \cdot n_i \qquad (10.15)$$

Үүнд, x_i-вариантууд (тасралтгүй санамсаргүй хэмжигдэхүүний тохиолдолд интервалын дундаж), n_i-давтамжууд (зарим тохиолдолд жин гэнэ).
Хэрэв вариантууд их тоогоор илэрхийлэгдэх бол тооцоог хялбарчлах зорилгоор (10.15) томъёотой эквивалент дараах томъёог ашигладаг.

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{k} \left(\frac{x_i - C}{m} \cdot n_i \right) \cdot m + C. \qquad (10.16)$$

Үүнд, m-зэрэгцээ вариантуудын хоорондох интервал буюу таблицын алхам, C-дурын тогтмол тоо. (10.16) томъёог вариацын цуваа зөвхөн таблицын тогтмол алхамтай үед хэрэглэх ба хамгийн их давтамжтай вариантыг C тоогоор сонгон авна. Үүний адилаар (10.14) ба (10.15) томъёог дараах байдлаар хялбарчилж болно.

$$\begin{aligned} \overline{X} &= \frac{1}{n}\sum_{i=1}^{n}(x_i - C) + C, \\ \overline{X} &= \frac{1}{n}\sum_{i=1}^{k}(x_i - C) \cdot n_i + C. \end{aligned} \qquad (10.17)$$

Үүнд, C тогтмолыг сонгон авахдаа түүврийн утгууд орших интервалын дундажтай ойрхон утгаар сонгоно. (10.16) ба (10.17) томъёогоор түүврийн дунджийг бодох нь абсолют хэмжигдэхүүнээрээ харьцангуй бага тоонууд дээр ажиллаж буйгаараа давуутай юм.

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{X})^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{X}^2. \qquad (10.18)$$

(10.18) хэмжигдэхүүнийг түүврийн дисперс гэж нэрлэнэ. Вариацын цувааны хувьд түүврийн дисперсийг бичвэл:

$$S^2 = \frac{1}{n}\sum_{i=1}^{k}(x_i - \overline{X})^2 n_i = \frac{1}{n}\sum_{i=1}^{k} x_i^2 n_i - \overline{X}^2. \qquad (10.19)$$

Түүврийн дисперс олох томъёог (10.16) томъёоны адилаар хувиргавал:

$$S^2 = \frac{1}{n}\sum_{i=1}^{k}\left[\left(\frac{x_i - C}{m}\right)^2 \cdot n_i\right] \cdot m^2 - (\overline{X} - C)^2. \qquad (10.20)$$

Мөн (10.18), (10.19) томъёог тооцоо хийхэд хялбар байдлаар бичвэл:

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - C)^2 - (\overline{X} - C)^2 \qquad (10.21)$$

$$S^2 = \frac{1}{n}\sum_{i=1}^{k}(x_i - C)^2 n_i - (\overline{X} - C)^2 \qquad (10.22)$$

$S = \sqrt{S^2}$ хэмжигдэхүүнийг түүврийн дундаж квадрат хазайлт буюу стандарт хазайлт гэнэ.
Эх олонлогийн тоон үзүүлэлтүүдийн нэг нь моод ба медиан болохыг бид мэдэх билээ. Хэрэв дискрет вариацын цуваа авч үзэж байгаа бол хамгийн их давтамжтай вариант нь түүврийн моод болно. Харин интервалын вариацын цувааны хувьд моодыг дараах ойролцоо томъёогоор олно.

$$\overline{M}_0 = x_0 + m \cdot \frac{n_i - n_{i-1}}{(n_i - n_{i-1}) + (n_i - n_{i+1})} \qquad (10.23)$$

Үүнд, x_0- хамгийн их давтамжтай интервалын (моодын интервал гэнэ) эхлэл, m-таблицын алхам, n_i-хамгийн их давтамж, n_{i-1}, n_{i+1}-хамгийн их давтамж бүхий интервалтай хөрш интервалын давтамжууд.

Дискрет вариацын цуваа өгөгдсөн үед, хэрэв түүврийн хэмжээ сондгой бол вариацын цувааны голын элемент (вариант) нь тухайн түүврийн медиан болно. Интервалын вариацын цувааны медианыг дараах ойролцоо томъёогоор олно.

$$\overline{m}_e = x_0 + m\frac{n/2 - (n_1 + n_2 + \cdots + n_{i-1})}{n_i} \qquad (10.24)$$

Үүнд, x_0-голын элементийг агуулсан интервалын (медианы интервал гэнэ) эхлэл, m-таблицын алхам, n-түүврийн хэмжээ, n_i-медианы интервалын давтамж.

Жишээ-10.3 Өмнөх зүйлийн Жишээ-10.1-д авч үзсэн A-түүврийн тоон үзүүлэлтүүдийг бод.

Түүврийн дундаж ба дисперсийг олохын тулд (10.16) ба (10.20) томъёог ашиглая. Үүний тулд $m = 1$, $C = 3$ болохыг анхааран дараах таблицыг зохиоё.

x_i	n_i	$\frac{x_i-C}{m}$	$\frac{x_i-C}{m}n_i$	$\left(\frac{x_i-C}{m}\right)^2$	$\left(\frac{x_i-C}{m}\right)^2 n_i$
0	4	-3	-12	9	36
1	13	-2	-26	4	52
2	14	-1	-14	1	14
3	24	0	0	0	0
4	16	1	16	1	16
5	3	2	6	4	12
6	3	3	9	9	27
7	2	4	8	16	32
$\sum$	79	-	-13	-	189

(10.16), (10.20) томъёо ёсоор түүврийн дундаж:

$$\overline{X} = \frac{-13}{79} \cdot 1 + 3 = -0.16 + 3 = 2.84$$

түүврийн дисперс: $S^2 = \frac{189}{79} \cdot 1 - (2.84 - 3)^2 = 2.3668$ болно. Тэгвэл түүврийн стандарт хазайлт $S = \sqrt{2.3668} = 1.54$. Харин моод ба медиан нь харгалзан $\overline{M}_0 = 3, \quad \overline{m}_0 = 3$ болно.

Жишээ-10.4 Жишээ10.2-д авч үзсэн B түүврийн тоон үзүүлэлтүүдийг ол.

Энэ тохиолдолд $C = 70, \ m = 2$ тул (10.16), (10.20) томъёог хэрэглэхэд дөхөмтэй болгох зорилгоор дараах хүснэгтийг ашиглая.

интервал	интер валын дундаж.	n_i	$\frac{x_i - C}{m}$	$\frac{x_i-C}{m}n_i$	$\left(\frac{x_i-C}{m}\right)^2$	$\left(\frac{x_i-C}{m}\right)^2 n_i$
$[59, 61[$	60	1	-5	-5	25	25
$[61, 63[$	62	2	-4	-8	16	32
$[63, 65[$	64	7	-3	-21	9	63
$[65, 67[$	66	16	-2	-32	4	64
$[67, 69[$	68	27	-1	-27	1	27
$[69, 71[$	70	40	0	0	0	0
$[71, 73[$	72	38	1	38	1	38
$[73, 75[$	74	38	2	76	4	152
$[75, 77[$	76	18	3	54	9	164
$[77, 79[$	78	9	4	36	16	144
$[79, 81[$	80	3	5	15	25	75
$[81, 83]$	82	1	6	6	36	36
$\sum$	-	200	-	132	-	818

Хүснэгт ашиглан түүврийн дундаж ба дисперсийг олбол:

$$\overline{X} = \frac{132}{200} \cdot 2 + 70 = 71.32, \quad S^2 = \frac{818}{200} \cdot 4 - (71.32 - 70)^2 = 14.6176,$$

$S = \sqrt{14.6176} = 3.82$. Одоо (10.23), (10.24) томъёогоор түүврийн моод ба медианыг олъё.

$$\overline{M}_0 = 69 + 2 \cdot \frac{40-27}{(40-27)+(40-38)} = 69 + 2 \cdot \frac{13}{13+2} = 70.7,$$

$$\overline{m}_0 = 71 + 2 \cdot \frac{100-(1+2+7+16+27+40)}{38} = 71,4.$$

Түүврийн дундаж ба дисперс нь түүврийн дээд эрэмбийн моментуудын тухайн тохиолдлууд билээ.
Түүврийн q-эрэмбийн анхны момент дараах томъёогоор бодогдоно.

$$\overline{\nu}_q = \overline{X}^q = \frac{1}{n}\sum_{i8=1}^{n} x_i^q = \frac{1}{n}\sum_{i=1}^{k} x_i^q n_i \tag{10.25}$$

$q=0,\ q=1,\ q=2$ үед харгалзан $\overline{\nu}_0 = 1,\ \overline{\nu}_1 = \overline{X},\ \overline{\nu}_2 = \overline{X}^2 = \frac{1}{n}\sum\limits_{i=1}^{k} x_i^2 n_i$ болно.
Түүврийн q- эрэмбийн төвийн момент:

$$\overline{\mu}_q = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{X})^q = \frac{1}{n}\sum_{i=1}^{k}(x_i - \overline{X})^q \cdot n_i. \tag{10.26}$$

$\overline{\mu}_0 = 1,\ \overline{\mu}_1 = 0,\ \overline{\mu}_2 = S^2$ байхыг бид мэднэ.

Түүврийн q эрэмбийн төвийн момент олох тооцоог хялбарчилах зорилгоор

$$\overline{\mu}_q = \frac{m^q}{n}\sum_{i=1}^{k}\left(\frac{x_i - C}{m} - \frac{\overline{X}-C}{m}\right)^q \cdot n_i \tag{10.27}$$

томъёог ашиглана. (10.27) томъёог шууд хэрэглэхгүйгээр дотор нь задалж, 2 шаттай хэрэглэж болно. Үүний тулд анхны x_i вариантыг $x_i' = \dfrac{x_i - C}{m}$ вариантаар сольж x_i' вариантуудын хувьд q эрэмбийн төвийн момент $\overline{\mu}_q'$-ийг олоод дараа нь

$$\overline{\mu}_q = \overline{\mu}_q' \cdot m^q \tag{10.28}$$

томъёогоор $\overline{\mu}_q$-г олдог. Ингэж, (10.27) томъёог 2 шаттай хэрэглэх нь түүврийн дээд эрэмбийн төвийн моментуудыг, дээд эрэмбийн харгалзах анхны моментуудаар олох боломжийг олгож буй юм. Энэ зорилгоор $q = 0,\ 1,\ 2,\ 3,\ 4$ тохиолдол бүрд түүврийн q эрэмбийн төвийн ба анхны моментуудын холбоог илэрхийлбэл:

$$\begin{aligned}
&\overline{\mu}_0 = \overline{\nu}_0 = 1,\ \ \overline{\mu}_1 = \overline{\nu}_1 - \overline{\nu}_1 = 0,\\
&\overline{\mu}_2 = \overline{\nu}_1 - (\overline{\nu}_1)^2,\ \ \overline{\mu}_3 = \overline{\nu}_3 - 3\overline{\nu}_1 \cdot \overline{\nu}_2 + 2(\overline{\nu}_1)^3,\\
&\overline{\mu}_4 = \overline{\nu}_4 - 4\overline{\nu}_1 \cdot \overline{\nu}_3 + 6(\overline{\nu}_1)^2 \cdot \overline{\nu}_2 - 3(\overline{\nu})^4.
\end{aligned} \tag{10.29}$$

(10.27)÷(10.29) томъёог практикт хэрхэн ашиглахыг үзүүлэх зорилгоор 10.2-р зүйлийн жишээнд авч үзсэн A-түүврийн 2, 3, 4-р эрэмбийн төвийн моментуудыг олъё.

x_i	n_i	$\frac{x_i-C}{m}$	$\frac{x_i-C}{m}n_i$	$\left(\frac{x_i-C}{m}\right)^2$	$\left(\frac{x_i-C}{m}\right)^2 n_i$	$\left(\frac{x_i-C}{m}\right)^3$	$\left(\frac{x_i-C}{m}\right)^3 n_i$
0	4	-3	-12	9	36	-27	-108
1	13	-2	-26	4	52	-8	-104
2	14	-1	-14	1	14	-1	-14
3	24	0	0	0	0	0	0
4	16	1	16	1	16	1	16
5	3	2	6	4	12	8	24
6	3	3	9	9	27	27	81
7	2	4	8	16	32	64	128
$\sum$	79	-	-73	-	189	-	23

x_i	n_i	$\left(\frac{x_i-C}{m}\right)^4$	$\left(\frac{x_i-C}{m}\right)^4 n_i$
0	4	81	324
1	13	16	208
2	14	1	14
3	24	0	0
4	16	1	16
5	3	16	48
6	3	81	243
7	2	256	512
$\sum$	79	-	1365

$x_i' = \dfrac{x_i - C}{m}$ вариантуудын хувьд анхны моментуудыг олбол

$\overline{\nu}_1' = \frac{1}{n}\sum\limits_{i=1}^{k} x_i' n_i = \frac{-13}{79}$,

$\overline{\nu}_2' = \frac{1}{n}\sum\limits_{i=1}^{k} (x_i')^2 n_i = \frac{189}{79}$,

$$\overline{\nu}_3' = \frac{1}{n}\sum_{i=1}^{k}(x_i')^3 n_i = \frac{23}{79}, \qquad \overline{\nu}_4' = \frac{1}{n}\sum_{i=1}^{k}(x_i')n_i = \frac{1365}{79}.$$

Төвийн моментуудыг x_i' вариантуудын хувьд (10.29) томъёогоор олбол:

$$\overline{\mu}_2' = \overline{\nu}_1' - (\overline{\nu}_1')^2 \approx 2.36 \qquad \overline{\mu}_3' = \overline{\nu}_3' - 3\overline{\nu}_1' \cdot \overline{\nu}_2' + 2(\overline{\nu}_1')^3 \approx 1.45,$$

$$\overline{\mu}_4' = \overline{\nu}_4' - 4\overline{\nu}_1' \cdot \overline{\nu}_3' + 6(\overline{\nu}_1')^2 \overline{\nu}_2' - 3(\overline{\nu}_1')^4 \approx 17,84.$$

(10.28) томъёо ёсоор анхны түүврийн төвийн моментуудыг олбол $\overline{\mu}_2 = \overline{\mu}_2' \cdot 1^2 \approx 2.36$, $\overline{\mu}_3 = \overline{\mu}_3' \cdot 1^3 \approx 1.45$, $\overline{\mu}_4 = \overline{\mu}_4' \cdot 1^4 \approx 17.84$ болно.

Түүврийн 3-р эрэмбийн төвийн моментыг стандарт хазайлтын кубд харьцуулсан харьцааг түүврийн асимметрийн коэффициент гэнэ.

$$\overline{A} = \frac{\overline{\mu}_3}{S^3} = \frac{\overline{\mu}_3}{\sqrt{(\overline{\mu}_2)^3}} = \frac{\sum\limits_{i=1}^{k}(x_i - \overline{X})^3 \cdot n_i}{n \cdot S^3} \tag{10.30}$$

(10.30) томъёоноос үзвэл, вариацын цуваанд $\overline{X}$-аас бага вариантуудын тоо давамгайлбал $\overline{A}$ коэффициент сөрөг гарна. Энэ тохиолдолд зүүн өрөөсгөл асимметртэй байна гэж ярих ба полигоны зүүн салаа баруунаасаа урт байна. Харин вариацын цуваанд $\overline{X}$-аас их вариантуудын тоо давамгайлбал $\overline{A}$ коэффициент эерэг гарна. Энэ тохиолдолд баруун өрөөсгөл асимметртэй байна гэж ярих ба полигоны баруун салаа зүүнээсээ урт байна.

Түүврийн 4-р эрэмбийн төвийн моментыг стандарт хазайлтын 4-р зэрэгт харьцуулсан харьцааг 3-аар хорогдуулсныг түүврийн эксцесс (эксцессийн коэффициент) гэнэ.

$$\overline{E} = \frac{\overline{\mu}_4}{S^4} - 3 = \frac{\overline{\mu}_4}{\mu_2^2} - 3 \tag{10.31}$$

Хэрэв $\overline{E} < 0$ бол тархалтын муруй "хавтгай" оройтой, $\overline{E} > 0$ бол "хурц" оройтой байна. Хэвийн тархалтын хувьд $A = E = 0$ байх ба хэвийн тархалтын муруйтай харьцуулах зорилгоор A, E коэффициентүүдыг авч үздэг болохыг 6.4-р зүйлд дурдсан билээ.

10.4 Тархалтын үл мэдэгдэх параметрүүдийн цэгэн үнэлэлт

Их тооны хуульд үндэслэн, мэдэгдэхгүй байгаа онолын тархалтын функцийн оронд ажиглагдсан утгуудын тусламжтайгаар байгуулсан туршилтын тархалтын функцийг авч болох боловч практикт ихэвчлэн $F(x)$ функцийн аналитик хэлбэр мэдэгдээд хэд хэдэн үл мэдэгдэх праметрийг агуулсан байдаг. Өөрөөр хэлбэл, онолын тархалтын функц

$$F(x) = F(x, a_1, a_2, \dots, a_k) \tag{10.32}$$

хэлбэртэй байна. Ийнхүү эх олонлогийн тархалтын функц $F(x)$-ийг олох асуудал $a_1, a_2, \dots, a_k$ параметрүүдийг үнэлэх асуудалд шилжинэ. Үнэлэх гэдэг нь ажиглалтын түүврийн утгуудаар эх олонлогийн тархалтын параметрүүдийг яг тодорхойлж буй хэрэг биш харин түүний үнэлэлт гэж нэрлэгдэх ойролцоо утгыг олохыг хэлж буй юм.
X гэсэн санамсаргүй хэмжигдэхүүн дээр туршилт (ажиглалт) хийсэн гэе. Нэг удаагийн туршилтын утгууд $x_1, x_2, \dots, x_n$ нь n хэмжээст түүвэр болох бөгөөд тус бүр нь n хэмжээтэй түүврүүдийн олонлог авч үзье. i-р түүврийн тусламжтайгаар олсон a параметрийн үнэлэлтийг $\overline{a}_i$ гэвэл түүврийн бүрэлдэхүүн санамсаргүй учраас $\overline{a}_i$ өөрөө санамсаргүй хэмжигдэхүүн байна. Ийнхүү эх олонлогийн n хэмжээст түүврийн санамсаргүй шинж чанарыг тусгахын тулд ажиглалтын утга бүрийг санамсаргүй хэмжигдэхүүн мэтээр үзнэ. Өөрөөр хэлбэл, түүврийн утгууд x_i бүр нь нэг ижил (эх олонлогтой адил) тархалттай, үл хамаарах санамсаргүй хэмжигдэхүүн юм. Тэгвэл эх олонлогийн a_i параметрийн үнэлэлт $\overline{a}_i$ нь түүврийн утгуудаас хамаарсан функц

болно.

$$\overline{a}_i = F(x_1, x_2, \ldots, x_n) \tag{10.33}$$

Одоо X санамсаргүй хэмжигдэхүүний тархалтын функцийн ямар нэг параметр $a = a(x_1, x_2, \ldots, x_n)$-г үнэлэх зорилго тавья. Аливаа параметрийн үнэлэлтийг хазайлтгүй, эрчимтэй, зохимжтой гэж ангилна.

a параметрийн үнэлэлт $\overline{a}$-ийн хувьд

$$M(\overline{a}) = a \tag{10.34}$$

нөхцөл биелэгдэж байвал $\overline{a}$ үнэлэлтийг хазайлтгүй үнэлэлт гэнэ. Эсрэг тохиолдолд ($M\overline{a} < a, \quad M\overline{a} > a$) хазайлттай гэнэ. Параметрийн үнэлэлт хазайлтгүй байна гэдэг нь уг параметрийг үнэлэхэд санамсаргүй биш байнгын (хүн ба багажнаас хамаарах) алдаа гараагүй болохыг илэрхийлнэ.

$\overline{a}$ үнэлэлт a параметрийн хазайлтгүй үнэлэлт байхын тулд $\overline{a}$ утга түүврээс түүвэрт шилжихэд түүний математик дундаж нь a-ийн орчинд хэлбэлзэх ёстой. Үүнийг илэрхийлэх хэмжигдэхүүн нь $D\overline{a} = M(\overline{a} - a)^2$ юм. Ийнхүү хазайлтгүй үнэлэлтүүд ч үнэлж буй параметрийн жинхэнэ утгаас янз бүрээр сарнина. Үүнтэй уялдан эх олонлогийн янз бүрийн түүврүүдийн хувьд a параметрийн бүх боломжит хазайлтгүй үнэлэлтүүдийн дотроос хамгийн бага дисперстэй үнэлэлтийг эрчимтэй гэнэ. Өөрөөр хэлбэл, $M(\overline{a}_1 - a)^2 < M(\overline{a}_2 - a)^2$ бол $\overline{a}_1$-г a_2-оос эрчимтэй гэнэ. Үнэлэлт эрчимтэй байна гэдэг нь уг чанартаа, тухайн үнэлэлт хамгийн бага санамсаргүй алдаатай байна гэсэн үг юм. Үнэлэлт бүрийн хувьд эрчимтэй үнэлэлт оршин байх нь албагүй.

Хэрэв a параметрийн үнэлэлт $\overline{a}$ их тооны хуульд захирагдаж байвал $\overline{a}$-г зохимжтой үнэлэлт гэнэ. Өөрөөр хэлбэл, $\lim\limits_{n\to\infty} P(|\overline{a} - a| < \varepsilon) = 1$ нөхцөл биелэнэ гэсэн үг. Энэ нь түүврийн хэмжээ хүрэлцээтэй их байхад, үл мэдэгдэх параметрийг үнэлэхэд гарах алдаа багасахыг илэрхийлнэ.

Хэрэв a нь $F(x, a)$ тархалтын функц бүхий санамсаргүй хэмжигдэхүүний математик дундаж юмуу эсвэл дисперс бол туршилтаар олсон түүврийн дундаж, эсвэл түүврийн дисперсийг a-ийн үнэлэлт болгон сонгон авна.

Түүврийн дундаж $\overline{X} = \dfrac{1}{n}\sum\limits_{i=1}^{n} x_i$ (10.14) нь эх олонлогийн математик дунджийн хазайлтгүй, зохимжтой үнэлэлт болно. Харин түүврийн дисперс $S^2 = \dfrac{1}{n}\sum\limits_{i=1}^{n}(x_i - \overline{X})^2$ (10.18) нь эх олонлогийн дисперсийн хазайлттай боловч зохимжтой үнэлэлт болно. Түүврийн дисперсийн оронд

$$S^2 = \frac{n}{n-1}S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{X})^2 \tag{10.35}$$

үнэлэлтийг авбал эх олонлогийн дисперсийн хазайлтгүй, зохимжтой үнэлэлт болох ба түүнийг түүврийн засварласан дисперс гэнэ.

$\dfrac{n}{n-1}$-коэффициентийг Бесселийн засвар гэнэ.

Эх олонлогийн мат.дундаж MX мэдэгдэж буй тохиолдолд

$$S_*^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - MX)^2 \tag{10.36}$$

үнэлэлт эх олонлогийн дисперсийн хазайлтгүй, зохимжтой үнэлэлт болно. Эх олонлог зөвхөн хэвийн тархалттай тохиолдолд $\overline{X}$, S_*^2 үнэлэлтүүд эрчимтэй байна. Дээр авч үзсэн үнэлэлтүүд нь үл мэдэгдэх параметр a-г зөвхөн ганц $\overline{a}$ тоогоор үнэлж буй тул цэгэн үнэлэлт гэж нэрлэгдэнэ.

10.5 Хамгийн их үнэний хувь бүхий арга

Тархалтын үл мэдэгдэх параметрийн үнэлэлтийг олох практикт өргөн нэвтэрсэн аргуудын нэг бол хамгийн их үнэний хувь бүхий арга юм. Энэ аргын үндсэн санаа нь, n-хэмжээст санамсаргүй түүврийн утгуудаар, хамгийн их үнэний хувь бүхий гэж нэрлэгдэх функцийн утга хамгийн их байхаар тархалтын үл мэдэгдэх параметрүүдийг олоход оршино.

Санамсаргүй хэмжигдэхүүн X-ийн онолын тархалтын нягт $f(x, a)$ хэлбэртэй байг. a үл мэдэгдэх параметр. Түүнчлэн n хэмжээст түүврийн утгууд $x_1, x_2, \ldots, x_n$ мэдэгдэж байг. Дурын бага эерэг тоо ε-г авч $(x_i - \frac{\varepsilon}{2},\ \ x_i + \frac{\varepsilon}{2})$ гэсэн, x_i цэгийн $\frac{\varepsilon}{2}$ радиустай орчныг байгуулъя $(i = \overline{1, n})$. Тэгвэл санамсаргүй хэмжигдэхүүн $(x_i - \frac{\varepsilon}{2};\ \ x_i + \frac{\varepsilon}{2})$ завсраас утгаа авах магадлал ойролцоогоор $f(x_i, a) \cdot \varepsilon$ болно. Хэрэв n удаа ажиглалт хийсэн бол 1-р ажиглалтын үр дүн 1-р завсарт байх, 2-р ажиглалтын үр дүн 2-р завсарт байх гэх мэтчилэн n-р ажиглалтын үр дүн n-р завсарт байх үзэгдлүүд зэрэг явагдах магадлал буюу хамтын тархалтын нягт нь

$$P(x_1, x_2, \ldots, x_n, a) = f(x_1, a) \cdot f(x_2, a), \ldots, f(x_n, a) \cdot \varepsilon^n \tag{10.37}$$

болно. Энэ функц экстремумтай байх зайлшгүй нөхцөл нь

$$\frac{\partial P(x_1, x_2, \ldots, x_n, a)}{\partial a} = 0 \tag{10.38}$$

бөгөөд a цэг дээр максимумтай байх хүрэлцээтэй нөхцөлийг бичвэл

$$\frac{\partial P(x_1, x_2, \ldots, x_n, a)}{\partial a^2} < 0 \tag{10.39}$$

Хэрэв хэд хэдэн цэг дээр максимумтай бол тэдгээрийн дотроос функц хамгийн их утгатайг нь сонгоно.

$$L(x_1, x_2, \ldots, x_n, a) = Ln\frac{P(x_1, x_2, \ldots, x_n, a)}{\varepsilon^n} = \sum_{i=1}^{n} Lnf(x_i, a) \tag{10.40}$$

томъёогоор тодорхойлогдох функцийг хамгийн их үнэний хувь бүхий функц гэнэ.

$$\frac{\partial L}{\partial a} = \frac{\partial (LnP)}{\partial a} = \frac{1}{P}\cdot\frac{\partial P}{\partial a}, \quad (P>0) \tag{10.41}$$

нөхцөл биелэгдэх учраас $P(x_1, x_2, \ldots, x_n, a)$ ба $L(x_1, x_2, \ldots, x_n, a)$ функцийн максимумын цэгүүд ижил байна. Ерөнхий тохиолдолд $L(x_1, x_2, \ldots, x_n, a)$ функц $a_1, a_2, \ldots, a_k$ гэсэн k параметраас хамаарч болох бөгөөд энэ тохиолдолд экстремум орших нөхцөлийг бичвэл:

$$\frac{\partial L(x_1, x_2 \ldots, x_n, a_1, a_2, \ldots, a_k)}{\partial a_j} = 0 \quad (j=\overline{1,k}) \tag{10.42}$$

Хэрэв X нь x_i утгуудыг $p_i(a) \quad (i=\overline{1,n})$ магадлалтай хүлээж авах дискрет санамсаргүй хэмжигдэхүүн бол хамгийн их үнэний хувь бүхий функц

$$L(x_1, x_2, \ldots, x_n, a) = p_1^{q_1}(a)p_2^{q_2}(a)\ldots p_k^{q_k}(a) \tag{10.43}$$

хэлбэртэй байна. Үүнд, q_i $(i=\overline{1,k})$ нь түүвэр дэх x_i утгын давтамж.

Жишээ-10.6 Түүврийн $x_1, x_2, \ldots, x_n$ утгуудаар $f(x) = \lambda e^{-\lambda x}$ нягт бүхий илтгэгч тархалтын параметр λ-г хамгийн их үнэний хувь бүхий аргаар үнэл.

Хамгийн их үнэний хувь бүхий функцийг бичвэл

$$L(x_1, x_2, \ldots, x_n, \lambda) = \sum_{i=1}^{n} Ln(\lambda e^{-\lambda x_i}) = n\cdot Ln\lambda - \lambda\sum_{i=1}^{n} x_i$$

λ-аар дифференциалчилбал:

$$\frac{\partial L}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i = 0 \Rightarrow \lambda = \frac{n}{\sum\limits_{i=1}^{n} x_i} = \frac{1}{\overline{X}} \quad \text{буюу} \quad \overline{X} = \frac{1}{\lambda}.$$

Өөрөөр хэлбэл, түүврийн дунджийн урвуу нь λ параметрийн хамгийн оновчтой үнэлэлт болно.

Жишээ-10.7 Хэвийн тархалттай санамсаргүй хэмжигдэхүүний a, σ^2 параметрүүдийг хамгийн их үнэний хувь бүхий аргаар үнэл.

Хамтын нягт

$P(x_1, x_2, \ldots, x_n, a, \sigma^2) = \dfrac{\varepsilon^n}{(2\pi\sigma^2)^{n/2}} e^{-\frac{1}{2\sigma^2}\sum\limits_{i=1}^{n}(x_i-a)^2}$ учир

$$L(x_1, x_2, \ldots, x_n, a, \sigma^2) = -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-a)^2$$

a ба σ^2-аар дифференциалчилбал:

$$\begin{cases} \dfrac{\partial L}{\partial a} = \dfrac{-2}{2\sigma^2}\sum\limits_{i=1}^{n}(x_i - a) = 0 \\ \dfrac{\partial L}{\partial \sigma^2} = \dfrac{-n}{2\sigma^2} + \dfrac{1}{2(\sigma^2)^2}\sum\limits_{i=1}^{n}(x_i - a)^2 = 0 \end{cases}$$

Тэгшитгэл тус бүрээс a ба σ^2-ийг олбол:

$$a = \frac{1}{n}\sum_{i=1}^{n} x_i = \overline{X}, \qquad \sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{X})^2 = S^2$$

Эндээс үзвэл, түүврийн дундаж ба түүврийн дисперс нь хэвийн тархалтын параметрүүдийн хамгийн оновчтой үнэлэлтүүд болж байна.

10.6 Параметрийн завсран үнэлэлт. Итгэх магадлал, итгэх завсар.

Туршилтын тоо хүрэлцээтэй их байхад цэгэн үнэлэлт үл мэдэгдэх параметрийн жинхэнэ утгаас бага ялгагдах боловч туршилтын тоо цөөн тохиолдолд $\overline{a}$-ийн санамсаргүй шинж чанараас хамааран a, $\overline{a}$ утгуудын хооронд мэдэгдэхүйц зөрөө гарна. Үүнээс үндэслэн үл мэдэгдэх параметр a-г зөвхөн ганц a утгаар биш ямар нэг интервалаар үнэлэх шаардлага гардаг. Ийм үнэлэлтийг завсран (интервалан) үнэлэлт гэнэ.
Үл мэдэгдэх параметр a-ийн хазайлтгүй үнэлэлт $\overline{a}$-г түүврийн утгуудаар гарган авсан гэе. a параметрийг $\overline{a}$-аар солиход гарч болох алдааг үнэлэхийн тулд p-магадлалтайгаар бараг гарцаагүй явагдах дараах үзэгдлийн магадлалаас δ-утгыг тодорхойлъё.

$$P(|a - \overline{a}| < \delta) = p. \tag{10.44}$$

Энэхүү δ тоо нь үнэлгээний нарийвчлалыг илэрхийлнэ. (10.44) нөхцөлийг хангах p магадлалыг итгэх магадлал буюу найдвар гэнэ. Өөрөөр хэлбэл, үл мэдэгдэх параметр a нь $(a - \overline{\delta}, a + \overline{\delta})$ завсраас утгаа авах магадлал юм. p нь урьдчилан өгөгдсөн байх бөгөөд практикт 0.9, 0.95, 0.99, 0.975, 0.999 гэсэн тоонуудыг ихэвчлэн авдаг. Ийнхүү эх олонлогийн үл мэдэгдэх параметр a-ийн жинхэнэ утгыг p магадлалтайгаар агуулсан завсрыг түүний итгэх завсар, p тоог итгэх магадлал, $\alpha = 1 - p$ тоог итгэх түвшин гэнэ.
(10.44)-өөс $P(a - \overline{\delta} < a < a + \overline{\delta}) = p$ буюу $P(\overline{a}_1 < a < \overline{a}_2) = p$ болно. $\overline{a}_1$, $\overline{a}_2$-г итгэх завсрын хилүүд гэх ба эдгээр нь ажиглалтын үр дүнгээр тодорхойлогдох санамсаргүй хэмжигдэхүүн юм. Итгэх завсрын хил $\overline{a}_1$, $\overline{a}_2$-г итгэх магадлал p-ийн тусламжтай олохдоо, тархалт нь a болон бусад үл мэдэгдэх параметрүүдээс хамаарахгүй, зөвхөн түүврийн утгаас хамаарсан

функц мэтээр үзэж олно.
Тухайлбал, $\overline{a}$ үнэлэлтийн тархалтын хууль мэдэгдэж буй бол итгэх завсрын хилийг дараах нөхцөлөөс олно.

$$P(\overline{a}-\delta<a<\overline{a}+\delta)=F(\delta)-F(-\delta)=\int\limits_{-\delta}^{\delta} f(x)dx=p. \tag{10.45}$$

Энд жишээ болгож a, σ^2 параметр бүхий хэвийн тархалттай санамсаргүй хэмжигдэхүүний дисперс (σ^2) мэдэгдэж буй тохиолдолд түүний математик дундаж a-ийн итгэх завсрыг хэрхэн байгуулахыг авч үзье.
n хэмжээст түүврийн утга $x_1, x_2, \ldots, x_n$ тус бүр нь a, σ^2 параметр бүхий хэвийн тархалттай тул, $\overline{X} = \dfrac{1}{n}\sum\limits_{i=1}^{n} x_i$ хэмжигдэхүүн нь a, $\dfrac{\sigma^2}{n}$-параметрүүд бүхий хэвийн тархалттай байна ($\overline{X} \sim N(a, \dfrac{\sigma^2}{n})$).
Хэвийн тархалттай хэмжигдэхүүний хувьд $P(|X-a| < \delta) = 2\Phi\left(\dfrac{\delta}{\sigma}\right)$ байдгийг анхаарвал $P(|\overline{X}-a|<\delta)=2\Phi\left(\dfrac{\delta\sqrt{n}}{\sigma}\right)$ нөхцөл биелэнэ.
Үүнд, $\Phi(x)$-Лапласын функц. $\dfrac{\delta\sqrt{n}}{\sigma}=t$ гэвэл
$P(|\overline{X} - a| < \delta) = 2\phi(t)$ буюу $\delta = \dfrac{\sigma t}{\sqrt{n}}$ учир $P\left(|\overline{X} - a| < \dfrac{\sigma t}{\sqrt{n}}\right) = 2\Phi(t)$.
Эндээс $P\left(\overline{X} - \dfrac{\sigma t}{\sqrt{n}} < a < \overline{X} + \dfrac{\sigma t}{\sqrt{n}}\right) = 2\Phi(t) \quad (*)$
Гэтэл нөгөө талаас итгэх магадлал p өгөгдсөн тул

$$P\left(\overline{X} - \frac{\sigma t}{\sqrt{n}} < a < \overline{X} + \frac{\sigma t}{\sqrt{n}}\right) = p.$$

Иймд математик дунджийн итгэх завсар нь $\left]\overline{X} - \dfrac{\sigma t}{\sqrt{n}};\ \overline{X} + \dfrac{\sigma t}{\sqrt{n}}\right[$ болно.
Үүнд t-утгыг $2\Phi(t) = p \quad (**)$ нөхцөлөөс олох ба

$$\delta = \frac{\sigma t}{\sqrt{n}} \tag{10.46}$$

нь үнэлэлтийн нарийвчлал юм.

Жишээ-10.8 $\sigma^2=4$ дисперстэй, a гэсэн үл мэдэгдэх математик дундажтай, хэвийн тархалттай санамсаргүй хэмжигдэхүүн дээр 5 удаа туршилт хийсэн дүнг хүснэгтэд өгчээ.

$i:$	1	2	3	4	5
$x_i:$	-25	34	-20	10	21

Санамсаргүй хэмжигдэхүүний математик дунджийн итгэх завсрыг $p = 0.9$ итгэх магадлалтай байгуул.

Түүврийн дунджийг олбол $\overline{X} = \frac{1}{5}(-25 + 34 - 20 + 10 + 21) = 4$, $2\Phi(t) = 0.9$ нөхцөл ба Лапласын функцийн таблицаас $t = 1.65$ гэж олдоно.
$\delta = \frac{\sigma t}{\sqrt{n}} = \frac{1.65 \cdot 2}{\sqrt{5}} = 1.47$ Тэгвэл математик дунджийн итгэх завсар
$\left]\overline{X} - \frac{\sigma t}{\sqrt{n}};\ \overline{X} + \frac{\sigma t}{\sqrt{n}}\right[=]-2.35; 5.47[$ гэсэн интервал болно.
Дээр авч үзсэний адилаар, эх олонлог хэвийн тархалттай нөхцөлд түүний параметрүүдийн завсран үнэлэлтийг хэрхэн байгуулах, тэдгээр нь ямар тархалттай болохыг дараах хүснэгтэд харуулав.

Үнэлэх пара- метр	Ямар нөхцөлд	Пара- метрийн үнэлэлт	Итгэх завсар	Үнэлэлтийн тархалт
a	σ^2- мэдэгдэх	$\overline{X}$	$\overline{X} - \frac{\sigma t}{\sqrt{n}} < a < \overline{X} + \frac{\sigma t}{\sqrt{n}}$	$\overline{X} \sim N(a, \frac{\sigma^2}{n})$
a	σ^2- үл мэдэгдэх	$\overline{X}$	$\overline{X} - \frac{t\widehat{S}}{\sqrt{n}} < a < \overline{X} + \frac{t\widehat{S}}{\sqrt{n}}$	$\frac{\overline{X} - a}{\widehat{S}}\sqrt{n} \sim S_{n-1}(t)$
σ^2	$a-$ мэдэгдэх	S_*^2	$\frac{nS_*^2}{\chi^2_{1-\alpha/2}(n)} < \sigma^2 < \frac{nS_*^2}{\chi^2_{\alpha/2}(n)}$	$S_*^2 \sim \frac{\sigma^2}{n}\chi^2(n)$
σ^2	$a-$ үл мэдэгдэх	$\widehat{S}^2$	$\frac{(n-1)\widehat{S}^2}{\chi^2_{1-\alpha/2}(n-1)} < \sigma^2 < \frac{(n-1)\widehat{S}^2}{\chi^2_{\alpha/2}(n-1)}$	$\widehat{S}^2 \sim \frac{\sigma^2}{n-1}\chi^2(n-1)$

Үүнд, $N\left(a, \frac{\sigma^2}{n}\right) - a; \frac{\sigma^2}{n}$ параметрүүд бүхий хэвийн тархалт,
$S_{n-1}(t) - (n-1)$ чөлөөний зэрэг бүхий Стьюдентийн тархалт,
$\chi^2(n) - n$ чөлөөний зэрэг бүхий хи-квадрат тархалт.

Бином тархалтын p параметрийн завсран үнэлэлт

Бернуллийн схемийн нөхцөлд A үзэгдэл явагдах магадлал $p = P(A)$ мэдэгдэхгүй байг. Туршилтыг n удаа давтан хийхэд A үзэгдэл k удаа явагдсан бол $p^* = \frac{k}{n}$ нь A үзэгдлийн харьцангуй давтамж юм. Тэгвэл p^* нь үл мэдэгдэх магадлал p-ийн цэгэн үнэлэлт болно ($p^* \approx p$). Үл мэдэгдэх p магадлалын итгэх завсар $(\overline{p}_1, \overline{p}_2)$-г өгөгдсөн $(1-\alpha)$ итгэх магадлалтайгаар олох шаардлага практикт олонтаа тааралдана. Хэрэв $n > 50$, $np^* > 5$, $n(1-p^*) > 5$ нөхцөлүүд биелэгдэх бол $Z = \frac{p^* - p}{\sqrt{pq/n}}$ $(q = 1-p)$-санамсаргүй хэмжигдэхүүн $N(0,1)$-тархалттай байдаг. Энэ үр дүнг ашиглан $P(\overline{p}_1 < p < \overline{p}_2) \geq 1-\alpha$ нөхцөлийг хангах $\overline{p}_1, \overline{p}_2$ хилүүдийг дараах томъёогоор олж болно.

$$\overline{p}_{1,2} = \frac{1}{1 + t^2/n}\left(p^* + \frac{t^2}{2n} \mp t \cdot \sqrt{\frac{p^*(1-p^*)}{n} + \frac{t^2}{4n^2}}\right) \tag{10.47}$$

Үүнд t-г $2\Phi(t) = 1 - \alpha$ нөхцөлөөс олно. Зарим тохиолдолд

$$\overline{p}_{1,2} \approx p^* \mp \frac{t}{\sqrt{n}}\sqrt{p^*(1-p^*)} \tag{10.48}$$

томъёог ашиглах ба бином тархалттай санамсаргүй хэмжигдэхүүний хувьд түүврийн хэмжээг тодорхойлоход ихэвчлэн хэрэглэнэ.

Жишээ-10.9 Үйлдвэрлэсэн деталуудаас 100 детал авч шалгахад 10 нь гологдол байв. Тэгвэл нийт деталуудын дотор гологдол деталын эзлэх хувийн итгэх завсрыг $1-\alpha = 0.95$ итгэх магадлалтай байгуул.
$n > 50$, $p^* = 0.1$, $np^* = 10$, $p^*(1-p^*) = 0.09$. Лапласын функцийн таблиц ашиглан $2\Phi(t) = 0.95$ нөхцөлөөс t-г олбол $t = 1.96$. (10.47) томъёогоор итгэх завсруудын хилийг олбол $0.055 < p < 0.174$ болно.
Энэ бодлогын хувьд, "нийт бүтээгдэхүүний доторхи гологдол деталийн хувь (магадлал), түүвэр дэх гологдол деталын харьцангүй давтамжаас 0.01-ээс ихгүйгээр хазайна" гэж 0.95 магадлалтайгаар батлан хэлэхийн тулд түүврийн хэмжээ хамгийн багадаа хэд байх вэ? гэдгийг тогтооё.
(10.48) томъёог ашиглан итгэх завсрын хилүүдийг дараах тэнцэтгэл бишээр илэрхийлж болно.

$$|p^* - p| < t \cdot \sqrt{\frac{p^*(1-p^*)}{n}}$$

Өгөгдсөн ёсоор $|p^* - p| \leq 0.01$ тул $t \cdot \sqrt{\dfrac{p^*(1-p^*)}{n}} \leq 0.01$
Эндээс түүврийн хэмжээг олбол $n \geq (0.3 \cdot 196)^2 = 3457.44$ буюу түүврийн хэмжээ хамгийн багадаа 3458 байхад дээрх нотолгоог хэлж болно гэсэн үг.
Санамж: $n < 50$ юмуу эсвэл $n(1-p^*)$, np^* тоонуудын аль нэг нь 5-аас бага бол (10.47), (10.48) томъёог үл хэрэглэнэ.

Пуассоны тархалтын λ параметрийн завсран үнэлэлт

x нь λ параметр бүхий Пуассоны тархалттай санамсаргүй хэмжигдэхүүний ажиглагдсан утга байг. λ параметрийн итгэх завсрын хилүүд $\overline{\lambda}_1, \overline{\lambda}_2$-ыг олохдоо Пуассоны тархалт хязгаартаа χ^2-тархалт уруу нийлдэг болохыг ашиглана.

$$\begin{aligned} \overline{\lambda}_1 &= \frac{1}{2}\chi^2_{\alpha/2}(2x) \\ \overline{\lambda}_2 &= \frac{1}{2}\chi^2_{1-\alpha/2}(2x+2) \end{aligned} \tag{10.49}$$

Практикт λ параметрийн үнэлэлтийг олохдоо эхлээд
$x = \sum_{i=1}^{n} x_i$ хэмжигдэхүүний $n\lambda$ -параметрийн итгэх завсрын хилүүдийг олно.
$n\lambda$ параметрийн итгэх завсрын хилүүд нь $\overline{\lambda}_1 < n\lambda < \overline{\lambda}_2$ бол λ параметрүүдийн итгэх хилүүд

$$\frac{\overline{\lambda_1}}{n} < \lambda < \frac{\overline{\lambda_2}}{n} \tag{10.50}$$

болно.

Жишээ-10.10 Шинэ деталыг турших явцад A деталын хувьд 2 эвдрэл, B ба C деталуудын хувьд тус бүр нэг эвдрэл бүртгэгджээ. Аль ч деталын хувьд эвдрэл бүртгэгдэх тоо нь λ параметр бүхий Пуассоны тархалттай гэвэл параметрийн итгэх завсрыг $p = 0.9$ итгэх магадлалтайгаар байгуул. Ажиглалтаар тогтоосон эвдрэлийн ерөнхий тоо $x = \sum_{i=1}^{n} x_i = 4$ учир (10.49) томьёо болон $\chi^2(n)$ тархалтын таблицыг ашиглавал $\chi^2_{0.05}(8) = 2.73$, $\chi^2_{0.95}(10) = 18.3$ гэж олдоно. Тэгвэл $n\lambda$ параметрийн итгэх хилүүд $1.365 < 5\lambda < 9.15$ тул λ параметрийн итгэх завсар нь $0.273 < \lambda < 1.83$ болно.

Түүврийн хэмжээг олох

Бид үүнээс өмнө түүврийн хэмжээ n өгөгдсөн үед итгэх завсрыг хэрхэн байгуулах тухай авч үзсэн. Гэтэл анхнаасаа түүврийн хэмжээ ямар байх үед шаардлагатай үр дүнгээ гарган авч болох вэ? гэсэн асуулт гарна. Их тооны хууль ёсоор хүрэлцээтэй их байхаар сонгон авч болох боловч ингэх нь боловсруулалтанд ихээхэн хүндрэлтэй. Ийм учраас, практикт хангалттай үр дүн өгөхөөр хамгийн бага хэмжээг ашиглахыг эрмэлздэг. Тухайлбал, хэвийн тархалтын математик дунджийн итгэх завсрыг байгуулахад олсон (10.46) томьёог ашиглан түүврийн хэмжээг тогтоож болно:

$$n = t^2 \frac{\sigma^2}{\delta^2} \quad (2\Phi(t) = p) \tag{10.51}$$

Энэ томъёоноос үзвэл түүврийн хэмжээ нь σ^2 ба t^2 утгуудад шууд пропорционал, δ^2 утганд урвуу пропорционал байна. Энэ томьёог зөвхөн эх олонлогийн дисперс мэдэгдэж байхад ашиглах нь ойлгомжтой. Хэрэв σ^2 үл мэдэгдэх бол түүврийн хэмжээг

$$n = t_p^2 \frac{\widehat{S}^2}{\delta^2} \tag{10.52}$$

томъёогоор олно. t_p утгыг Стьюдентийн тархалтын таблицаас олно.

(10.51), (10.52) томъёо нь хэвийн тархалттай санамсаргүй хэмжигдэхүүний хувьд буцаалттай түүврийн хэмжээг тодорхойлно. Гэхдээ (10.51) томъёо мэдэгдэж буй σ^2 ба p магадлалаар, туршилт явуулахаас өмнө түүврийн хэмжээг урьдчилан тодорхойлж байгаа бол (10.52) томъёо анхны сорилтын түүвэр дээр тулгуурлан түүврийн хэмжээг тодорхойлж байгаагаараа тус тус ялгаатай юм.

10.7 Статистик таамаглал шалгах шинжүүрүүд

Байгал шинжлэл, техник, эдийн засгийн бодлогуудад авч үзэж буй санамсаргүй хэмжигдэхүүний тоон үзүүлэлтүүдийн хоорондын харьцаа болон тархалтын хуулийн хэлбэрийн тухай статистик таамаглал дэвшүүлэх шаардлага ямагт гардаг. Таамаглал гэдэг нь эсвэл үнэн эсвэл худал байх ямар нэг урьдчилсан дүгнэлт билээ. Мэдэгдэхгүй байгаа тархалтын хуулийн хэлбэрийн тухай юмуу эсвэл хэлбэр нь мэдэгдэж буй тархалтын хуулийн параметрүүдийн тухай таамаглалыг статистик таамаглал гэнэ. Энэ тодорхойлолтыг үндэслэн статистик таамаглалуудыг тархалтын параметрүүдийн тухай таамаглал, тархалтын хуулийн тухай таамаглал гэж ерөнхийд нь 2 ангилж болно. Жишээлбэл, " техник, технологийн адил нөхцөлд нэг ижил ажил үүрэг гүйцэтгэж байгаа ажилчдын хөдөлмөрийн бүтээмж хэвийн тархалттай " гэсэн таамаглал нь тархалтын хуулийн хэлбэрийн тухай таамаглал юм. Харин " ижил суурь машинууд дээр зорогдсон нэг хэлбэрийн деталуудын дундаж хэмжээ өөр хоорондоо үл ялгагдана" гэсэн таамаглал нь тархалтын параметрийн тухай таамаглал болно.

10.7.1 Тархалтын параметрийн тухай шинжүүр

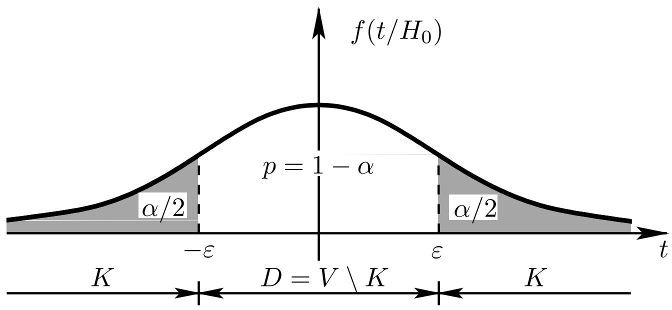
Санамсаргүй хэмжигдэхүүн X-ийн тархалтын функц $F(x, a)$ нь зөвхөн ганц a параметрээс хамаарсан байг.
Тархалтын үл мэдэгдэх параметрийн тухай таамаглал нь энгийн ба нийлмэл гэсэн 2 янз байна. a параметр тодорхой нэг утгатай $(a = a_0)$ гэсэн таамаглал нь энгийн, $a \neq a_0$, $a < a_0$, $a > a_0$ гэсэн таамаглалуудын аль нь ч нийлмэл таамаглал болно. Хэрэв үл мэдэгдэх параметр a тодорхой a_0 гэсэн утгатай тэнцүү гэж үзэх үндэс байгаа бол $a = a_0$ гэсэн таамаглал дэвшүүлж болно. Энэ таамаглалыг тэг таамаглал буюу анхны таамаглал гэж нэрлээд H_0 гэж тэмдэглэе. Математик томъёоллоор $H_0 : a = a_0$ гэж бичнэ. Анхны таамаглалтай харшлах юмуу түүний эсрэг таамаглалыг өрсөлдөгч таамаглал буюу альтернатив таамаглал гэж нэрлээд H_1 гэе. Энэ тохиолдолд өрсөлдөгч таамаглал H_1 нь $H_1 : a \neq a_0$, $H_1 : a < a_0$, $H_1 : a > a_0$ гэсэн 3 янз байж болно. Бодлогын шинж чанараас хамааруулж эдгээрийн аль нэгийг сонгон авна. Бидний зорилго бол эдгээр статистик таамаглалуудыг санамсаргүй түүврийн утгуудаар шалгахад оршино. Эх олонлогийн түүврийн санамсаргүй шинж чанараас хамаарч статистик шалгалтын дүнд хоёр тохиолдолд буруу шийдвэр гаргасан байж болно. Өөрөөр хэлбэл, дараах хоёр төрлийн алдаа гарч болно. Үүнд:
1. Зөв таамаглалыг хэрэгсэхгүй болгох. Энэ нь H_0 таамаглал үнэн байхад H_1 таамаглалыг хүлээн зөвшөөрсөн гэсэн үг юм. Ийм алдааг I төрлийн алдаа гэнэ.
2. Буруу таамаглалыг хүлээн зөвшөөрөх. Өөрөөр хэлбэл, H_1 таамаглал үнэн байхад H_0 таамаглалыг хүлээн зөвшөөрсөн гэсэн үг. Ийм алдааг II төрлийн алдаа гэнэ.

I төрлийн алдаа гарах магадлалыг α гэж тэмдэглээд таамаглал шалгах итгэх түвшин, $p = 1 - \alpha$ тоог статистик найдвар гэж тус тус нэрлэнэ. Практикт техник, эдийн засгийн асуудлуудыг шийдвэрлэхэд ихэвчлэн $\alpha = 0.05$, $\alpha = 0.01$-ээр харин эмнэлэг, анаагаах ухааны судалгаанд $\alpha = 0.001$ гэж авдаг.

Дэвшүүлсэн таамаглалыг шалгахын тулд тархалтын жинхэнэ буюу ойролцоо утга нь урьдчилан мэдэгдэж байгаа тусгайлан сонгон авсан, санамсаргүй хэмжигдэхүүнийг ашигладаг. Энэ хэмжигдэхүүнийг статистик шинжүүр гэнэ (T). Өөрөөр хэлбэл, H_0 таамаглалыг шалгах үүрэг гүйцэтгэх санамсаргүй хэмжигдэхүүнийг статистик шинжүүр гэнэ. I ба II төрлийн алдаа гарах магадлалууд бага байхаар шинжүүрийг сонгон авна. Цаашид ямагт энгийн таамаглал авч үзэх учраас H_0 таамаглал үнэн үед шинжүүрийн тархалтыг мэдэгдэж байгаа гэж үзнэ. Тархалтын параметрийн тухай шинжүүрийг итгэх шинжүүр, тархалтын хуулийн тухай шинжүүрийг зөвшөөрлийн шинжүүр гэж нэрлэнэ.
Тодорхой шинжүүр сонгон авсны дараа түүний боломжит бүх утгуудын олонлогийг (V) үл огтлолцох 2 дэд олонлогт хуваана $(V = D \cup K,\ D \cap K = \varnothing)$.

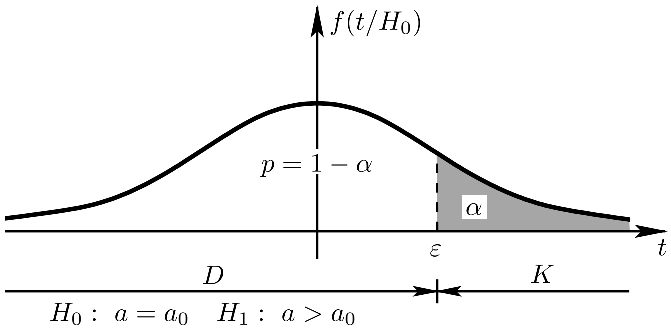
Анхны таамаглалыг хүлээн зөвшөөрөх утгуудын олонлогийг шинжүүрийн утгын муж (D), анхны таамаглалыг хэрэгсэхгүй байлгах утгуудын олонлогийг шинжүүрийн критик муж (K) гэж нэрлэнэ.

Итгэх түвшин α-ийн хувьд критик муж K-г сонгон авах олон янзын боломжууд байдаг. Итгэх түвшин α нь критик мужийн хэмжээг тодорхойлно. Шинжүүрийн утгын олонлогт критик мужийн байрлал нь өрсөлдөгч таамаглал H_1-ээс хамаарна. Тухайлбал, анхны таамаглал $H_0 : a = a_0$ бөгөөд өрсөлдөгч таамаглал нь $H_1 : a \neq a_0$ бол критик муж K-г $K : |t| > \varepsilon$ нөхцөл биелэгдэж байхаар сонгож авна. Энэ мужийг хоёр талт муж гэх ба Зураг 10.9-д дүрсэлсэн тэгш хэмт муж болно.

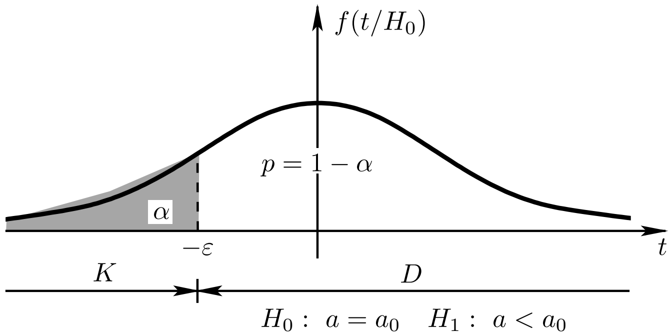


Зураг 10.9:

Критик мужийн дээр байгаа зураасласан муж тус бүрийн талбай $\alpha/2$-той тэнцүү байна. Хэрэв өрсөлдөгч таамаглал нь $H_1 : a > a_0$ $(H_1 :\ a < a_0)$ бол тэгш хэмт тархалтын хувьд критик муж нь $K :\ t > \varepsilon$ $(K :\ t < \varepsilon)$ нөхцөлөөр тодорхойлогдоно. (Зураг 10.10, Зураг 10.11)
Эдгээр мужуудыг нэг талт мужууд гэх ба харгалзан баруун өрөөсгөл, зүүн өрөөсгөл критик муж гэнэ. Зураасласан муж тус бүрийн талбай α-тай тэнцүү байна.

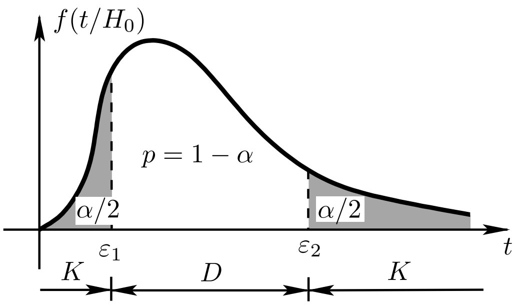


Зураг 10.10:



Зураг 10.11:

Тархалт нь тэгш хэмтэй биш боловч эерэг нягттай, шинжүүрийн өрсөлдөгч таамаглал нь $H_1 : a \neq a_0$ үед критик мужийг $K : 0 < t < \varepsilon_1, \ \varepsilon_2 < t$ нөхцөл биелэгдэж байхаар сонгон авна. Ийм мужийг квази тэгш хэмт критик муж гэнэ. (Зураг 10.12)



Зураг 10.12:

Эдгээр зураасласан талбайнууд мөн $\dfrac{\alpha}{2}$-той тэнцүү байна. Статистик шийдвэр гаргахдаа хэрэв, шинжүүрийн ажиглагдсан утгууд критик мужид харьяалагдаж байвал таамаглалыг хэрэгсэхгүй,
харин шинжүүрийн ажиглагдсан утгууд шинжүүрийн утгын мужид харьяалагдаж байвал таамаглалыг хүлээн зөвшөөрнө.

Одоо дээр тайлбарласныг математик тэмдэглэгээ ашиглан нэгтгэе. Санамсаргүй хэмжигдэхүүн X дээр туршилт хийж, n хэмжээст түүврийн утгууд $x_1, x_2, \ldots, x_n$ -ийг олсон гэе. Ажиглалтын дүнд H_0 таамаглал дэвшигдсэн бөгөөд дээрх түүврийн утгуудаас хамаарсан $T(x_1, x_2, \ldots, x_n, H_0)$ гэсэн шинжүүр сонгон авсан гэе. Итгэх түвшин α өгөгдсөн байг. Тэгвэл критик муж K нь $P_{H_0}(T \in K) = \alpha \quad (*)$ нөхцөлөөр тодорхойлогдоно. Хэрэв шинжүүрийн ажиглагдсан утга: $\widetilde{T} = T(x_1, x_2, \ldots, x_n, H_0)$ K мужид харьяалагдаж байвал

$(\widetilde{T} \in K)$ H_0 таамаглалыг хэрэгсэхгүй. $\widetilde{T} \notin K$ бол H_0 таамаглалыг хүлээн зөвшөөрөх ба энэ тохиолдолд H_0 таамаглалыг хазайхгүй гэнэ.
Өрсөлдөгч таамаглал үнэн үед шинжүүрийн утга критик мужид харьяалагдах магадлал $(P_{H_1}(T \in K))$-ыг шинжүүрийн чадал гэнэ. Өөрөөр хэлбэл, өрсөлдөгч таамаглал үнэн байхад тэг таамаглалыг хэрэгсэхгүй болгох магадлал юм. Таамаглал шалгах явцад итгэх түвшин ба түүврийн хэмжээ нь тогтмол, харин критик мужаа янз бүрээр сонгон авах боломж байгаа учраас шинжүүрийн чадал хамгийн их байхаар критик мужийг сонгон авна. Шинжүүрийн чадал хичнээн их байх тутам II төрлийн алдаа гарах магадлал төдий чинээн бага байна. Энэ утгаар критик мужийн оновчтой байрлал нь альтернатив таамаглал H_1-ийг яаж сонгон авснаас хамаардаг болохыг дахин тэмдэглэе.

Жишээ-10.11 Мэдэгдэж байгаа σ^2 дисперс бүхий хэвийн тархалттай санамсаргүй хэмжигдэхүүний математик дунджийн тухай: "Математик дундаж нь a_0 утгатай тэнцүү гэсэн" таамаглалыг p итгэх магадлалтай шалга.

$H_0 : a = a_0$ Өрсөлдөгч таамаглалыг $H_1 : a \neq a_0$ гэе.
$T = \dfrac{\overline{X} - a_0}{\sigma/\sqrt{n}}$ функцийг статистик шинжүүрээр сонгон авбал H_0 таамаглал үнэн үед T нь нормчлогдсон хэвийн тархалттай байна. Өгөгдсөн итгэх түвшин α-ийн хувьд критик муж нь $K : |t| > \varepsilon$ гэсэн тэгш хэмтэй муж болно. Мужийн хил ε-г $P(|t| < \varepsilon) = 2\Phi(\varepsilon) = 1 - \alpha$ нөхцөлөөс олно. $\Phi(t)$-но'рмчлогдсон хэвийн тархалтын функц.
Хэрэв $\widetilde{T} = \dfrac{\overline{X} - a_0}{\sigma/\sqrt{n}}$ утга K мужид харьяалагдаж байвал H_0 таамаглалыг хэрэгсэхгүй.
Тухайн тохиолдолд, $\sigma^2 = 16,\ n = 9, \overline{X} = 18$ бөгөөд $a_0 = 16$ үед H_0 таамаглалыг $\alpha = 0.05$ итгэх түвшинтэйгээр шалгая гэвэл: $2\Phi(\varepsilon) = 0.95$ нөхцөл ёсоор но'рмчлогдсон хэвийн тархалтын таблицаас $\varepsilon = 1.96$ гэж олдоно. Тэгвэл $K : |t| \geq 1.96$ буюу $\widetilde{T} = \dfrac{18 - 16}{4/3} = 1.5 < 1.96$ учир H_0 таамаглал хазайхгүй ба хүлээн зөвшөөрч болно гэсэн үг юм.
Хэвийн тархалттай эх олонлогийн параметрүүдийн тухай зарим таамаглалыг шалгах шинжүүр ба тэдгээрийн тархалт, критик мужийг олох нөхцөл зэргийг дараах хүснэгтэнд үзүүлэв.

N	шалгах таамаглал	урьдчилан өгөгдсөн нөхцөл	статистик шинжүүр	критик мужийн хилийг олох нөхцөл	H_0 таамаглал үнэн үед шинжүүрийн тархалт
1.	$X \sim N(a, \sigma^2)$ $H_0 : a = a_0$	σ^2- мэдэгдэх	$T = (\overline{X} - a_0)\frac{\sqrt{n}}{\sigma}$	$H_1 : a \neq a_0$ үед $2\Phi(\varepsilon) = 1 - \alpha$. $H_1 : a > a_0$ үед $2\Phi(\varepsilon) = 1 - 2\alpha$. $H_1 : a < a_0$ үед $2\Phi(-\varepsilon) = 1 - 2\alpha$.	$T \sim N(0, 1)$
2.	$X \sim N(a, \sigma^2)$ $H_0 : a = a_0$	σ^2-үл мэдэгдэх	$T = (\overline{X} - a_0)\frac{\sqrt{n}}{\widehat{S}}$	$H_1 : a \neq a_0$ үед $S_{n-1}(\varepsilon) = 1 - \alpha$. $H_1 : a > a_0$ үед $S_{n-1}(\varepsilon) = 1 - 2\alpha$. $H_1 : a < a_0$ үед $S_{n-1}(-\varepsilon) = 1 - 2\alpha$.	$T \sim S_{n-1}(t)$
3.	$X \sim N(a, \sigma^2)$ $H_0 : \sigma^2 = \sigma_0^2$	$a-$ мэдэгдэх	$T = nS_*^2/\sigma_0^2$ $S_*^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - a)^2$	$H_1 : \sigma^2 \neq \sigma_0^2$ үед $\chi^2_{1-\alpha/2;\ n} < \chi^2 <$ $< \chi^2_{\alpha/2;\ n}$, (H_0+) $H_1 : \sigma^2 > \sigma_0^2$ үед $\chi^2 \geq \chi^2_{\alpha;\ n}$, (H_0-) $H_1 : \sigma^2 < \sigma_0^2$ үед $\chi^2 \leq \chi^2_{1-\alpha;\ n}$, (H_0-)	$T \sim \chi^2_n(t)$
4.	$X \sim N(a, \sigma^2)$ $H_0 : \sigma^2 = \sigma_0^2$	$a-$үл мэдэгдэх	$T = (n-1)\widehat{S}^2/\sigma_0^2$	$H_1 : \sigma^2 \neq \sigma_0^2$ үед $\chi^2_{1-\alpha/2;\ n-1} < \chi^2 <$ $< \chi^2_{\alpha/2;\ n-1}$, (H_0+) $H_1 : \sigma^2 > \sigma_0^2$ үед $\chi^2 \geq \chi^2_{\alpha;\ n-1}$, (H_0-) $H_1 : \sigma^2 < \sigma_0^2$ үед $\chi^2 \leq \chi^2_{1-\alpha;\ n-1}$, (H_0-)	$T \sim \chi^2_{n-1}(t)$
5.	$X \sim N(a_1, \sigma_x^2)$ $Y \sim N(a_2, \sigma_y^2)$ $H_0 : a_1 = a_2$	$\sigma_x^2,\ \sigma_y^2-$ мэдэгдэх $\sigma_x^2 = \widehat{S}_x^2$ $\sigma_y^2 = \widehat{S}_y^2$	$T = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{\sigma_x^2}{n_1} + \frac{\sigma_y^2}{n_2}}}$ n_1, n_2-түүврийн хэмжээ	$H_1 : a_1 \neq a_2$ үед $2\Phi(\varepsilon) = 1 - \alpha$. $H_1 : a_1 > a_2$ үед $2\Phi(\varepsilon) = 1 - 2\alpha$. $H_1 : a_1 < a_2$ үед $2\Phi(-\varepsilon) = 1 - 2\alpha$.	$T \sim N(0, 1)$

6.	$X \sim N(a_1, \sigma_x^2)$ $Y \sim N(a_2, \sigma_y^2)$ $H_0 : a_1 = a_2$	$\sigma_x^2,\ \sigma_y^2$−үл мэдэгдэх $\sigma_x^2 = \widehat{S}_x^2$ $\sigma_y^2 = \widehat{S}_y^2$	$T = (\overline{X} - \overline{Y}) / \sqrt{\frac{n_1\sigma_x^2 + n_2\sigma_y^2}{n_1+n_2-2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$ n_1, n_2−түүврийн хэмжээ	$H_1 : a_1 \neq a_2$ үед $S_k(\varepsilon) = 1 - \alpha.$ $H_1 : a_1 > a_2$ үед $S_k(\varepsilon) = 1 - 2\alpha.$ $H_1 : a_1 < a_2$ үед $S_k(-\varepsilon) = 1 - 2\alpha.$	$T \sim S_k(t)$ $k = n_1 + n_2 - 2$
7.	$X \sim N(a_1, \sigma_x^2)$ $Y \sim N(a_2, \sigma_y^2)$ $H_0 : \sigma_x^2 = \sigma_y^2$	$a_1,\ a_2$− мэдэгдэх	$T = \frac{S_{*1}^2}{S_{*2}^2}$ $S_{*1}^2 = \frac{1}{n_1}\sum_{i=1}^{n_1}(x_i - a_1)^2$ $S_{*2}^2 = \frac{1}{n_2}\sum_{i=1}^{n_2}(x_i - a_2)^2$	$H_1 : \sigma_x^2 \neq \sigma_y^2$ үед $F_{1-\alpha/2}(n_1; n_2) < F <$ $< F_{\alpha/2}(n_1; n_2)$, (H_0+) $H_1 : \sigma_x^2 > \sigma_y^2$ үед $F > F_\alpha(n_1; n_2)$, (H_0-) $H_1 : \sigma_x^2 < \sigma_y^2$ үед $F_{1-\alpha}(n_1; n_2) = 1/F_\alpha(n_2; n_1)$ $F < F_{1-\alpha}(n_1; n_2)$, (H_0-)	$T \sim F(n_1; n_2)$
8.	$X \sim N(a_1, \sigma_x^2)$ $Y \sim N(a_2, \sigma_y^2)$ $H_0 : \sigma_x^2 = \sigma_y^2$	$a_1,\ a_2$−үл мэдэгдэх $a_1 = \overline{X}$ $a_2 = \overline{Y}$	$T = \frac{\widehat{S}_x^2}{\widehat{S}_y^2}$	$H_1 : \sigma_x^2 \neq \sigma_y^2$ үед $F_{1-\alpha/2}(k_1; k_2) < F <$ $< F_{\alpha/2}(k_1; k_2)$, (H_0+) $H_1 : \sigma_x^2 > \sigma_y^2$ үед $F > F_\alpha(k_1; k_2)$, (H_0-) $H_1 : \sigma_x^2 < \sigma_y^2$ үед $F_{1-\alpha}(k_1; k_2) = 1/F_\alpha(k_2; k_1)$ $F < F_{1-\alpha}(k_1; k_2)$, (H_0-)	$T \sim F(k_1; k_2)$ $k_1 = n_1 - 1$ $k_2 = n_2 - 1.$

Үүнд, (H_0+) нь H_0 таамаглалыг хүлээн зөвшөөрнө, (H_0-) нь H_0 таамаглалыг няцаана гэсэн товчилсон тэмдэглэгээ, $F(n_1, n_2)$ нь n_1 ба n_2 чөлөөний зэргүүд бүхий Фишерийн тархалт (F-тархалт). Фишерийн тархалтын хүснэгтийг хавсралтаас үз (Таблиц N8).

Статистик таамаглал шалгах явцад ямар нэг H_0 таамаглалыг хүлээн зөвшөөрч байгаа нь түүнийг яг баталсан гэсэн үг биш, харин түүнийг үгүйсгэх бүрэн үндэс байхгүй гэсэн үг юм. Өөрөөр хэлбэл, H_0 таамаглал судалгааны үр дүнд харшлахгүй гэдгийг л нотоллоо гэсэн үг. Үүнчлэн H_0 таамаглалыг хэрэгсэхгүй болгож байгаа нь түүнийг хүлээн зөвшөөрөх бүрэн үндэслэл байхгүй гэсэн үг юм.

Энэ бүгдээс үндэслэн энгийн таамаглалыг хүлээн зөвшөөрсний дараа нэмэлт судалгаа шаардлагатай болохыг тэмдэглэе.

Санамж: Тархалтын параметрүүдийн тухай дээрх шинжүүрүүдэд, $H_0 : a = a_0$ таамаглалыг α-итгэх түвшинтэйгээр хүлээн зөвшөөрөх муж нь a параметрийн $1 - \alpha$ итгэх магадлал бүхий итгэх завсартай давхцана. Өрөөсгөл муж бүхий итгэх шинжүүрт өрөөсгөл итгэх завсар, хоёр талт муж бүхий итгэх шинжүүрт хоёр талт итгэх завсар тус тус харгалзана. Энэ бүхэн нь дэвшүүлсэн H_0 таамаглалыг a параметрийн итгэх интервалаар шалгах боломжийг олгож буй юм.

Хэрэв a параметрийн a_0 утга харгалзах итгэх интервалаар хучигдаж байвал H_0 таамаглалыг хүлээн зөвшөөрнө. Эсрэг тохиолдолд H_0 таамаглал хазайна. $H_0 : a_1 = a_2$ гэсэн таамаглал шалгаж байгаа бол $a_1 - a_2$ ялгаврын итгэх интервалыг байгуулна. Байгуулсан интервал $a_1 - a_2$ ялгаврын тэг утгыг хучиж байвал H_0 таамаглалыг хүлээн зөвшөөрнө. Харин дисперсүүдийн тэнцлийн тухай $H_0 : \sigma_x^2 = \sigma_y^2$ таамаглалын хувьд итгэх интервал нь дисперсүүдийн харьцааны хувьд байгуулагдах учраас байгуулсан итгэх интервал параметрүүдийн нэгтэй тэнцүү утгыг хучиж байвал H_0 таамаглалыг хүлээн зөвшөөрнө.

10.7.2 Тархалтын хуулийн тухай шинжүүр

Судлаж буй санамсаргүй хэмжигдэхүүний тархалтын хууль мэдэгдэхгүй тохиолдолд статистик цувааны тусламжтайгаар гистограмм байгуулж, санамсаргүй хэмжигдэхүүн онолын тархалтын ямар хуультай байж болох тухай таамаглал дэвшүүлнэ. Өөрөөр хэлбэл, онолын тархалтын функц нь статистик шинжүүр шаардах статистик таамаглал байдаг. Энэ таамаглалыг шалгахдаа онолын тархалтын нягтын функц ба гистограммыг харьцуулж үзэх графикийн аргыг хэрэглэж болно. Жишээлбэл, хэмжилтийн үр дүнгээр зохиосон интервалын вариацын цуваа авч үзье.

$x:$	$[-4;-3[$	$[-3;-2[$	$[-2;-1[$	$[-1;0[$	$[0;1[$	$[1;2[$	$[2;3[$	$[3;4]$
$p^*:$	0.011	0.051	0.140	0.270	0.244	0.172	0.090	0.022

Гистограммыг байгуулж, графиктай харьцуулбал: "санамсаргүй хэмжигдэхүүн хэвийн тархалттай" гэсэн таамаглалыг дэвшүүлж болохоор байна. Одоо тархалтын параметрүүдийг олъё. Үүний тулд түүврийн элементүүдийг хэрчмийн дунджуудаар сонгон авч түүврийн дундаж ба дисперсийг бодъё:

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}x_i=\frac{1}{n}\sum_{i=1}^{k}n_ix_i=\sum_{i=1}^{8}p_i^*x_i\approx 0.171,\quad S^2=\sum_{i=1}^{8}x_i^2p_i^*-\overline{X}^2\approx 2.18675$$

Хамгийн их үнэний хувь бүхий арга ёсоор хэвийн тархалтын параметрүүд $a=\overline{X},\quad \sigma^2=S^2$ байдаг учир $S=\sqrt{2.18675}\approx 1.479$. Тэгвэл онолын тархалтын нягтын функц

$$F(x)=\frac{1}{1.479\sqrt{2\pi}}e^{-\frac{(x-0.171)^2}{2\cdot 2.18675^2}}$$

болно.

$y=f(x)$ функцийн утгуудыг хэрчмийн захын цэгүүд дээр олж таблиц зохиовол

$x:$	-4	-3	-2	-1	0	1	2	3	4
$f(x):$	0.004	0.025	0.090	0.199	0.274	0.234	0.124	0.042	0.008

Энэ таблицаас $f(x)$ функцын графикийг байгуулж өмнөх байгуулсан гистограммтай харьцуулах замаар дүгнэлт гаргаж болно.

Одоо хэвийн тархалтын хуулийг шалгах хэд хэдэн практик шинжүүрүүдийг авч үзье.

1. Дундаж абсолют хазайлт ашиглах шинжүүр

Түүврийн хэмжээ их биш ($n<120$) үед, "хэрэв түүврийн тархалт нь хэвийн тархалттай ойролцоо тархалттай бол

$$\left|\frac{\overline{\overline{X}}}{\overline{\sigma}}-0.7979\right|<\frac{0.4}{\sqrt{n}} \qquad (**)$$

нөхцөл биелэгдэнэ" гэсэн энгийн зөвлөмжөөр шалгаж болно.
Үүнд, $\overline{\overline{X}} = \frac{1}{n}\sum\limits_{i=1}^{n}|x_i - \overline{X}|$ дундаж абсолют хазайлт,
$\overline{\sigma} = S$ (зөвхөн хэвийн тархалттай нөхцөлд энэ тэмдэглэгээг хэрэглэнэ).
Жишээлбэл, нэг ангийн оюутнуудын өндрийн хэмжээгээр зохиогдсон статистик цуваа авч үзье.

N	$[X_{i-1}; x_i]$	n_i	p_i^* (давтамж)	$\overline{x}_i$ (инт.дундаж)
1	[165;169[	5	0.089	167
2	[169;173[	13	0.232	171
3	[173;177[	15	0.268	175
4	[177;181[	14	0.250	179
5	[181;185[	5	0.089	183
6	[185;189]	4	0.072	187

Түүврийн параметрүүдийг олъё.

$$\overline{X} = \sum_{i=1}^{6} p_i^* \overline{x}_i \approx 175.93, \quad S = \sqrt{\sum_{i=1}^{6} p_i^* \overline{x}_1^2 - \overline{X}^2} \approx 5.55,$$

$$\overline{\overline{X}} = \frac{1}{56}\sum_{i=1}^{56}|x_i - \overline{X}| \approx \frac{247}{56} = 4.41$$

Эдгээр холбогдлуудаар (**) нөхцөлийг шалгавал $\left|\frac{4.41}{5.55} - 0.7979\right| < \frac{0.4}{\sqrt{56}}$ буюу $0.0032 < 0.0535$ болно.
Иймд түүврийн тархалт нь хэвийн тархалттай гэсэн таамаглалыг хүлээн зөвшөөрч болно.

2. Интервалын далайц ашиглах шинжүүр

Түүврийн хэмжээ $3 < n < 1000$ үед хэвийн тархалтын хуулийг интервалын далайц ашиглан шалгаж болно. Интервалын далайц $R = x_{max} - x_{min}$-ийг олсны дараа $R/\overline{\sigma}$ харьцааг бодож хавсралтын таблицаас (табл.N6) энэ харьцааны харгалзах дээд, доод утгуудыг олж жишнэ. Хэрэв $R/\overline{\sigma}$ харьцаа алдааны 0.1 магадлалтайгаар дээд ба доод хязгаарын хооронд оршиж байвал хэвийн тархалтын тухай таамаглалыг энэ шинжээр хүлээн зөвшөөрч болно. Тухайлбал, өмнөх жишээнд $R = 189 - 165 = 24$. $R/\overline{\sigma} = \frac{24}{5.55} \approx 4.324$, $n = 56$, $p = 0.1$ үед таблицаас харгалзах дээд ба доод утгуудыг олбол 5.23 ба 4.03 болно. Манай тохиолдолд $4.03 < 4.324 < 5.23$ учраас түүвэр хэвийн тархалттай гэж үзэж болно.

3. Түүврийн эксцесс ба асимметрийн коэффициент ашиглах шинжүүр

Түүврийн хэмжээ бага ($n < 20$) үед хэвийн тархалтын хуулийг эксцесс ба асимметрийн коэффициентуудын тусламжтайгаар шалгаж болно. Түүврийн эдгээр коэффициентуудыг дараах томъёогоор олдог болохыг бид мэдэх билээ.

$$\overline{A} = \frac{1}{nS^3}\sum_{i=1}^{n}(x_i - \overline{X})^3, \quad \overline{E} = \frac{1}{nS^4}\sum_{i=1}^{n}(x_i - \overline{X})^4 - 3$$

Үүнд S^2 түүврийн дисперс. Эдгээр санамсаргүй хэмжигдэхүүн тус бүрийн дисперсүүд нь харгалзан

$$D(\overline{A}) = \frac{6(n-2)}{(n+1)(n+3)} \tag{10.53}$$

$$D(\overline{E}) = \frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)} \tag{10.54}$$

томъёогоор бодогдоно. Хэрэв

$$|\overline{A}| \le 3\sqrt{D(\overline{A})} \tag{10.55}$$

$$|\overline{E}| \le 5\sqrt{D(\overline{E})} \tag{10.56}$$

нөхцөлүүд биелэгдэж байвал уг түүврийг хэвийн тархалттай гэж үзэж болно. Өмнө авч үзсэн жишээний хувьд $\overline{A} = 0.262$, $\overline{E} = -0.861$, $D(\overline{A}) = 0.098$, $D(\overline{E}) = 0.329$ ба $|\overline{E}| \le 2.867$, $|\overline{A}| \le 0.939$ учир өгөгдсөн түүвэр хэвийн тархалттай байна.

4. Пирсоны хи-квадрат (χ^2) шинжүүр

Санамсаргүй хэмжигдэхүүний тархалтын хуулийн тухай таамаглалыг шалгахад практикт өргөн хэрэглэгддэг, тохиромжтой бөгөөд ерөнхий арга нь К.Пирсоны хи-квадрат шинжүүр юм. Энэ шинжүүр, туршилтаар тогтоосон тархалтын хууль нь урьдчилан өгөгдсөн юмуу эсвэл урьдчилан төсөөлж буй тархалтын хуультай тохирох эсэх тухай таамаглалыг шалгана.

Эхлээд n хэмжээст түүврээр зохиогдсон санамсаргүй хэжигдэхүүний өөрчлөгдөх мужийг k интервалд хуваана. Интервалын тоог $8 < k < 20$ байхаар авбал тохиромжтой бөгөөд практикт ихэвчлэн $k = [1 + 3.2\ln n]$ томъёогоор олдог. Үүнд $[x]$ нь x тооны бүхэл хэсэг. i-р интервалд харьяалагдах түүврийн элементийн тоог n_i гэе. Үүний дараа түүврийн тусламжтай гистограммыг байгуулан, санамсаргүй хэмжигдэхүүний тархалтын хуулийн хэлбэрийн тухай таамаглал дэвшүүлж, тархалтын үл мэдэгдэх параметрүүдийг олно. Параметрүүдийг олохдоо хамгийн их үнэний хувь бүхий аргыг хэрэглэж болно.

Сонгон авсан тархалтын хуулийн тусламжтайгаар санамсаргүй хэмжигдэхүүн i-р завсраас утгаа авах магадлал p_i-г олно. Түүврээр зохиогдсон тархалт ба сонгон авсан тархалтын хоорондох хазайлтыг тодорхойлохын тулд χ^2 гэж нэрлэгдэх дараах шинжүүрийг авч үздэг.

$$\chi^2 = \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i} = \sum_{i=1}^{k} \frac{n}{p_i}(p_i^* - p_i)^2 \qquad (10.57)$$

Математик статистикт дээрх хэмжигдэхүүнийг $l = k - r - 1$ чөлөөний зэрэг бүхий χ^2-тархалттай гэж баталдаг. Үүнд, k- интервалын тоо, n-түүврийн хэмжээ, r-сонгон авсан тархалтын хуулийн параметрийн тоо.
Өгөгдсөн итгэх түвшин α ба χ^2 тархалтын таблицаас дээрх хэмжигдэхүүний онолын утга $\chi^2_{l,\alpha}$-г олно.
Хэрэв $\chi^2 \leq \chi^2_{1,\alpha}$ нөхцөл биелэгдэж байвал сонгон авсан тархалтын хуулийг хүлээн зөвшөөрнө. Эсрэг тохиолдол нь уг түүвэр энэ тархалтын хуульд нийцэхгүйг харуулна.

χ^2- шинжүүрийг түүврийн хэмжээ $n \geq 50 \div 150$ ба $n_i \geq 5 \div 8$ үед хэрэглэвэл тохиромжтой. Хэрэв аль нэг i-р интервалд харьяалагдах элементүүдийн тоо $n_i < 5$ байвал хөрш зэргэлдээ орших хоёр буюу хэд хэдэн интервалуудыг нэгтгэнэ. Эх олонлог хэвийн тархалттай тохиолдолд $r = 2$ бөгөөд $\overline{X}$, $\overline{\sigma}^2$ параметрүүд нь түүврээр тодорхойлогдохын гадна p_i-г олохдоо дараах томъёог ашиглана.

$$p_i = P(x_i \leq X \leq x_{i+1}) = \Phi\left(\frac{x_{i+1} - \overline{X}}{\overline{\sigma}}\right) - \Phi\left(\frac{x_i - \overline{X}}{\overline{\sigma}}\right)$$

Жишээ-10.12 Хайлуулж буй гангийн чанарыг удирдахын тулд түүний механик чанарын тархалтын хуулийг мэдэх шаардлагатай болдог байна. 35ГС маркийн ганг судалж $n = 374$ удаа ажиглалт хийсний дүнд гангийн механик чанаруудын нэг болох "урсах хязгаар хэвийн тархалттай" гэсэн таамаглал дэвшүүлжээ. Туршилтаар $\overline{X} = 42.37$кг/мм^2, $\overline{\sigma} = 0.94$кг/мм^2 гэж тогтоосон бөгөөд интервалын вариацын цувааг дараах хүснэгтэнд өгөв.

N	$[x_i; x_{i+1}]$	n_i
1	$[40; 41[$	20
2	$[41; 42[$	112
3	$[42; 43[$	154
4	$[43; 44[$	73
5	$[44; 45[$	13
6	$[45; 46]$	2
$\sum$	-	374

Тархалтын хуулийн тухай энэхүү таамаглалыг $\alpha = 0.01$ итгэх түвшинтэйгээр шалгая. Интервалын вариацын цуваанаас харвал 6-р интервалын ажиглалтын утга 5-аас бага учраас 5, 6 дугаар интервалуудыг нэгтгэе.

$[x_i; x_{i+1}]$	$[40; 41[$	$[41; 42[$	$[42; 143[$	$[43; 44[$	$[44; 46]$
n_i	20	112	154	73	15

Өгөгдсөн санамсаргүй хэмжигдэхүүний утга хуваалтын интервал тус бүрд харьяалагдах магадлал p_i-г олъё.

$p_1 = P(40 < X < 41) = \Phi\Big(\frac{41-42.37}{0.94}\Big) - \Phi\Big(\frac{40-42.37}{0.94}\Big) = 0.4940 - 0.4280 = 0,066.$
Үүнчлэн $p_2 = 0.2762, p_3 = 0.3971,\ p_4 = 0.2128,\ p_5 = 0,0417.$
χ^2 шинжүүрийн утгыг олохын тулд дараах таблицыг зохиох нь тохиромжтой байдаг.

N	$[x_i \div x_{i+1}]$	n_i	p_i	np_i	$\|n_i - np_i\|$	$(n_i - np_i)^2$	$\frac{(n_i - np_i)^2}{np_i}$
1	[40; 41[	20	0.066	24.76	4.76	21.56	0.871
2	[41; 42[	112	0.2762	103.70	8.30	68.89	0.664
3	[42; 43[	154	0.3971	148.52	5.48	29.03	0.127
4	[43; 44[	73	0.2128	79.59	6.59	43.43	0.546
5	[44; 46]	15	0.0417	15.60	0.60	0.36	0.023
$\sum$	-	374	0.9938	372.17	-	-	$\chi^2 = 2.231$

Манай тохиолдолд интервалын тоо $k = 5$, эх олонлог хэвийн тархалттай гэж үзэж буй тул $r = 2$. Иймд чөлөөний зэрэг $l = 5 - 2 - 1 = 2$. χ^2-тархалтын таблицаас $\alpha = 0.01$, $l = 2$ байх үеийн утгыг олбол $\chi^2_{2;0.01} = 9.2$, $2.231 < 9.2$ учир 35ГС -маркийн гангийн "урсах хязгаар хэвийн тархалттай" гэсэн таамаглалыг хүлээн зөвшөөрч болно.

5. Колмогоровын шинжүүр

n удаагийн үл хамаарах туршилтын үр дүнгээр, тасралтгүй санамсаргүй хэмжигдэхүүний туршилтын тархалтын функц $F_n(x)$-ийг байгуулсан гэе. Тэгвэл туршилтын тоо хүрэлцээтэй их үед $F_n(x)$ функц эх олонлогийн тархалтын функц $F(x)$ уруу нийлдэг болохыг бид мэдэх билээ. Оросын математикч А.Н.Колмогоров $n \to \infty$ үед $D_n = max|F_n(x) - F(x)|\sqrt{n}$ хэмжигдэхүүн $F(x)$ функцийн хэлбэрээс хамаарахгүйгээр

$$K(\lambda) = \sum_{k=-\infty}^{+\infty} (-1)^k e^{-2k^2\lambda^2} \tag{10.58}$$

хязгаартай байна гэж баталжээ. (10.58) функцийг Колмогоровын функц гэнэ. Санамсаргүй хэмжигдэхүүний тархалтын функц ба Колмогоровын функцийн тодорхойлолтоос, хүрэлцээтэй их n ба дурын $\lambda > 0$ тооны хувьд

$$P(D_n < \lambda) \approx K(\lambda) = \sum_{k=-\infty}^{+\infty} (-1)^k e^{-2k^2\lambda^2} \tag{10.59}$$

$$P(D_n \geq \lambda) \approx 1 - K(\lambda) = 1 - \sum_{k=-\infty}^{+\infty} (-1)^k e^{-2k^2\lambda^2} \tag{10.60}$$

нөхцөлүүд мөрднө.
(10.60) функцийн таблицыг λ-ийн утгуудад зохиосон байдаг. (табл.N5)
Одоо онолын ба туршилтын тархалтын функцүүдын хоорондын хазайлтыг Колмогоровын шинжүүрээр шалгах дараалыг авч үзье. Эхлээд туршилтын

тархалтын функцийн утга онолын функцээс хазайх хамгийн их абсолют хазайлтыг дараах томъёогоор олно.

$$D = max|F_n(x) - F(x)| \tag{10.61}$$

Дараа нь

$$\lambda = \sqrt{n} \cdot D \tag{10.62}$$

нөхцөлөөс λ-хэмжигдэхүүнийг олно. Өгөгдсөн итгэх түвшин α-д харгалзах λ-ийн онолын утгыг дараах таблицаас олно.

α	0.01	0.02	0.05	0.1	0.15	0.20	0.25	0.3
$\lambda_{1-\alpha}$	1.63	1.52	1.36	1.22	1.14	1.07	1.02	0.97

α	0.4	0.5	0.6	0.7	0.8	0.9	0.99
$\lambda_{1-\alpha}$	0.89	0.83	0.77	0.71	0.64	0.57	0.44

Хэрэв $\lambda < \lambda_{1-\alpha}$ нөхцөл биелэгдэж байвал санамсаргүй хэмжигдэхүүний онолын тархалтын функц $F(x)$ хуультай байна. Колмогоровын шинжүүрийн давуу тал нь $K(\lambda)$ магадлал урьдчилан сонгон авсан тархалтын хуулиас хамаардаггүйд оршино.

Жишээ-10.13 Санамсаргүй хэмжигдэхүүний туршилтын үр дүнгээр зохиосон статистик цуваа авч үзье.

N	$[x_i; x_{i+1}]$	$\overline{x}_i = \dfrac{x_i + x_{i+1}}{2}$	n_i	p_i^*
1	$[-20; -15[$	-17.5	7	0.035
2	$[-15; -10[$	-12.5	11	0.055
3	$[-10; -5[$	-7.5	15	0.075
4	$[-5; 0[$	-2.5	24	0.120
5	$[0; 5[$	2.5	49	0.245
6	$[5; 10[$	7.5	41	0.205
7	$[10; 15[$	12.5	26	0.130
8	$[15; 20[$	17.5	17	0.085
9	$[20; 25[$	22.5	7	0.035
10	$[25; 30]$	27.5	3	0.015

Түүврийн параметрүүдийг олбол

$$\overline{X} = \sum_{i=1}^{10} p_i^* \overline{x}_i = 4.30, \qquad \overline{\sigma} = \sqrt{\sum_{i=1}^{10} p_i^* (\overline{x}_i - \overline{X})^2} = 9.71$$

Санамсаргүй хэмжигдэхүүн хэвийн тархалттай байх тухай таамаглалыг χ^2 ба Колмогоровын шинжүүрээр давхар шалгая. Үүний тулд дараах таблицыг ашиглавал тохиромжтой.

N	$[x_i; x_{i+1}]$	n_i	np_i	$nF_n(x)$	$nF(x)$	$n\|F_n(x)-F(x)\|$	$\frac{(n_i-np_i)^2}{np_i}$
1	$]-\infty;-15[$	7	4.7	7	4.7	2.3	1.12
2	$[-15;-10[$	11	9.5	18	14.2	3.8	0.23
3	$[-10;-5[$	15	19.5	33	33.7	0.7	1.18
4	$[-5;0[$	24	31.6	57	65.3	8.3	1.81
5	$[0;5[$	49	40.3	106	105.6	0.4	1.87
6	$]5;10[$	41	38.9	147	144.5	2.5	0.11
7	$[10;15[$	26	29.2	173	173.7	0.7	0.20
8	$[15;20[$	17	15.6	190	189.3	0.7	0.002
9	$[20;25[$	7	7.7	197	197	0.0	0.06
10	$[25;+\infty[$	3	3	200	200	0.0	0.00
	-	200	-	-	-	$nD = 8.3$	$\chi^2 = 6.642$

Энд p_i магадлалуудыг χ^2 шинжүүрт авч үзсэн томъёогоор олсон болно. Тухайлбал,

$$p_2 = \Phi\left(\frac{-10-4.3}{9.71}\right) - \Phi\left(\frac{-15-4.3}{9.71}\right) = 0.0475,\ np_2 = 200 \cdot 0.0475 = 9.5\ ,$$

$\chi^2 = \sum\limits_{i=1}^{10} \frac{(n_i - np_i)^2}{np_i} = 6,642$. Чөлөөний зэрэг $l = 10 - 2 - 1 = 7$. $\chi^2_{7,0.05} = 14,1$. Эцэст нь $6.642 < 14.1$ тул авч үзэж буй санамсаргүй хэмжигдэхүүнийг хэвийн тархалттай гэж үзэж болно.

Одоо дээрх таамаглалыг Колмогоровын шинжүүрээр шалгая. $nD = max\ n|F_n(x) - F(x)| = 8.3$, $\lambda = \frac{nD}{\sqrt{n}} = 0.59$, $\alpha = 0.05$ үед өмнөх хүснэгтээс $\lambda_{0.95} = 1.36$ гэж олдоно. $0.59 < 1.36$ учир өгөгдсөн санамсаргүй хэмжигдэхүүн мөн л хэвийн тархалттай гэдэг нь дахин батлагдаж байна.

Хэвийн тархалтын тухай дээр авч үзсэн шинжүүрүүдээс, эксцесс ба асимметрийн коэффициентийг ашиглах шинжүүр нь моментуудын хэрэглэгээг үзүүлэхээс гадна компьютер дээр тооцоог хийхэд ихээхэн тохиромжтой. Харин Колмогоровын шинжүүрийг ховор хэрэглэнэ. Практикт нилээд хэрэглэгддэг тохиромжтой шинжүүр нь интервалын далайц ашиглах шинжүүр ба χ^2 шинжүүрүүд юм.

Бүлэг 11

Дисперсийн шинжилгээ

11.1 Дисперсийн шинжилгээний үндсэн санаа

Нэгэн зэрэг үйлчлэх янз бүрийн хүчин зүйлээс хамаарсан ажиглалтуудын үр дүнг боловсруулах, гол нөлөө бүхий хүчин зүйлүүдийг сонгох, тэдгээрийн нөлөөллийг үнэлэх, өөр хооронд нь жиших аргыг дисперсийн шинжлэлд авч үздэг. Туршилтанд нөлөөлж буй гадаад нөхцлүүдийг фактор буюу хүчин зүйл гэж нэрлэнэ. Тухайлбал, t^0, атмосферийн даралт, таталцлын хүч, тоног төхөөрөмжийн төрөл гэх мэт ерөнхий хүчин зүйлүүдээс гадна тухайн туршилтанд нөлөөлөх ямарч параметрийг үндсэн хүчин зүйл болгон сонгон авч болно. Гэхдээ нөлөө нь мэдэгдэхүйц бөгөөд хянаж болох тийм хүчин зүйлүүдийг сонирхоно. Туршилтанд хүчин зүйлүүд нь хувьсаж болно. Энэ тохиолдолд хүчин зүйл нь янз бүрийн түвшинд хувьсаж байна гэх буюу эсвэл хүчин зүйл хэд хэдэн түвшинтэй байна гэж ярьдаг.

Ямар ч туршилтын хувьд, судалж байгаа хэмжигдэхүүний дундаж утга нь үндсэн хүчин зүйлийн тоон ба чанарын өөрчлөлтөөс хамаарахаас гадна санамсаргүй хүчин зүйлүүдээс хамаарна. Инхүү үндсэн болон санамсаргүй хүчин зүйлүүд ажиглалтын дундаж утганд хэрхэн нөлөөлөхийг судлах нь дисперсийн шинжилгээний үндсэн бодлого юм.

Шинжилж буй хүчин зүйлийн тооноос хамаарч нэг хүчин зүйлт, хоёр хүчин зүйлт, олон хүчин зүйлт шинжилгээ гэж ангилж үзнэ. Хэд хэдэн хүчин зүйлийн хувьд дисперсийн шинжилгээ хийх нь тохиромжтой байдаг. Сонгодог аргад зөвхөн ганц хүчин зүйлийг өөрчлөгддөг мэтээр авч үзэж, бусдыг тогтмол гэж тооцдог. Судалгааны ийм аргад, хүчин зүйлүүд нэгэн зэрэг өөрчлөгдөх үед тэдгээрийн харилцан үйлчлэлийг тооцож болдоггүй учраас өрөөсгөл юм. Ийм учраас дисперсийн шинжилгээнд түүврийн нийт дисперсийг, хоорондоо үл хамаарах хүчин зүйлүүдтэй уялдсан нэмэгдэхүүнд задалж, эдгээр нэмэгдэхүүн бүрийн нөлөөллийг хооронд нь харьцуулах замаар гол нөлөө бүхий хүчин зүйлүүдийг сонгон авдаг. Жишээлбэл, ямар нэг хэмжигдэхүүн M-ийн утгыг хэмжиж X гэсэн утга гарган авсан ба хэмжилтийн процесст A, B гэсэн үл хамаарах санамсаргүй хүчин зүйлүүд нөлөөлдөг гэе. Тэгвэл хазайлт нь $M - X = \alpha + \beta + \gamma$ болох ба дисперсийг бодвол

$D(M - X) = D(\alpha + \beta + \gamma) = D\alpha + D\beta + D\gamma$. Үүнд:

$D\alpha - A$ хүчин зүйлийн нөлөө,

$D\beta - B$ хүчин зүйлийн нөлөө,

$D\gamma -$тооцогдоогүй үлдсэн бусад санамсаргүй хүчин зүйлүүдийн нөлөө.

$D\gamma$-г үлдэгдэл дисперс гэнэ.

Эдгээр дисперс нь эх олонлогийн дисперсийн үнэлэлтүүд болох ба A, B хүчин зүйлүүдийн нөлөөг үнэлэхийн тулд $D\alpha, D\beta$ дисперсүүдийг $D\gamma$ дисперстэй харьцуулж үздэг.

Дисперсийн шинжилгээ хийхийн тулд:

1. Ажиглалт буюу хэмжилтийн үр дүнгүүд нь хэвийн тархалттай, хамааралгүй санамсаргүй хэмжигдэхүүн байх
2. Хүчин зүйлүүд зөвхөн дундаж утгуудын өөрчлөлтөнд нөлөөлөх ба ажиглалтуудын дисперс ижил байх

гэсэн үндсэн шаардлагуудыг хангасан байх ёстой. Зөвхөн ийм нөхцөлд гарган авсан математик дундаж ба дисперсийн нөлөөг үнэлж, итгэх завсруудыг байгуулж болно.

11.2 Нэг хүчин зүйлт дисперсийн шинжилгээ.

Ямар нэг автомат шугамын суурь машинууд дээр ижил төрлийн деталууд үйлдвэрлэдэг гэе. Цаашдын боловсруулалтыг зөв төлөвлөхийн тулд суурь машин бүр дундаж хэмжээ нь хоорондоо ижил байх деталиуд гаргаж чадаж байна уу?, үгүй юу? гэдгийг мэдэх шаардлагатай. Энэ тохиолдолд деталын дундаж хэмжээнд нөлөөлөх үндсэн хүчин зүйл нь тэдгээрийг үйлдвэрлэж буй суурь машинууд юм.

Тэгвэл энэхүү хүчин зүйл деталын дундаж хэмжээнд хир их нөлөөлж байна вэ? Энэ хүчин зүйлийн нөлөөг арилгаж болох уу? гэсэн асуудлыг сонирхож болно. Суурь машин бүр дээр үйлдвэрлэсэн деталын хэмжээ хэвийн тархалттай бөгөөд тэнцүү дисперстэй гэе.

m тооны суурь машинтай гэвэл бидний сонирхсон хүчин зүйл m янзын утга авч болох учраас m-түвшинтэй гэсэн үг. i-р түвшин бүр дээр n_i $(i = \overline{1, m})$ ажиглалт хийвэл нийт ажиглалтын тоо $N = n_1 + n_2 + \cdots + n_m$ болох ба хялбарыг бодож түвшин бүр дээр ижилхэн n удаа ажиглалт хийсэн гэе. Тэгвэл түүврийн нийт $N = m \cdot n$ элементийг ажиглалтын матриц хэмээх дараах матрицаар дүрсэлж болно.

$$X = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ x_{i1} & x_{i2} & \dots & x_{in} \\ \dots & \dots & \dots & \dots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{pmatrix} \tag{11.1}$$

Ямар ч ажиглалтын дүнг дараах хэлбэртэй дүрсэлж болдог гэе.

$$x_{ij} = \mu + \gamma_i + \varepsilon_{ij} = \beta_i + \varepsilon_{ij} \tag{11.2}$$

Үүнд:

μ−ерөнхий дундаж,

γ_i−хүчин зүйлийн i-р түвшний нөлөөгөөр үүссэн хазайлт,

ε_{ij}−i-р түвшний ажиглалтын (хэмжилтийн) алдаа,

$\beta_i = \mu + \gamma_i$−i-р түвшний ажиглалтын дундаж.

ε_{ij} нь тооцогдоогүй үлдсэн бусад санамсаргүй хүчин зүйлүүдийн нөлөөгөөр тодорхойлогдоно. Дисперсийн шинжлэлийн ерөнхий бодлого ёсоор x_{ij} хувьсагчийн ерөнхий хазайлтыг (алдааг) γ хүчин зүйлийн нөлөө ба бусад санамсаргүй хүчин зүйлийн нөлөөгөөр тус тус тодорхойлогдох хоёр нэмэгдэхүүнд задлах ёстой. Үүний тулд ерөнхий дундаж μ ба түвшин бүрийн дундаж β_i-ийн үнэлэлтүүдийг олох шаардлагатай.

i-р түвшний ажиглалтын дунджийг $\overline{x}_i$ гэж тэмдэглэвэл

$$\overline{x}_i = \frac{1}{n}\sum_{j=1}^{n} x_{ij} \tag{11.3}$$

болох ба энэхүү дунджийг бүлгийн дундаж гэж нэрлэнэ. Нийт ажиглалтын арифметик дундаж нь ерөнхий дунджийн (μ-ийн) үнэлэлт болно. Энэ үнэлэлтийг $\overline{X}$ гэж тэмдэглээд цаашид ажиглалтын ерөнхий дундаж гэж нэрлээ.

$$\overline{X} = \frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij} = \frac{1}{m}\sum_{i=1}^{m}\overline{x}_i \tag{11.4}$$

Түүврийн элементүүд ба $\overline{X}$ дунджийн ялгаврын квадратуудын нийлбэр

$$Q = \sum_{i=1}^{m}\sum_{j=1}^{n}(x_{ij} - \overline{X})^2 \tag{11.5}$$

-ыг дараах үл хамаарах хоёр нэмэгдэхүүнд задалж болно.

$$Q = \sum_{i=1}^{m}\sum_{j=1}^{n}(x_{ij} - \overline{X})^2 = n\sum_{i=1}^{m}(\overline{x}_i - \overline{X})^2 + \sum_{i=1}^{m}\sum_{j=1}^{n}(x_{ij} - \overline{x}_i)^2$$

Өөрөөр хэлбэл:

$$Q = Q_1 + Q_2. \tag{11.6}$$

$$Q_1 = n\sum_{i=1}^{m}(\overline{x}_i - \overline{X})^2 \tag{11.7}$$

(11.7) нийлбэрийг бүлэг хоорондын хазайлтын квадратуудын нийлбэр гэж нэрлэх ба бүлэг хоорондын сарнилыг илэрхиийлнэ.

$$Q_2 = \sum_{i=1}^{m}\sum_{j=1}^{n}(x_{ij} - \overline{x}_i)^2 \tag{11.8}$$

нийлбэрийг бүлэг доторхи хазайлтын кватратуудын нийлбэр гэж нэрлэх ба i-р түвшний ажиглалтуудын хоорондох сарнилыг илэрхийлнэ. Q_2 нь тооцогдоогүй үлдсэн бусад санамсаргүй хүчин зүйлүүдийн нөлөөг тодорхойлох бөгөөд үлдэгдэл сарнил гэнэ. Q-г хазайлтуудын квадратын бүтэн нийлбэр буюу ерөнхий нийлбэр гэнэ.

Q ба Q_1, Q_2 нийлбэрүүдийг олсноор, тэдгээрийн харгалзах дисперсүүд болох түүврийн дисперс-S^2 , бүлэг хоорондын дисперс буюу хүчин зүйлийн дисперс-S_1^2, бүлэг доторхи дисперс буюу үлдэгдэл дисперс S_2^2-ийг тус тус олж болно:

$$S^2 = \frac{1}{mn-1}\sum_{i=1}^{m}\sum_{j=1}^{n}(x_{ij} - \overline{X})^2 \tag{11.9}$$

$$S_1^2 = \frac{1}{m-1}\sum_{i=1}^{m}(\overline{x}_i - \overline{X})^2 = \frac{Q_1}{m-1} \tag{11.10}$$

$$S_2^2 = \frac{1}{m(n-1)}\sum_{i=1}^{m}\sum_{j=1}^{n}(x_{ij} - \overline{x}_i)^2 = \frac{Q_2}{m(n-1)} \tag{11.11}$$

γ хүчин зүйлийн нөлөөг тодорхойлохын тулд өгөгдсөн итгэх түвшин α-ийн хувьд, дундаж утгууд нь хүчин зүйлийн түвшин бүр дээр тэнцүү байх тухай таамаглалыг шалгах ёстой. Хэрэв γ хүчин зүйлийн нөлөө түвшин бүр дээр ижил байвал S_1^2 ба S_2^2 нь ерөнхий дисперсийн үнэлэлтүүд болно. Иймээс

$$H_0 : S_1^2 = S_2^2 \tag{11.12}$$

гэсэн таамаглалыг шалгахад хангалттай. Үүний тул

$$F = \frac{S_1^2}{S_2^2} \tag{11.13}$$

гэсэн шинжүүрийг сонгон авна. Анхны таамаглал үнэн үед санамсаргүй хэмжигдэхүүн F нь $k_1 = m-1$ ба $k_2 = m(n-1)$ чөлөөний зэрэг бүхий Фишерийн тархалттай байна. Фишерийн тархалтын таблицаас α, k_1, k_2-утгуудад

харгалзах онолын утга $F_{\alpha;k_1;k_2}$-ийг олно.
Хэрэв $F > F_{\alpha;k_1;k_2}$ нөхцөл биелэгдэж байвал H_0 таамаглалыг үгүйсгэж, γ хүчин зүйл мэдэгдэхүйц нөлөөлж буй тухай дүгнэлт гаргана.
$F < F_{\alpha;k_1;k_2}$ бол H_0 таамаглалыг үгүйсгэх үндэсгүй бөгөөд γ хүчин зүйлийн нөлөөг тооцохгүй байж болно. Хүчин зүйлийн дисперсийг үлдэгдэл дисперсэд харьцуулсан харьцаагаар γ хүчин зүйл хир хүчтэй нөлөөлж буйг тогтооно. Түвшин бүр дэх туршилтын тоо нь ижил байх нэг хүчин зүйлт дисперсийн шинжилгээний үр дүнг дараах хүснэгтэнд нэгтгэн харуулав.

дисперс	квадратуудын нийлбэр	чөлөөний зэрэг	дундаж квадрат	дисперсийн үнэлэлт
хүчин зүйлийн	$n\cdot\sum\limits_i(\overline{x}_i-\overline{X})^2=Q_1$	$m-1$	$\dfrac{\sum\limits_i(\overline{x}_i-\overline{X})^2}{m-1}$	$S_1^2=\dfrac{Q_1}{m-1}$
үлдэгдэл	$\sum\limits_{i,j}(x_{ij}-\overline{x}_i)^2$	$m(n-1)$	$\dfrac{\sum\limits_{i,j}(\overline{x}_{ij}-\overline{x}_i)^2}{m(n-1)}$	$S_2^2=\dfrac{Q_2}{m(n-1)}$
ерөнхий	$\sum\limits_{ij}(x_{ij}-\overline{X})^2=Q$	$mn-1$	$\dfrac{\sum\limits_{ij}(x_{ij}-\overline{X})^2}{mn-1}$	$S^2=\dfrac{Q}{mn-1}$

Хэдийгээр хялбарыг бодож түвшин бүр дэх туршилтын тоог ижил гэж тооцсон ч гэсэн (11.5)÷(11.8) томъёогоор нийлбэрүүдийг олох нь түвэгтэй байдаг. Практикт ихэвчлэн түүврийн утгуудаар Q ба Q_1 нийлбэрийг олж Q_2-ыг тэдгээрийн ялгавраар авдаг. (11.5) ба (11.7) томъёог хувиргавал дараах хэлбэртэй болно.

$$Q=\sum_{i=1}^{m}\Big(\sum_{j=1}^{n}x_{ij}^2\Big)-\frac{1}{mn}\left[\sum_{i=1}^{m}\Big(\sum_{j=1}^{n}x_{ij}\Big)\right]^2 \tag{11.14}$$

$$Q_1=\frac{1}{n}\sum_{i=1}^{m}\Big(\sum_{j=1}^{n}x_{ij}\Big)^2-\frac{1}{mn}\left[\sum_{i=1}^{m}\Big(\sum_{j=1}^{n}x_{ij}\Big)\right]^2 \tag{11.15}$$

$\sum\limits_{j=1}^{n}x_{ij}^2=R_i,\quad L_i=\sum\limits_{j=1}^{n}x_{ij}$ гэж тэмдэглэвэл (11.14), (11.15) томьёо:

$$Q=\sum_{i=1}^{m}R_i-\frac{1}{mn}\left(\sum_{i=1}^{m}L_i\right)^2 \tag{11.16}$$

$$Q_1=\frac{1}{n}\sum_{i=1}^{m}L_i^2-\frac{1}{mn}\left(\sum_{i=1}^{m}L_i\right)^2 \tag{11.17}$$

хэлбэртэй болно. Түүврийн өгөгдлүүдээс хамаарч бага тоонууд дээр ажиллах шаардлага гарвал (11.16) ба (11.17) томьёонд $y_{ij}=x_{ij}-C$ (*) орлуулга хийж хялбарчилж болно. Үүнд, C-ээр ерөнхий дунджийн орчим байх бүхэл тоог сонгон авна. Тэгвэл (11.16) ба (11.17) томьёог дараах байдлаар бичиж болно.

$$Q = \sum_{i=1}^{m} P_i - \frac{1}{N}\left(\sum_{i=1}^{m} T_i\right)^2 \tag{11.18}$$

$$Q_1 = \frac{1}{n}\sum_{i=1}^{m} T_i^2 - \frac{1}{N}\left(\sum_{i=1}^{m} T_i\right)^2 \tag{11.19}$$

Үүнд , $N = mn, \quad P_i = \sum\limits_{j=1}^{n} y_{ij}^2, \qquad T_i = \sum\limits_{j=1}^{n} y_{ij}, \quad (i = \overline{1,m})$

Ингэж хялбарчилсаны дүнд дээрх бүх нийлбэрүүдийг хүснэгт ашиглан гараар бодож болно (Жишээг хар).

Практикт тохиолдох өөр нэг хүндрэл нь хүчин зүйлийн түвшин бүр ижилхэн тооны элементтэй байх албагүй байдагт оршино. Энэ тохиолдолд дээр авч үзсэн томьёонууд хэрхэн өөрчлөгдөхийг тэмдэглэе. i-р түвшний элементийн тоо n_i $(i = \overline{1,m})$ ба түүврийн хэмжээ $N = \sum\limits_{i=1}^{m} n_i$ болохыг анхааран (11.19) томъёог бичвэл:

$$Q_1 = \sum_{i=1}^{m} \frac{1}{n_i} T_i - \frac{1}{N}\left(\sum_{i=1}^{m} T_i\right)^2 \tag{11.20}$$

болох ба P_i, T_i нийлбэрүүд $P_i = \sum\limits_{j=1}^{n_i} y_{ij}^2, \quad T_i = \sum\limits_{j=1}^{n_i} y_{ij}$ хэлбэртэй болно.

Түүврийн дисперсүүд

$$S^2 = \frac{Q}{N}; \quad S_1^2 = \frac{Q_1}{m-1}; \quad S_2^2 = \frac{Q_2}{N-m} \tag{11.21}$$

болно.

Бусад томьёонууд хэвээр хадгалагдана. (11.18)-(11.20) томьёог хэрхэн ашиглахыг дараах жишээгээр үзүүлье.

Жишээ-11.1 Өгөгдсөн түүврийн утгуудаар "хүчин зүйлийн дундаж утгууд тэнцүү" гэсэн таамаглалыг итгэх түвшин $\alpha = 0.05$ үед шалга.

Туршилтын	Хүчин зүйлийн түвшин			
тоо	1	2	3	4
1	-19	-31	-35	-31
2	-28	-33	-32	-27
3	-39	-35	-26	-28
4	-36	-25	-35	-35
5	-44	-28	-30	-40
6	-39	-31	-17	-31

Өгөгдсөн түүврийн хувьд $N = 24$, $n = 6$, $m = 4$. Хүчин зүйлийн түвшин бүрийн дунджийг бодож ($\overline{x}_1 = -34.2$, $\overline{x}_2 = -30.5$, $\overline{x}_3 = -29.2$, $\overline{x}_4 = -32.0$), дараа нь түвшин хоорондын (бүлэг хоорондын) дунджийг бодвол түүврийн ерөнхий дундаж $\overline{x} = -31.5$ болно. $C = -31$ гэж сонгон авч дараах хүснэгтийг зохиое.

i \ j	Хүчин зүйлийн түвшин								$\sum$
	1		2		3		4		
	y_{j1}	y_{j1}^2	y_{j2}	y_{j2}^2	y_{j3}	y_{j3}^2	y_{j4}	y_{j4}^2	
1	12	144	0	0	-4	16	0	0	
2	3	9	-2	4	-1	1	4	16	
3	-8	64	-4	16	5	25	3	9	
4	-5	25	6	36	-4	16	-4	16	
5	-13	169	3	9	1	1	-9	81	
6	-8	64	0	0	14	169		0	
T_i	-19		3		11		-6		-11
T_i^2/n	60.2		1.5		20.2		6		87.9
P_i		475		65		225		122	917

(11.18), (11.19) томъёог ашиглан нийлбэрүүдийг олбол:

$$Q = 917 - \frac{121}{24} = 912.0, \quad Q_1 = 87.9 - \frac{121}{24} = 82.9, \quad Q_2 = 893.1$$

Одоо (11.10)-(11.11) томъёогоор дисперсүүдийн үнэлэлтийг олъё.

$$S_1^2 = \frac{82.9}{4-1} = 27.63; \quad S_2^2 = \frac{829.1}{24-4} = 41.46$$

(11.13) ёсоор шинжүүрийн утга нь $F = \dfrac{27.63}{41.46} = 0.67$ болно. Манай тохиолдолд $\alpha = 0.05$, $k_1 = 3$, $k_2 = 20$ учир таблицаас $F_{0.05,3,20} = 3.10$ гэж олдоно.

$F < 3.10$ учраас шинжүүрийн утга критик мужид харьяалагдахгүй болж анхны таамаглалыг үгүйсгэх үндэсгүй болно. Ингэхлээр авч үзэж буй хүчин зүйлийн нөлөөг тооцохгүй байж болно. Туршилтын үр дүнгүүд ижил дисперс бүхий хэвийн тархалтай байх ёстой учраас анхааралтай хандаж тодорхой шалгалт нотолгоон дээр үндэслэн эцсийн дүгнэлтээ гаргах шаардлагатай болохыг тэмдэглэе.

11.3 Хоёр хүчин зүйлт дисперсийн шинжилгээ

Хэрэв туршилтын үр дүнд хэд хэдэн хүчин зүйл зэрэг нөлөөлж байвал олон хүчин зүйлт дисперсийн шинжилгээ хийх шаардлагатай бөгөөд энэ тохиолдолд хүчин зүйл хоорондын харилцан үйлчлэлийг тооцох учраас нэг хүчин зүйлт шинжлэлээс онцлог ялгаатай.

Нэгэн зэрэг үйлчлэх хоёр хүчин зүйлийн нөлөөг тооцох бодлого авч үзье. Үүний тулд ижил төрлийн хэд хэдэн суурь машин, нэг төрлийн хэд хэдэн хэсэг түүхий эд байдаг гэе. Эдгээр суурь машин болон түүхий эдийн хэсгүүд үйлдвэрлэн гаргаж буй бүтээгдэхүүний чанарт нөлөөлж буй эсэх тухай бодлого нь хоёр хүчин зүйлт дисперсийн шинжлэлийн нийтлэг бодлогын нэг болно. Дисперсийн шинжлэл хийх үндсэн шаардлагууд биелэгдсэн гэж саная. Суурь машинуудаас үзүүлэх нөлөөг A хүчин зүйл, түүхий эдийн нөлөөг B хүчин зүйл болгон сонгож авья. r тооны суурь машин, v төрлийн түүхий эд байдаг гэвэл A хүчин зүйл r түвшинтэй, B хүчин зүйл v түвшинтэй болно.

Ажиглалтын матрицыг дараах хүснэгтээр үзүүлэв. (i, j)-р нүдэнд A ба B хүчин зүйлүүдийн харгалзан i ба j-р түвшний ажиглагдсан утга x_{ij}-г байрлуулсан болно.

	B факторын түвшин				
A факторын түвшин	B_1	B_2	$\cdots B_j \cdots$	B_v	$\overline{x}_{i*}$
A_1	x_{11}	x_{12}	$\cdots x_{1j} \cdots$	x_{1v}	$\overline{x}_{1*}$
A_2	x_{21}	x_{22}	$\cdots x_{2j} \cdots$	x_{2v}	$\overline{x}_{2*}$
A_i	x_{i1}	x_{i2}	$\cdots x_{ij} \cdots$	x_{iv}	$\overline{x}_{i*}$
A_r	x_{r1}	x_{r2}	$\cdots x_{rj} \cdots$	x_{rv}	$\overline{x}_{r*}$
$\overline{x}_{*j}$	x_{*1}	x_{*2}	$\cdots \overline{x}_{*j} \cdots$	$\overline{x}_{*v}$	$\overline{X}$

Хялбарчлах зорилгоор, эхлээд (i, j)-р нүдэнд ажиглалтын зөвхөн ганц x_{ij} утга байдаг гэе. Түүнчлэн A ба B хүчин зүйлүүд хоорондоо харилцан үйлчлэл үзүүлдэггүй (хамааралгүй) гэе. A хүчин зүйлийн i-р түвшний туршилтын дунджийг β_{jA}, B хүчин зүйлийн j-р түвшний туршилтын дунджийг β_{jB} гэвэл ямарч туршилтыг дараах хэлбэртэй дүрсэлж болно.

$$x_{ij} = \mu + \gamma_i + g_j + e_{ij} \tag{11.22}$$

Үүнд, $\mu -$ерөнхий дундаж,

$\gamma_i - A$ хүчин зүйлийн i-р түвшний нөлөөгөөр үүсэх хазайлт,

$g_j - B -$хүчин зүйлийн j-р түвшний нөлөөгөөр үүсэх хазайлт,

$e_{ij} - x_{ij}$ утгын ажиглалтын алдаа (зөвхөн ганц x_{ij} ажиглалтын үед алдаа тэгтэй тэнцүү). Хүчин зүйлийн харилцан үйлчлэлийг тооцоогүй (11.22) хэлбэрийн загварыг шугаман загвар гэнэ.

μ, β_{iA}, β_{iB} дунджуудын үнэлэлтийг бодвол:

$$\overline{X} = \frac{1}{rv} \sum_{i=1}^{r} \sum_{j=1}^{v} x_{ij} \tag{11.23}$$

$$\overline{x}_{i*} = \frac{1}{v}\sum_{j=1}^{v} x_{ij}\ , \quad i = \overline{1, r} \tag{11.24}$$

$$\overline{x}_{*j} = \frac{1}{r}\sum_{i=1}^{r} x_{ij}\ , \quad j = \overline{1, v} \tag{11.25}$$

Хоёр хүчин зүйлт дисперсийн шинжилгээнд, түүврийн ерөнхий дисперсийг A ба B хүчин зүйл тус бүрийн нөлөөтэй уялдсан хэсгүүд, мөн тооцогдоогүй үлдсэн бусад санамсаргүй хүчин зүйлийн нөлөөтэй уялдсан хэсэг гэсэн гурван нэмэгдэхүүнд задалдаг. Өөрөөр хэлбэл (11.6) томъёо дараах хэлбэртэй болно:

$$Q = Q_1 + Q_2 + Q_3 \tag{11.26}$$

Үүнд:

$$Q_1 = v\sum_{i=1}^{r}(\overline{x}_{i*} - \overline{X})^2 \tag{11.27}$$

Q_1 нь мөрийн дунджууд ба ерөнхий дунджийн ялгаврын квадратуудын нийлбэр бөгөөд A хүчин зүйлийн өөрчлөлтөөр тодорхойлогдоно.

$$Q_2 = r\sum_{j=1}^{v}(\overline{x}_{*j} - \overline{X})^2 \tag{11.28}$$

нь баганы дунджууд ба ерөнхий дунджийн ялгаврын квадратуудын нийлбэр бөгөөд B хүчин зүйлийн өөрчлөлтөөр тодорхойлогдоно.

$$Q_3 = \sum_{i=1}^{r}\sum_{j=1}^{v}(x_{ij} - \overline{x}_{i*} - \overline{x}_{*j} + \overline{X})^2 \tag{11.29}$$

Q_3 нь квадратуудын үлдэгдэл нийлбэр гэж нэрлэгдэх ба тооцогдоогүй үлдсэн санамсаргүй хүчин зүйлүүдийн өөрчлөлтөөр тодорхойлогдоно. Q-г ерөнхий нийлбэр гэнэ.

Харгалзах дисперсүүдийн үнэлэлтийг бодвол:

$$S^2 = \frac{1}{rv-1}\sum_{i=1}^{r}\sum_{j=1}^{v}(x_{ij} - \overline{X})^2 = \frac{Q}{rv-1} \tag{11.30}$$

$$S_1^2 = \frac{v}{r-1}\sum_{i=1}^{r}(\overline{x}_{i*} - \overline{X})^2 = \frac{Q_1}{r-1} \tag{11.31}$$

$$S_2^2 = \frac{r}{v-1}\sum_{j=1}^{v}(\overline{x}_{*j} - \overline{X})^2 = \frac{Q_2}{v-1} \tag{11.32}$$

$$S_3^2 = \frac{1}{(r-1)(v-1)}\sum_{i=1}^{r}\sum_{j=1}^{v}(x_{ij} - \overline{x}_{i*} - \overline{x}_{*j} + \overline{X})^2 = \frac{Q}{(r-1)(v-1)} \tag{11.33}$$

Хоёр хүчин зүйлт дисперсийн шинжилгээнд A ба B хүчин зүйлийн нөлөөг тооцохдоо хүчин зүйл тус бүрийн дисперсийг үлдэгдэл дисперст харьцуулна. Тухайлбал A хүчин зүйлийн нөлөөг тооцоход $F_A = \dfrac{S_1^2}{S_3^2}$ гэсэн шинжүүрийг ашиглах ба энэ санамсаргүй хэмжигдэхүүн $r–1$ ба $(r–1)(v–1)$ чөлөөний зэрэг бүхий Фишерийн тархалттай байна. B хүчин зүйлийн нөлөөг тооцохдоо $v–1$ ба $(r-1)(v-1)$ чөлөөний зэрэг бүхий $F_B = \dfrac{S_2^2}{S_3^2}$ гэсэн шинжүүрийг ашиглана. Факторуудын нөлөөний тухай эцсийн шийдвэр гаргах арга нь нэг хүчин зүйлт дисперсийн шинжилгээтэй адил байна. Дээр авч үзсэн, түвшин бүр дээрээ зөвхөн ганц туршилттай хоёр хүчин зүйлт дисперсийн шинжлэлийн үндсэн үр дүнг дараах хүснэгтэнд нэгтгэн харуулав.

Дисперс	Квадратуудын нийлбэр	Чөлөөний зэрэг	Дисперсүүдийн үнэлэлт
A хүчин зүйлийн	$Q_1 = v \cdot \sum\limits_{i=1}^{r} (\overline{x}_{i*} - \overline{X})^2$	$r-1$	$S_1^2 = \dfrac{Q_1}{r-1}$
B хүчин зүйлийн	$Q_2 = r \cdot \sum\limits_{j=1}^{v} (\overline{x}_{*j} - \overline{X})^2$	$v-1$	$S_2^2 = \dfrac{Q_2}{v-1}$
Үлдэгдэл	$Q_3 = \sum\limits_{i=1}^{r} \sum\limits_{j=1}^{v} (\overline{x}_{ij} - \overline{x}_{i*} - \overline{x}_{*j} + \overline{X})^2$	$(r-1)(v-1)$	$S_3^2 = \dfrac{Q_3}{(r-1)(v-1)}$
Ерөнхий	$Q = \sum\limits_{i=1}^{r} \sum\limits_{j=1}^{v} (x_{ij} - \overline{X})^2$	$rv-1$	$S^2 = \dfrac{Q}{rv-1}$

Жишээ-11.2 Бидэнд дараах ажиглалтын матриц өгөгдсөн гэе.

	B факторын түвшин			
A факторын түвшин	B_1	B_2	B_3	$\overline{x}_{i*}$
A_1	1	2	3	2
A_2	5	6	10	7
$\overline{x}_{*j}$	3	4	6.5	4.5

Манай тохиолдолд $r = 2,\ v = 3,\ n = rv = 6$ болно. Хүснэгтийн сүүлийн багана ба мөрийн элементүүдийг бодож олсон болно. Тухайлбал: A факторын A_2 түвшний хувьд (мөрийн дундаж) $i = 2,\ \ j = \overline{1,3}$ тул $\overline{x}_{2*} = \dfrac{5+6+10}{3} = 7$ гэх мэт. Ерөнхий дундаж $\overline{X} = 4.5$ (11.27)÷(11.29) томъёогоор $Q_1,\ Q_2,\ Q_3$ нийлбэрүүдийг олбол: $Q_1 = 37.5,\ Q_2 = 3,\ Q_3 = 53.5$ Дисперсүүдийн үнэлэлтийг бодъё.

$S_1^2 = \dfrac{37.5}{1} = 37.5; \quad S_2^2 = \dfrac{13}{2} = 6.5, \quad S_3^2 = \dfrac{3}{1 \cdot 2} = 1.5$

Гаргаж авсан үр дүнгүүдийг дараах хүснэгтэд нэгтгэвэл:

Дисперс	Квадратуудын нийлбэр	Чөлөөний зэрэг	Дисперсүүдийн үнэлэлт
A хүчин зүйлийн	37.5	1	37.5
B хүчин зүйлийн	13.0	2	6.5
Үлдэгдэл дисперс	3.0	2	1.5
Ерөнхий дисперс	53.5	5	10.7

Одоо F_A, F_B шинжүүрийн утгуудыг бодьё.

$$F_A = \frac{S_1^2}{S_3^2} = \frac{37.5}{1.5} = 25,\ F_B = \frac{S_2^2}{S_3^2} = \frac{6.5}{1.5} = 4,3.$$

Итгэх түвшин $\alpha = 0.05$ ба $k_1 = 1,\ k_2 = 2$ чөлөөний зэргийн хувьд Фишерийн таблицын утгуудыг олбол $F_{\alpha,A} = 18.5$;
$F_{\alpha,B} = 19.0$ болно. Ийнхүү $F_A > F_{\alpha,A}$; $F_B < F_{\alpha,B}$ харьцаа биелж байгаа учраас A хүчин зүйл судалж буй хэмжигдэхүүнд мэдэгдэхүйц нөлөөтэй, B хүчин зүйлийн нөлөөг тооцохгүй байж болно гэсэн дүгнэлтийг гаргаж болно.

Бид хоёр хүчин зүйлт дисперсийн шинжилгээний тухайн тохиолдлыг авч үзлээ. Бүр ерөнхий тохиолдолд A хүчин зүйлийн i-р түвшин, B хүчин зүйлийн j-р түвшин дээр хийсэн туршилтын тоо ((x_i, x_j)хос утгын тоо) нэг биш хэд хэд байж болохын гадна эдгээр тоонууд өөр өөр байж болно. Ийм туршилтыг давталттай туршилт буюу зэрэгцээ туршилт бүхий ажиглалт (туршилт) гэнэ. Үүнээс гадна A ба B хүчин зүйлүүд хоорондоо харилцан үйлчлэлтэй байж болно.

Ерөнхий тохиолдолд хоёр хүчин зүйлт шинжилгээний ажиглалтын утгыг дараах хэлбэртэй бичиж болно.

$$x_{ijk} = \mu + \gamma_i + g_j + \nu_{ij} + e_{ijk} \tag{11.34}$$

Үүнд, μ−ерөнхий дундаж,
γ_i−A хүчин зүйлийн i-р түвшний нөлөөтэй уялдсан хазайлт,
g_j−B хүчин зүйлийн j-р түвшний нөлөөтэй уялдсан хазайлт,
ν_{ij}−A ба B хүчин зүйлийн харилцан үйлчлэлээр үүссэн хазайлт,
e_{ijk}−i ба j түвшний хос дээрх зэрэгцээ туршилтуудын алдаа.
(11.34) загварыг шугаман биш загвар гэнэ.
(i, j) түвшний хос бүр дээр нэгэн ижил n тооны зэрэгцээ туршилт хийсэн гэвэл хоёр хүчин зүйлт дисперсийн шинжлэлийн үндсэн адилтгал

$$Q = Q_1 + Q_2 + Q_3 + Q_4 \tag{11.35}$$

хэлбэртэй болно.
Үүнд:

$$Q=\sum_{i=1}^{r}\sum_{j=1}^{v}\sum_{k=1}^{n}(x_{ijk}-\overline{X})^2 \tag{11.36}$$

$$Q_1=nv\sum_{i=1}^{r}(\overline{x}_{i**}-\overline{X})^2 \tag{11.37}$$

$$Q_2=nr\sum_{j=1}^{n}(\overline{x}_{*j*}-\overline{X})^2 \tag{11.38}$$

$$Q_3=n\sum_{i=1}^{r}\sum_{j=1}^{v}(\overline{x}_{ij*}-\overline{x}_{i**}-\overline{x}_{*j*}+\overline{X})^2 \tag{11.39}$$

$$Q_4=\sum_{i=1}^{r}\sum_{j=1}^{v}\sum_{k=1}^{n}(x_{ijk}-\overline{x}_{ij*})^2 \tag{11.40}$$

Q_3 нь A ба B хүчин зүйлийн харилцан үйлчлэлийг үнэлэх квадратуудын нийлбэр
Q_4 нь бусад хүчин зүйлүүдийн нөлөөг үнэлэх квадратуудын нийлбэр юм.

$$\left.\begin{array}{l}
\overline{x}_{ij*}=\dfrac{1}{n}\sum\limits_{k=1}^{n}x_{ijk}-(i,j)\text{ түвшний хос дээрх}\\
\qquad\qquad\qquad\text{зэрэгцээ туршилтын дундаж}\\
\overline{x}_{i**}=\dfrac{1}{v}\sum\limits_{j=1}^{v}\overline{x}_{ij*}-\text{ мөрийн дундаж}\\
\overline{x}_{*j*}=\dfrac{1}{r}\sum\limits_{i=1}^{r}\overline{x}_{ij*}-\text{ баганы дундаж}\\
\overline{X}=\dfrac{1}{rv}\sum\limits_{i=1}^{r}\sum\limits_{j=1}^{v}\overline{x}_{ij*}-\text{ ерөнхий дундаж}
\end{array}\right\} \tag{11.41}$$

Харгалзах дисперсүүдийн үнэлэлт ба чөлөөний зэргийг дараах хүснэгтэд харуулав.

Дисперс	Квадратуудын нийлбэр	Чөлөөний зэрэг	Дисперсүүдийн үнэлэлт
A хүчин зүйлийн	$Q_1=v\cdot n\sum\limits_{i=1}^{r}(\overline{x}_{i**}-\overline{X})^2$	$r-1$	$S_1^2=\dfrac{Q_1}{r-1}$
B хүчин зүйлийн	$Q_2=r\cdot n\sum\limits_{j=1}^{v}(\overline{x}_{*j*}-\overline{X})^2$	$v-1$	$S_2^2=\dfrac{Q_2}{v-1}$
Харилцан үйлчлэлийн	$Q_3=n\sum\limits_{i=1}^{r}\sum\limits_{j=1}^{v}(\overline{x}_{ij*}-\overline{x}_{i**}-\overline{x}_{*j*}+\overline{X})^2$	$(v-1)(r-1)$	$S_3^2=\dfrac{Q_3}{(r-1)(v-1)}$
Үлдэгдэл дисперс	$Q_4=\sum\limits_{i=1}^{r}\sum\limits_{j=1}^{v}\sum\limits_{k=1}^{n}(x_{ijk}-\overline{x}_{ij*})^2$	$rv(n-1)$	$S_4^2=\dfrac{Q_4}{rv(n-1)}$
Ерөнхий дисперс	$Q=\sum\limits_{i=1}^{r}\sum\limits_{j=1}^{v}\sum\limits_{k=1}^{n}(x_{ijk}-\overline{X})^2$	$rvn-1$	$S^2=\dfrac{Q}{rvn-1}$

Шугаман биш загварын хувьд дисперсийн шинжилгээ хийх дараалал шугаман загвартай адил. Харин Q_4 гэсэн шинэ нэмэгдэхүүн гарч ирсэн, Q_3 нийлбэрийн бүтцэнд өөрчлөлт (x_{ij}-ийн оронд $\overline{x}_{ij*}$-г авсан) орсон зэрэг онцлог байна. Q_4 нийлбэр нь (i, j) түвшний хос дээр хэд хэдэн зэрэгцээ туршилт хийсэнтэй холбоотой гарч ирсэн бөгөөд бусад санамсаргүй хүчин зүйлүүдийн нөлөөг илэрхийлнэ. Ийм учраас A ба B хүчин зүйлийн нөлөөг тодорхойлохдоо эдгээр хүчин зүйл тус бүрийн дисперсийг санамсаргүй хүчин зүйлийн дисперст харьцуулдаг. Өөрөөр хэлбэл, $F_A = \dfrac{S_1^2}{S_4^2}$, $F_B = \dfrac{S_2^2}{S_4^2}$, $F_{AB} = \dfrac{S_3^2}{S_4^2}$-гэсэн шинжүүрүүдийг авч үзэж, эдгээр утгуудыг таблицын утгатай жиших замаар эцсийн дүгнэлт гаргана.

Жишээ-11.3 Нэхмэлийн үйлдвэрт ээрсэн утасны чанарт нөлөөлж буй хүчин зүйлүүдийг илрүүлэх нь чухал ач холбогдолтой байдаг. Дараах хүснэгтэнд, утас тасрах ачааллын хэмжээ нь машины тохируулга ба түүхий эдийн хэсгүүдээс хамаарах хамаарал өгчээ.

Машины тохируулгын түвшин	Түүхий эдийн хэсэг									
	B_1					B_2				
A_1	190	260	170	170	170	190	150	210	150	150
A_2	150	250	220	140	180	230	190	200	190	200
A_3	190	185	135	195	195	150	170	150	170	180

Машины тохируулгын түвшин бүрд түүхий эдийн хэсэг бүрээс тав тавыг түүвэрлэн авч тасрах ачааллыг хэмжжээ. Тасрах ачааллын хэмжээнд дээрх 2 хүчин зүйл хир их нөлөөтэй болохыг тодорхойл.
(11.41) томьёогоор дундаж утгуудыг бодож үр дүнг дараах хүснэгтэнд үзүүлье.

Машины тохируулгын түвшин	Түүхий эдийн хэсэг		
	B_1	B_2	x_{j**}
A_1	$\overline{x}_{11*} = 192$	$\overline{x}_{12*} = 170$	$\overline{x}_{1**} = 181$
A_2	$\overline{x}_{21*} = 188$	$\overline{x}_{22*} = 202$	$\overline{x}_{2**} = 195$
A_3	$\overline{x}_{31*} = 180$	$\overline{x}_{32*} = 164$	$\overline{x}_{3**} = 172$
$\overline{x}_{*j*}$	$\overline{x}_{*1*} = 186.7$	$\overline{x}_{*2*} = 178.7$	$\overline{X} = 182.7$

Квадратуудын нийлбэрийг олбол:
$Q_1 = 2686.6,\ Q_2 = 480,\ Q_3 = 1860,\ Q_4 = 22360,\ Q = 27386,7.$
Тэгвэл дисперсүүдийн үнэлэлт: $S_1^2 = 2686.7,\ S_2^2 = 240,\ S_3^2 = 930,$

$S_4^2 = 931.7$, $S^2 = 944,4$. болно.
$F_A = 2686.7/931.7 = 2.88$, $F_B = 240/931.7 = 0.26$ F_A шинжүүрийн хувьд чөлөөний зэргүүд нь $k_1 = 1$, $k_2 = 24$ тул $\alpha = 0.05$ итгэх түвшинтэй үед $F_{\alpha,A} = 4.26$, F_B шинжүүрийн таблицын утга $k_1 = 2$, $k_2 = 24$ ба $\alpha = 0.05$ үед $F_{\alpha,B} = 3,40$. Ийнхүү $F_a < F_{\alpha,A}$; $F_B < F_{\alpha,B}$ учир машины тохируулгын түвшин ба түүхий эдийн хэсэг нь утас тасрах ачааралд нөлөөлөхгүй гэсэн дүгнэлтэнд хүрч байна.

Ажиглалтын матрицын зарим нүднүүд тэнцүү биш тооны элементтэй үед, өөрөөр хэлбэл хос түвшин бүр дээр ялгаатай тооны зэрэгцээ туршилт хийсэн нөхцөлд дисперсийн шинжлэл дээр авч үзсэнээс ялгаатай арга барилаар гүйцэтгэгдэнэ. Ажиглалтын матрицын зарим нүднүүд хоосон байх үед ч өөрийн онцлог ялгаатай бөгөөд энэ хэсэгт дээрх 2 тохиолдлыг авч үзээгүй болно.

Бүлэг 12

Корреляцийн шинжилгээ

12.1 Функцэн ба статистик хамаарал. Регрессийн муруйн тухай ойлголт

Байгаль ертөнцийн аливаа үзэгдлүүдийн холбоо хамаарал нийлмэл бөгөөд янз бүр байх боловч тэдгээрийг тодорхой байдлаар ангилж болдог. Техник болон байгалийн шинжлэх ухаанд, нэг хувьсах хэмжигдэхүүний утга бүрд нөгөө хэмжигдэхүүний тодорхой утга харгалзах хамаарал буюу функцэн хамаарлын тухай ихэвчлэн авч үздэг билээ. Гэтэл тус бүрийн үйлчлэл нь дангаараа үл мэдрэгдэх, төгсгөлгүй олон тооны хүчин зүйлүүдийн үйлчлэлийн үр дүнд явагддаг үзэгдэл процессууд бодит ертөнцөд байдаг. Энэ тохиолдолд дээрх функцэн хамаарал утгаа алдах бөгөөд судалж буй физик систем яг тодорхой төлөв байдалд шилжихгүй харин түүний боломжит бүх төлөв байдлын аль нэгд шилжинэ.
Ийнхүү нэг хувьсах хэмжигдэхүүний тодорхой утганд нөгөө хувьсах хэмжигдэхүүний тархалт харгалзаж байвал тэдгээрийг статистик хамааралтай гэнэ. Өөрөөр хэлбэл, нэг хувьсах хэмжигдэхүүний утга нөгөө хувьсах хэмжигдэхүүний тархалтын хуулиас хамаарна гэсэн үг. Судалж буй хувьсах хэмжигдэхүүн хяналтанд үл өртөх юмуу эсвэл тооцоологүй хүчин зүйлүүдийн нөлөөлөлд орох, түүнчлэн түүний утгуудыг хэмжихэд санамсаргүй алдаа зайлшгүй гарч болдог зэрэгтэй уялдан статистик хамаарлыг авч үзэх шаардлага гарч байгаа юм. Статистик хамааралтай хэмжигдэхүүний жишээ гэвэл: бордоо ба ургацын хэмжээ, хүний биеийн жин ба өндрийн хэмжээ, ургамлын масс ба түүний үрийн масс гэх мэт.
Дээрх тодорхойлолтоос үзвэл хоёр хэмжигдэхүүний статистик хамаарлыг судлахын тулд хоёр хэмжээст тархалтын хуулийн аналитик хэлбэрийг мэдэх шаардлагатай болж байна.
Гэтэл хязгаарлагдмал хэмжээтэй түүврийн утгуудаар хоёр хэмжээст тархалтын хуулийг байгуулахад түвэгтэй бөгөөд мэдэгдэхүйц алдаа гарахад хүрнэ. Ийм учраас практикт ξ ба η санамсаргүй хэмжигдэхүүний статис-

тик хамаарлыг судлахын тулд аль нэгнийх нь утга нөгөөгийнхөө нөхцөлт математик дунджаас хамаарах хамаарлыг авч үздэг. Ийнхүү нэг санамсаргүй хэмжигдэхүүний утга ба нөгөө санамсаргүй хэмжигдэхүүний нөхцөлт мат.дунджийн функцэн хамаарлыг ковариац буюу корреляц хамаарал гэнэ. Корреляц хамаарлыг дараах томьёогоор дүрсэлж болно.

$$M(\eta/\xi = x) = \varphi(x) \tag{12.1}$$

$$M(\xi/\eta = y) = \psi(y) \tag{12.2}$$

Үүнд, $M(\eta/\xi = x) - \xi$ нь x гэсэн утга авах үеийн η санамсаргүй хэмжигдэхүүний математик дундаж, $M(\xi/\eta = y) - \eta$ нь y гэсэн утга авах үеийн ξ санамсаргүй хэмжигдэхүүний математик дундаж.
(12.1) тэгшитгэлийг η-ийн ξ дээрх, (12.2) тэгшитгэлийг ξ-ийн η дээрх регрессийн тэгшитгэлүүд (загварын) гэнэ. Регрессийн загварын тэгшитгэл нь статистик хамааралтай хоёр хэмжигдэхүүний корреляц хамаарлыг ерөнхий хэлбэрээр илэрхийлсэн тэгшитгэл юм.
$\varphi(x)$, $\psi(y)$-функцүүдийг регрессийн загварын функцүүд, тэдгээрийн графикийг регрессийн загварын муруй гэнэ.

Дээр тэмдэглэсэн ёсоор (ξ, η) системийн тархалтын хуулийн аналитик илэрхийллийг олох боломжгүй учраас практикт регрессийн загварын функцүүдийг тэдгээрийн хамгийн сайн үнэлэлтээр орлуулдаг. Өөрөөр хэлбэл, хоёр хэмжээст санамсаргүй хэмжигдэхүүний ажиглагдсан хос утгууд (x_i, y_j)-ээр η-ийн ξ дээрх регрессийн муруй

$$\overline{y}_x = \widehat{\varphi}(x, b_0, b_1, \ldots, b_q) \tag{12.3}$$

-ыг байгуулна.
Үүнд, $\overline{y}_x$ нь $\xi = x$ утга авах үеийн η санамсаргүй хэмжигдэхүүний нөхцөлт математик дундаж (бүлгийн дундаж), $b_0, b_1, \ldots, b_q$ нь регрессийн муруйн параметрүүд. Үүнчлэн ξ -ийн η дээрх регрессийн муруй

$$\overline{x}_y = \widehat{\psi}(y, c_0, c_1, \ldots, c_q) \tag{12.4}$$

-ийг байгуулна. (12.3) ба (12.4) муруйг туршилтын регрессийн муруйнууд гэх ба цаашид товчоор, регрессийн муруйнууд гэж нэрлэнэ.

12.2 Корреляцийн шинжилгээний үндсэн бодлого. Корреляцийн коэффициент.

Санамсаргүй хэмжигдэхүүн хоорондын статистик хамаарлыг корреляцийн ба регрессийн шинжлэлийн аргаар судалж эдгээр аргуудын тусламжтайгаар практикт олон янзын бодлогуудыг шийдэж болдог.
Корреляцийн шинжлэлийн үндсэн бодлого нь

1. Санамсаргүй хэмжигдэхүүнүүдийн хоорондын корреляц хамаарлын хэлбэрийг тогтоох;

2. Корреляц хамаарлын хүч (хүчтэй, сул) буюу нягтралыг тогтоох явдал юм.

Хамаарлын хэлбэрийг тогтоохдоо корреляцийн янз бүрийн (хос, хамтын, тухайн) коэффициентуудын цэгэн ба завсран үнэлэлтийг ашиглахын гадна регрессийн тэгшитгэлд үнэлэлт өгнө. Корреляц хамаарлын хүч (нягтрал) нь санамсаргүй хэмжигдэхүүний утга өөрийнхөө нөхцөлт математик дунджаас сарних сарнилын хэмжээгээр тодорхойлогдоно. Корреляцийн шинжлэлийг:

1. Хувьсах хэмжигдэхүүн бүр санамсаргүй байх,

2. Санамсаргүй хэмжигдэхүүнүүд нь хамтын хэвийн тархалттай байх

гэсэн хоёр үндсэн шаардлага биелэгдсэн нөхцөлд явуулна. Өөрөөр хэлбэл, туршилт ба ажиглалтын утгууд нь олон хэмжээст хэвийн тархалттай эх олонлогоос сонгогдсон, санамсаргүй хэмжигдэхүүн байх тохиолдолд корреляцийн шинжилгээний аргуудыг хэрэглэнэ гэсэн үг юм.

Одоо корреляцийн шинжилгээний хамгийн хялбар тохиолдол болох хоёр хэмжээст загварыг авч үзье.

ξ, η санамсаргүй хэмжигдэхүүнүүд хамтын хэвийн тархалттай байг. Хоёр хэмжээст хамтын тархалтын нягтыг ерөнхий тохиолдолд бичвэл:

$$f(x,y)=\frac{1}{2\pi\sigma_\xi\sigma_\eta\sqrt{1-\rho^2}}e^{-L(x,y)}$$

$$L(x,y)=\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-a_\xi}{\sigma_\xi}\right)^2-2\rho\frac{x-a_\xi}{\sigma_\xi}\cdot\frac{y-a_\eta}{\sigma_\eta}+\left(\frac{y-a_\eta}{\sigma_\eta}\right)^2\right] \tag{12.5}$$

$a_\xi,\ a_\eta - \xi$ ба η-санамсаргүй хэмжигдэхүүний математик дунджууд,

$\sigma_\xi,\ \sigma_\eta - \xi$ ба η-санамсаргүй хэмжигдэхүүний дисперсүүд,

$\rho-$эх олонлогийн корреляцийн коэффициент.

Корреляцийн коэффициент ρ-г дараах томъёогоор олно.

$$\rho=\frac{K_{\xi\eta}}{\sigma_\xi\sigma_\eta} \tag{12.6}$$

Үүнд,

$$K_{\xi\eta}=M[(\xi-M\xi)(\eta-M\eta)] \tag{12.7}$$

$K_{\xi\eta}-$г ξ,η санамсаргүй хэмжигдэхүүнүүдийн корреляцийн момент буюу ковариац гэж нэрлэнэ. (12.5) томъёнд орсон таван параметр нь ξ,η санамсаргүй хэмжигдэхүүний корреляц хамаарлын тухай бүрэн мэдээлэл өгнө. $\rho-$параметр нь ξ,η-санамсаргүй хувьсагчдын эх олонлог дахь корреляц хамаарлын хүчийг тодорхойлно. Магадлалын онолд, ξ,η-санамсаргүй хэмжигдэхүүнүүд хамтын хэвийн тархалттай үед регрессийн загварын тэгшитгэл бүр дараах шугаман функцээр илэрхийлэгдэнэ гэж баталдаг.

$$\varphi(x)=a_\eta+\rho\frac{\sigma_\eta}{\sigma_\xi}(x-a_\xi) \tag{12.8}$$

$$\psi(y) = a_\xi + \rho\frac{\sigma_\xi}{\sigma_\eta}(y - a_\eta) \tag{12.9}$$

$\rho \cdot \dfrac{\sigma_\eta}{\sigma_\xi}$ ба $\rho \cdot \dfrac{\sigma_\xi}{\sigma_\eta}$ тоог регрессийн коэффициентүүд гэнэ. Эндээс, санамсаргүй хэмжигдэхүүн хоорондын хамаарал зөвхөн шугаман үед корреляцийн коэффициент ρ нь хамаарлын хүчийг (нягтралыг) илэрхийлнэ гэсэн чухал дүгнэлт мөрдөж буй юм.

Корреляцийн коэффициентийн үнэлэлт нь түүврийн корреляцийн коэффициент r бөгөөд түүнийг олохын тулд a_ξ, a_η, σ_ξ^2, σ_η^2 параметрүүдийн үнэлэлтийг олох шаардлагатай.

$$r = \frac{\sum\limits_{i=1}^{n}(x_i - \overline{X})(y_i - \overline{Y})}{nS_\xi S_\eta} \tag{12.10}$$

Үүнд, $\overline{X}, \overline{Y}-$түүврийн дунджууд, $S_\xi, S_\eta-$түүврийн стандарт хазайлтууд (X бүлгийн 10.3, 10.4 дэд бүлгийн томъёог үз.)

ρ-г хос корреляцийн коэффициент, r-ийг түүврийн хос корреляцийн коэффициент гэж нэрлэнэ. Эдгээр үнэлэлтүүдийг (12.8) ба (12.9) томъёонд тавибал харгалзан η-ийн ξ дээрх, ξ-ийн η дээрх регрессийн тэгшитгэлүүд гарна.

$$\overline{y}_x = \overline{Y} + r\frac{S_\eta}{S_\xi}(x - \overline{X}) \tag{12.11}$$

$$\overline{x}_y = \overline{X} + r\frac{S_\xi}{S_\eta}(y - \overline{Y}) \tag{12.12}$$

Корреляцийн коэффициентийг ашиглах нь санамсаргүй хэмжигдэхүүн хоорондын холбоо хамаарлыг тогтоох хамгийн өргөн дэлгэр тархсан арга юм.

Корреляцийн коэффициент хэмжилтийн нэгж ба тооллын эхийн сонголтоос хамаарахгүй бөгөөд дараах чанаруудыг хангана.

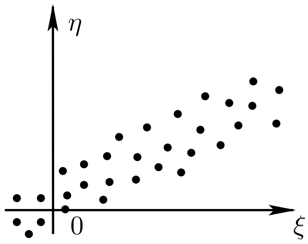
1. Түүврийн хэмжээ хүрэлцээтэй их үед хос корреляцийн коэффициентийн тоон утга $[-1, 1]$ завсарт байна. Өөрөөр хэлбэл,

$$-1 \leq r \leq 1 \tag{12.13}$$

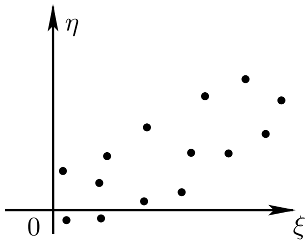
$|r|$ нь 1-д ойртох тутам корреляц хамаарал сулаас дунд зэрэг, мэдэгдэхүйц, хүчтэй, бүр хүчтэй гэсэн чиглэлд өөрчлөгдөнө. $r > 0$ үед (шууд хамааралтай) эерэг корреляц, $r < 0$ үед (урвуу хамааралтай) сөрөг корреляц гэнэ.

Эерэг корреляцтай үед нэг санамсаргүй хэмжигдэхүүний утга өсөхөд нөгөө санамсаргүй хэмжигдэхүүн дунджаараа өснө, сөрөг корреляцтай үед буурна. (Зураг 12.1, Зураг 12.2, Зураг 12.3)

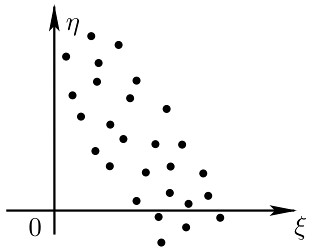
Эндээс үзвэл, корреляцийн коэффициентийн тэмдэг нь нэг санамсаргүй хэмжигдэхүүний нөхцөлт математик дундаж нөгөө санамсаргүй хэмжигдэхүүнээсээ хамаарах чиглэлийн өсөх буурахыг зааж байна.



Зураг 12.1: хүчтэй, эерэг корреляц



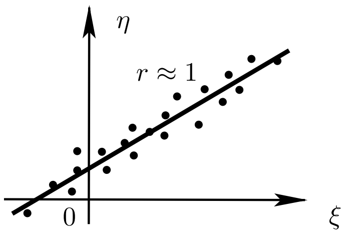
Зураг 12.2: сул, эерэг корреляц



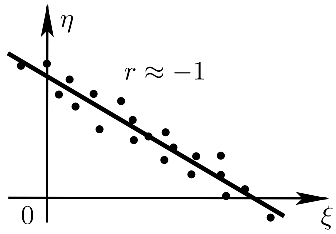
Зураг 12.3: сөрөг корреляц

2. Санамсаргүй хэмжигдэхүүн бүрийн бүх утгуудыг нэг ижил тоогоор юмуу эсвэл нэг ижил тоо дахин ихэсгэж багасгахад корреляцийн коэффициент өөрчлөгдөхгүй.

3. $r = \mp 1$ үед санамсаргүй хэмжигдэхүүний корреляц хамаарал шугаман хамаарлыг илэрхийлнэ. Өөрөөр хэлбэл, санамсаргүй хэмжигдэхүүнүүд корреляц хамааралгүй, тэдгээрийн хоорондын хамаарал ердийн шугаман функцээр илэрхийлэгдэнэ. Энэ тохиолдолд ξ, η санамсаргүй хэмжигдэхүүний бараг бүх ажиглагдсан утгууд регрессийн муруй дээр оршино. (Зураг 12.4, Зураг 12.5)

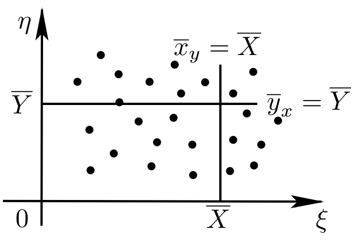


Зураг 12.4:



Зураг 12.5:

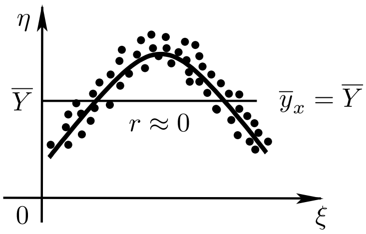
4. $r = 0$ бол санамсаргүй хэмжигдэхүүнүүд корреляц хамааралгүй байна. Энэ тохиолдолд хувьсагчдын бүлгийн дунджууд ерөнхий дунджуудтайгаа тэнцүү болох ба регрессийн шугамууд координатын тэнхлэгүүдтэй параллель болно.



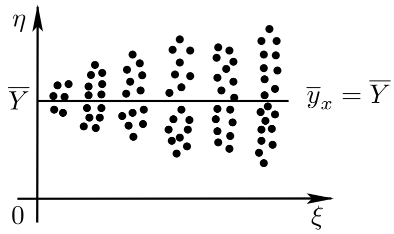
Зураг 12.6:

Энд сануулан хэлэхэд, корреляц хамааралгүй гэдгийг санамсаргүй хэмжигдэхүүний үл хамаарах чанартай хольж ойлгож болохгүй. Хамааралгүй санамсаргүй хэмжигдэхүүнүүд ямагт корреляц хамааралгүй байх боловч урвуу өгүүлбэр нь үл биелэгдэнэ.

Тухайлбал, корреляц хамааралгүй хэмжигдэхүүнүүд хамааралтай, тэгэхдээ бүр функцэн хамааралтай байж болохоос гадна шугаман биш корреляц (статистик) хамааралтай байж болно. Жишээлбэл, Зураг 12.7,12.8-д $r=0$ ба



Зураг 12.7:



Зураг 12.8:

η-ийн ξ дээрх регрессийн шугам абсцисс тэнхлэгтэй параллель боловч корреляцийн орны цэгүүдийн байрлал ба санамсаргүй хэмжигдэхүүний хооронд шугаманаас өөр корреляц хамаарал ажиглагдаж буйг харуулж байна.
Зураг 12.7 шугаман биш корреляц (бараг функцэн) хамаарлыг дүрсэлж байгаа бол, Зураг 12.8 ξ өөрчлөгдөхөд бүлгийн дунджууд нь өөрчлөгдөхгүй боловч корреляцийн орны цэгүүд, регрессийн шугамаас сарних сарнил нь өөрчлөгдөх статистик хамаарлыг илэрхийлж байна. Түүврийн корреляцийн коэффициент r нь эх олонлогийн корреляцийн коэффициент ρ-ийн үнэлэлт учраас түүврийн хэмжээ n хичнээн их байх тутам төдийчинээн нарийн үнэлэлт болно. Дээр тэмдэглэсэн чанарууд нь үнэн чанартаа ρ-коэффициентийн чанарууд боловч n-ийн хүрэлцээтэй их утганд r коэффициентод дэлгэрүүлсэн хэрэг юм.

12.3 Түүврийн хос корреляцийн коэффициент. Корреляцийн таблиц.

Түүврийн корреляцийн коэффициентийг (12.10) томъёогоор бодох боловч практикт хэрэглэхэд тохиромжтойг бодож түүний хувиргасан хэлбэрүүдийг ашигладаг. Гэхдээ анхны мэдээллүүд ямар хэлбэртэй өгөгдсөнөөс хамаарч аль хэлбэрийг ашиглахаа сонгоно. Хэрэв түүврийн хэмжээ бага (ажиглалтын тоо цөөн) бол дараах томъёог ашиглах нь тохиромжтой.

$$r = \frac{\overline{XY} - \overline{X} \cdot \overline{Y}}{S_\xi S_\eta} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - (\sum x_i)^2} \cdot \sqrt{n\sum y_i^2 - (\sum y_i)^2}} \tag{12.14}$$

Жишээ-12.1 (ξ, η) санамсаргүй хэмжигдэхүүний системийн утгууд үл хамаарах 10 удаагийн туршилтанд дараах байдлаар илэрчээ.

(2.1; 3.0) (2.1; 2.8) (2.0; 3.0) (2.5; 2.0) (2.8; 1.8)
(2.2; 2.5) (3.2; 1.5) (3.2; 1.1) (3.2; 1.0) (4.7;1.3).

Түүврийн корреляцийн коэффициентийг ол.

(12.14) томъёог ашиглан корреляцийн коэффициентийг бодоход дараах хүснэгтийг зохиож тооцоо хийх нь тохиромжтой байдаг.

Туршилтын дугаар i	x_i	y_i	x_iy_i	x_i^2	y_i^2
1	2.1	3.0	6.30	4.41	9.00
2	2.1	2.8	5.88	4.41	7.84
3	2.0	3.0	6.0	4.00	9.00
4	2.5	2.0	5.0	6.25	4.00
5	2.8	1.8	5.04	7.84	3.24
6	2.2	2.5	5.50	4.84	6.25
7	3.2	1.5	4.80	10.24	2.25
8	3.2	1.1	3.52	10.24	1.21
9	3.2	1.0	3.20	10.24	1.00
10	4.7	1.3	6.11	22.09	1.69
$\sum$	28	20	51.35	84.56	45.48

Хүснэгтийн сүүлчийн мөрнөөс үзвэл
$\sum x_i = 28$, $\sum y_i = 20$, $\sum x_iy_i = 51.35$, $\sum x_i^2 = 84.56$, $\sum y_i^2 = 45.48$ тул (12.14) томъёо ёсоор $r = \dfrac{-46.5}{7.84 \cdot 7.4} = -0.8$ гэж олдоно.

Түүврийн хэмжээ их үед өгөгдлүүдийг эрэмбэлж, бүлэглэдэг. Бүлэглэсэн түүвэр дискрет ба интервалын байж болно. Хэрэв нийт түүвэрт x_i утга n_i удаа $(i = \overline{1,k})$, y_j утга n_j удаа $(j = \overline{1,s})$, (x_i, y_i) хос утга n_{ij} удаа ажиглагдсан гэвэл санамсаргүй хэмжигдэхүүний утгуудын тухай мэдээллийг дараах корреляцийн таблиц хэлбэрээр дүрсэлж болно. (таблиц 12.1)

$X \backslash Y$	y_1	y_2	$\cdots$	y_j	$\cdots$	y_s	$n_i = \sum\limits_{j=1}^{s} n_{ij}$
x_1	n_{11}	n_{12}		n_{1j}		n_{1s}	n_1
x_2	n_{21}	n_{22}		n_{2j}		n_{2s}	n_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_i	n_{i1}	n_{i2}		n_{ij}		n_{is}	n_i
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_k	n_{k1}	n_{k2}		n_{kj}		n_{ks}	n_k
$n_j = \sum\limits_{i=1}^{k} n_{ij}$	n_1	n_2	$\cdots$	n_j	$\cdots$	n_s	$n = \sum\limits_{i=1}^{k}\sum\limits_{j=1}^{s} n_{ij}$

Үүнд, $n = \sum\limits_{i=1}^{k} n_i = \sum\limits_{j=1}^{s} n_j = \sum\limits_{i=1}^{k}\sum\limits_{j=1}^{s} n_{ij}$.

Интервалын вариацын цувааны тохиолдолд x_i, y_j утгуудын оронд харгалзах интервалын дунджуудыг авна. Санамсаргүй хэмжигдэхүүний утгуудын тухай мэдээлэл корреляцийн таблиц хэлбэрээр өгөгдсөн бол корреляцийн коэффициентийг дараах томъёогоор олно.

$$r = \frac{n\sum\limits_{i=1}^{k}\sum\limits_{j=1}^{s} x_i y_j n_{ij} - \left(\sum\limits_{i=1}^{k} x_i n_i\right)\cdot\left(\sum\limits_{j=1}^{s} y_j n_j\right)}{\sqrt{n\sum\limits_{i=1}^{k} x_i^2 n_i - \left(\sum\limits_{i=1}^{k} x_i n_i\right)^2}\cdot\sqrt{n\sum\limits_{j=1}^{s} y_j^2 n_j - \left(\sum\limits_{j=1}^{s} y_j n_j\right)^2}} \tag{12.15}$$

Үлгэр болгон эдийн засгийн агуулгатай жишээ авч үзье. Энэхүү жишээг дараагийн дэд бүлгүүдэд үргэлжлүүлэн авч үзэх болно.

Жишээ-12.2 Ижил төрлийн 50 үйлдвэрийн үйлдвэрлэлийн үндсэн фонд (сая төг)-X, хоногт үйлдвэрлэх бүтээгдэхүүний хэмжээ (тонн)-Y дараах хүснэгтэд өгөгдөв. X ба Y-хэмжигдэхүүний статистик хамаарлыг судла. (таблиц 12.2)

Үндсэн фондын хэмжээ X	Интер-валуудын дундаж	Хоногт үйлдвэрлэх бүтээгдэхүүний хэмжээ					Нийт n_i	Бүлгийн дунджууд $\overline{y}_j$
		7-11	11-15	15-19	19-23	23-27		
	y_j / x_i	9	13	17	21	25		
20-25	22.5	2	1	-	-	-	3	10.3
25-30	27.5	3	6	4	-	-	13	13.3
30-35	32.5	-	3	11	7	-	21	17.8
35-40	37.5	-	1	2	6	2	11	20.3
40-45	42.5	-	-	-	1	1	2	23.0
Нийт n_j		5	11	17	14	3	50	—
Бүлгийн дунджууд $\overline{x}_j$		25.5	29.3	31.9	35.4	39.2	—	—

Бүлгийн дунджууд гэсэн мөр баганыг эс тооцвол дээрх өгөгдөл нь корреляцийн таблиц хэлбэрээр өгөгдсөн мэдээлэл юм. Бүлгийн дундаж тус бүрээр туршилтын муруй байгуулбал, X, Y хэмжигдэхүүнүүд шугаман корреляц хамааралтай байх боломжтой нь харагдаж байна. Одоо энэ хоёр хэмжигдэхүүний корреляц хамаарлын коэффициентийг ольё.

$$\sum_{i=1}^{5} x_i n_i = 22.5\cdot 3 + 27.5\cdot 13 + 32.5\cdot 21 + 37.5\cdot 11 + 42.5\cdot 2 = 1605$$

$$\sum_{i=1}^{5} x_i^2 n_i = 22.53^2\cdot 3 + 27.5^2\cdot 13 + 32.5^2\cdot 21 + 37.5^2\cdot 11 + 42.5^2\cdot 2 = 52612.5$$

$$\sum_{j=1}^{5} y_i n_j = 9\cdot 5 + 13\cdot 11 + 17\cdot 17 + 21\cdot 14 + 25\cdot 3 = 846$$

$$\sum_{j=1}^{5} y_i^2 n_j = 9^2\cdot 5 + 13^2\cdot 11 + 17^2\cdot 17 + 21^2\cdot 14 + 25^2\cdot 3 = 15226$$

$$\sum_{i=1}^{5}\sum_{j=1}^{5} x_i y_i n_{ij} = 22.5\cdot 9\cdot 2 + 22.5\cdot 13\cdot 1 + \cdots + 42.5\cdot 25\cdot 1 = 27895.$$

(12.15) томъёогоор түүврийн корреляцийн коэффициентийг олбол $r = 0.74$ болно. Корреляцийн коэффициент 1-д ойр гарсан тул үйлдвэрлэлийн үндсэн фонд ба хоногт үйлдвэрлэх бүтээгдэхүүний хэмжээ хоорондоо эерэг, шугаман корреляц хамааралтай төдийгүй энэхүү корреляц хамаарал хүчтэйд тооцогдоно.

12.4 Корреляцийн коэффициентийн итгэлтэй эсэхийг шалгах

Практикт, санамсаргүй хэмжигдэхүүн хоорндын онолын корреляцийн коэффициент үл мэдэгдэх тул тэдгээрийн хамаарлын хүчийг тогтоохдоо корреляцийн коэффициентийн цэгэн үнэлэлт болох r-ийг ашиглана. Эх олонлогоос авсан санамсаргүй түүврийн утгуудаар r-ийг олж буй учраас түүний ρ коэффициентоос ялгагдах үндсэн ялгаа нь r өөрөө санамсаргүй хэмжигдэхүүн байдагт оршино. $r = 0$ гарах нь бидний сонирхож буй коэффициент $\rho = 0$ байхыг нотолж буй хэрэг биш юм. Улмаар ξ ба η санамсаргүй хэмжигдэхүүнүүд корреляц хамааралгүй гэдгийг ч нотлохгүй болохыг өмнөх зүйлд тэмдэглэсэн билээ.

Түүнчлэн $r \neq 0$ гарах нь үнэхээр ξ, η санамсаргүй хэмжигдэхүүнүүдийн хооронд шугаман корреляц хамаарал оршин байна уу? үгүй юу? гэдгийг нэг утгатай шийдэхгүй. Ийм учраас санамсаргүй хэмжигдэхүүнүүд корреляц хамааралтай эсэхийг тайлбарлахын тулд түүврийн корреляцийн коэффициент r-ийн итгэлтэй чанарыг шалгах хэрэгтэй. Өөрөөр хэлбэл, санамсаргүй хэмжигдэхүүнүүд корреляц хамааралтай байх тухай үндэслэлтэй дүгнэлт гаргахад, r-коэффициентийн утга хангалттай эсэхийг тогтоох хэрэгтэй. Үүний тулд эх олонлогоос авсан санамсаргүй хэмжигдэхүүнүүд шугаман корреляц хамааралгүй байх тухай $H_0 : \rho = 0$ гэсэн таамаглалыг шалгана. Энэ таамаглалыг шалгахын тулд дараах шинжүүрийг авч үзнэ.

$$t = \frac{r}{S_r} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = r\sqrt{(n-2)/(1-r^2)} \tag{12.16}$$

(12.16) шинжүүр нь $k = n-2$ чөлөөний зэрэг бүхий Стьюдентийн (t-тархалт) тархалттай байна. Өгөгдсөн итгэх түвшин α-ийн хувьд Стьюдентийн таблицыг ашиглан $P(|t| \geq t_{\alpha,k}) = \alpha$ нөхцлийг хангах онолын утга $t_{\alpha,k}$-г олно.

Хэрэв $|t| \geq t_{\alpha,k}$ нөхцөл биелэгдэж байвал H_0 таамаглалыг хэрэгсэхгүй. Өөрөөр хэлбэл, түүврийн корреляцийн коэффициент тэгээс мэдэгдэхүйц ялгагдах учраас санамсаргүй хэмжигдэхүүнүүд корреляц хамааралтай гэсэн дүгнэлт хийж болно гэсэн үг. Энэ тохиолдолд түүврийн корреляцийн коэффициент r-ийг итгэлтэй коэффициент гэх бөгөөд итгэлтэй чанартай байна гэж ярьдаг. $|t| < t_{\alpha,k}$ үед H_0 таамаглалыг үгүйсгэх үндэслэлгүй тул хүлээн зөвшөөрнө.

Түүврийн корреляцийн коэффициент итгэлтэй тохиолдолд эх олонлогийн корреляцийн коэффициентийн итгэх завсрыг байгуулж болно. Үүний тулд түүврийн корреляцийн коэффициентийн тархалтын хуулийг мэдэх шаардлагатай бөгөөд түүний магадлалын нягт түвэгтэй хэлбэртэй байх ба $\rho \neq 0$ үед тэгш хэмгүй байхын гадна n-ийг өсөхөд хэвийн тархалт уруу маш аажуу нийлдэг.
Ийм учраас r коэффициентийн тархалтын хуулийн оронд, сайн судлагдсан тархалтууд руу нийлдэг тусгайлан сонгон авсан функцүүдийг ашигладаг. Тухайлбал, Фишерийн z хувиргалтыг сонгон авдаг.

$$z = \frac{1}{2}\ln\Big(\frac{1+r}{1-r}\Big) \tag{12.17}$$

Фишерийн z хувиргалт нь:

$$M(z) = \frac{1}{2}\ln\Big(\frac{1+\rho}{1-\rho}\Big) + \frac{\rho}{2(n-1)} \tag{12.18}$$

гэсэн математик дундажтай,

$$\sigma_z^2 = \frac{3}{n-3} \tag{12.19}$$

гэсэн дисперстэй хэвийн тархалт уруу нийлнэ.
Тэгвэл эх олонлогийн ρ коэффициентийн итгэх завсар
$th(z_1) < \rho < th(z_2)$ хэлбэртэй байна. $th(z_1)$, $th(z_2)$ утгуудыг таблицаас (Таблиц N9) z_1, z_2-ын дараах холбогдлуудад олно.

$$\begin{aligned} z_1 &= \frac{1}{2}\ln\Big(\frac{1+r}{1-r}\Big) - t_\alpha\frac{1}{\sqrt{n-3}} \\ z_2 &= \frac{1}{2}\ln\Big(\frac{1+r}{1-r}\Big) + t_\alpha\frac{1}{\sqrt{n-3}} \end{aligned} \tag{12.20}$$

Үүнд, $\Phi(t_\alpha) = 1-\alpha$, $\Phi(x)$-Лапласын функц.
Түүврийн корреляцийн коэффициент итгэлтэй биш тохиолдолд эх олонлогийн ρ параметрийн итгэх завсрыг байгуулах тухай асуудлыг авч үзэхгүй.

Жишээ-12.3 Өмнөх дэд бүлэгт авч үзсэн жишээний $X, Y-$санамсаргүй хэмжигдэхүүн хоорондын, түүврийн корреляцийн коэффициент итгэлтэй эсэхийг шалгаж, онолын коэффициент $\rho-$ийн итгэх завсрыг байгуул. Итгэх түвшин $\alpha = 0,05$.

($r = 0.74$ гэж олсон билээ.) (12.16) ёсоор Стьюдентийн шинжүүрийг бичвэл

$$t = \frac{0.74\sqrt{50-2}}{\sqrt{1-0.74^2}} = 7.62$$

Стьюдентийн таблицаас $\alpha = 0.05$, $k = 50-2 = 48$ үед онолын утгыг олбол $t_{0.05;48} = 2.01$

Туршилтын утга ба онолын утгын хувьд $t > t_{0.05;48}$ нөхцөл биелэгдэж буй тул r коэффициент итгэлтэй коэффициент болно. Одоо ρ коэффициентийн итгэх завсрыг байгуулья. Үүний тулд (12.17) томъёогоор Фишерийн z хувиргалтын утгыг бодъё.

$$z = \frac{1}{2}\ln\frac{1+0.74}{1-0.74} = 0.9505, \qquad \Phi(t_\alpha) = 0.95$$

нөхцлөөс $t_{0.05} = 1.96$ тул (12.20) томъёо ёсоор

$$z_1 = 0.9505 - 1.96 \cdot \frac{1}{\sqrt{50-3}} = 0.6646,$$
$$z_2 = 0.9505 - 1.96 \cdot \frac{1}{\sqrt{50-3}} = 1.2364$$

Гипербол тангенс $th\ (z)$ функцийн таблиц юмуу эсвэл $th(z) = \dfrac{e^{2z}-1}{e^{2z}+1}$ томъёог ашиглавал эх олонлогийн корреляцийн коэффициентийн итгэх завсар $0.51 \leq \rho \leq 0.844$ болно.

12.5 Корреляцийн харьцаа ба индекс

Бидний дээр авч үзсэн корреляцийн коэффициент нь зөвхөн хамтын хэвийн тархалттай бөгөөд санамсаргүй хэмжигдэхүүнүүд хоорондоо шугаман корреляц хамааралтай тохиолдолд л хамаарлын хүчийг илэрхийлэх үзүүлэлт болж чадна. Гэтэл практикт судалж буй хэмжигдэхүүнүүд өөр хоорондоо статистик хамааралтай боловч аль нэг нь санамсаргүй биш ч юмуу эсвэл хамтдаа хэвийн тархалттай биш гэх мэтчилэн корреляцийн шинжилгээний урьдчилсан нөхцлийг хангаагүй байж болдог. Энэ тохиолдолд хэмжигдэхүүнүүдийн харилцан хамаарлыг илэрхийлэх ерөнхий үзүүлэлт нь корреляцийн харьцаа юм.

Нэг санамсаргүй хэмжигдэхүүний утга нь нөгөөгөөсөө, функцийн утга байдлаар өөрчлөгдөх биш зөвхөн өөр хоорондын уялдаа байдлаар илрэх хамаарлыг харилцан хамаарал гэж ойлгоно. Энэ үзүүлэлтийг гаргахын тулд η санамсаргүй хэмжигдэхүүний түүврийн ерөнхий дисперсийг авч үзье.

$$S_\eta^2 = \frac{1}{n}\sum_{j=1}^{s}(y_j - \overline{Y})^2 n_j \tag{12.21}$$

Хэрэв η санамсаргүй хэмжигдэхүүний туршилтын утгууд хоорондоо үл огтлолцох хэд хэдэн бүлгээс тогтож байвал түүний ерөнхий дисперсийг дараах хоёр нийлбэрт тавьж болно.

$$S_\eta^2 = M\Big(M(\eta|\xi = x) - M\eta\Big)^2 + M\Big(\eta - M(\eta|\xi = x)\Big)^2$$

товчоор

$$S_{\eta}^2 = \overline{\delta}_{\eta|\xi}^2 + \overline{\sigma}_{\eta|\xi}^2 \tag{12.22}$$

$\overline{\delta}_{\eta|\xi}^2$-нь бүлэг хоорондын дисперс бөгөөд ξ хүчин зүйл η-д нөлөөлөх нөлөөллийг илэрхийлнэ.

$$\overline{\delta}_{\eta|\xi}^2 = \frac{1}{n}\sum_{i=1}^{k}(\overline{y}_i - \overline{Y})^2 n_i \tag{12.23}$$

$\overline{\sigma}_{\eta|\xi}^2$-үлдэгдэл дисперс. Үлдэгдэл дисперс нь бүлгийн дисперсүүдийн дундаж бөгөөд тооцогдоогүй бусад хүчин зүйлүүдийн нөлөөллийг илэрхийлнэ.

$$\begin{aligned} \overline{\sigma}_{\eta|\xi}^2 &= \frac{1}{n}\sum_{i=1}^{k}\overline{S}_{i\eta}^2 n_i \\ \overline{S}_{i\eta}^2 &= \frac{1}{n_i}\sum_{j=1}^{s}(y_j - \overline{y}_i)^2 n_{ij}, \quad (i = \overline{1,k}) \end{aligned} \tag{12.24}$$

$$\tau_{\eta\xi} = \sqrt{\frac{\overline{\delta}_{\eta|\xi}^2}{S_{\eta}^2}} \tag{12.25}$$

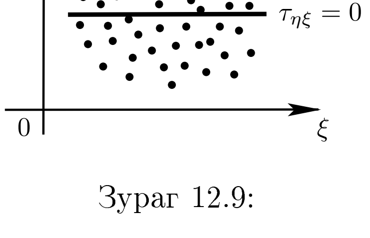
хэмжигдэхүүнийг туршилтын буюу түүврийн корреляцийн харьцаа (η-ийн ξ дээрх) гэж нэрлээд, ξ ба η хэмжигдэхүүний харилцан хамаарлын нягтралыг илэрхийлэх үзүүлэлт болгон авдаг.
$\tau_{\eta\xi}$ хэмжигдэхүүний оронд $\tau_{\eta\xi}^2$ хэмжигдэхүүнийг хэрэглэж заншсан бөгөөд түүнийг детерминацийн коэффициент гэнэ.
Санамсаргүй хэмжигдэхүүнүүдийн харилцан хамаарал хичнээн хүчтэй байх тутам $\tau_{\eta\xi}$ харьцаа их байна. Энэ нь ξ санамсаргүй хэмжигдэхүүний өөрчлөлт, тооцогдоогүй үлдсэн санамсаргүй хүчин зүйлүүдийг бодвол илүү ихээр η санамсаргүй хэмжигдэхүүний өөрчлөлтөнд нөлөөлөхийг илэрхийлнэ.
Үүнчлэн,

$$\tau_{\xi\eta} = \sqrt{\overline{\delta}_{\xi|\eta}^2 / S_{\xi}^2} \tag{12.26}$$

гэсэн корреляцийн харьцааг авч үзнэ. Корреляцийн $\tau_{\eta\xi}, \tau_{\xi\eta}$ харьцааны чанаруудыг дурдвал:

1. Корреляц харьцаа сөрөг биш байна. $(0 \leq \tau \leq 1)$

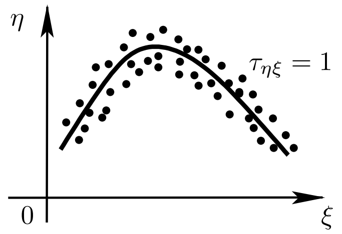
2. $\tau_{\eta\xi} = 0$ байх зайлшгүй бөгөөд хүрэлцээтэй нөхцөл нь $\overline{y}_i = \overline{Y} = const$ байх явдал юм.(Зураг 12.9)



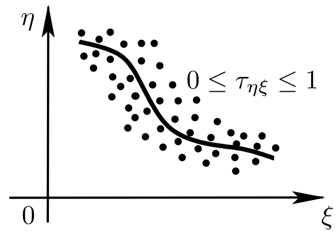
Зураг 12.9:

Өөрөөр хэлбэл, η-ийн ξ дээрх регрессийн шугам тархалтын төвийг дайрсан хэвтээ шугам байна гэсэн үг. Энэ тохиолдолд ξ, η хэмжигдэхүүнүүд корреляц хамааралгүй байна. $\tau_{\eta\xi}$-коэффициентийн хувьд мөн ийм чанар биелнэ.

3. $\tau_{\eta\xi} = 1$ байх зайлшгүй бөгөөд хүрэлцээтэй нөхцөл нь $\overline{\sigma}^2_{\eta|\xi} = 0$ байх явдал юм. Өөрөөр хэлбэл, санамсаргүй хүчин зүйлүүдийн нөлөө байхгүй бөгөөд бүх утгууд нь регрессийн муруйн орчимд төвлөрсөн байна гэсэн үг.



Зураг 12.10:



Зураг 12.11:

4. $\tau_{\eta\xi} \neq \tau_{\xi\eta}$ байна. Энэ нь корреляцийн коэффициентоос ялгагдах ялгаа юм. (корреляцийн коэффициентийн хувьд $r_{\xi\eta} = r_{\eta\xi} = r$.) Өөрөөр хэлбэл, корреляцийн харьцааг бодоход ямар хувьсагчийг үл хамаарах хувьсагч, ямрыг нь хамааран хувьсагч болгон сонгон авахаас хамаарна гэсэн үг.
Туршилтын корреляцийн харьцаа $\tau_{\eta\xi}$ нь $\overline{y}_i$ утгуудыг холбосон регрессийн шугамаас (хугархай шугам байна) корреляцийн орны цэгүүдийн сарних сарнилыг үзүүлсэн үзүүлэлт юм. Гэвч санамсаргүй хүчин зүйлийн нөлөөллөөс болж $\overline{y}_i$ утгуудын өөрчлөлтийн хууль нь зөрчигдөх учраас $\tau_{\eta\xi}$ харьцаа ихэсдэг. Үүнтэй уялдан онолын корреляцийн харьцаа буюу η-ийн ξ дээрх корреляцийн индекс $R_{\eta\xi}$ гэсэн үзүүлэлтийг авч үздэг.
Корреляцийн индекс $R_{\eta\xi}$ нь корреляцийн орны цэгүүд жинхэнэ регрессийн шугам $\overline{y}_x$-ээс сарних сарнилыг тодорхойлно:

$$R_{\eta\xi} = \sqrt{\frac{\delta^2_{\eta|\xi}}{S^2_\eta}} = \sqrt{1 - \frac{\sigma^2_{\eta|\xi}}{S^2_\eta}} \tag{12.27}$$

Үүнд, $\delta^2_{\eta|\xi}$ ба $\sigma^2_{\eta|\xi}$ дисперсүүд нь (12.23), (12.24) томьёогоор бодогдох боловч бүлгийн дунджууд $\overline{y}_i$-г регрессийн тэгшитгэлээр олсон нөхцөлт математик дундаж $\overline{y}_x$-ээр солино. Үүнтэй төстэйгээр

$$R_{\xi\eta} = \sqrt{\frac{\delta^2_{\xi|\eta}}{S^2_\xi}} = \sqrt{1 - \frac{\sigma^2_{\xi|\eta}}{S^2_\xi}} \tag{12.28}$$

гэсэн корреляцийн индексийг авч үзнэ.
τ, R үзүүлэлтүүдийн давуу тал нь ξ, η хэмжигдэхүүнүүд дурын харилцан хамааралтай байхад тэдгээрийг олж болдогт оршино. τ үзүүлэлтийн утга R-тэй харьцангуйгаар өндөр гарах боловч түүнийг олоход регрессийн тэгшитгэлийг мэдэх шаардлагагүй учраас давуу талтай юм.
r ба τ, R үзүүлэлтүүд дараах харьцаагаар холбогдоно.

$$0 \leq |r| \leq R \leq \tau \leq 1 \tag{12.29}$$

Хувьсагчдын харилцан хамаарал шугаман үед $R_{\xi\eta} = R_{\eta\xi} = |r|$ байна. Корреляцийн харцаа τ- ийн итгэлтэй эсэхийг шалгахад дараах шинжүүрийг ашиглана.

$$F_\tau = \frac{\tau^2(n-m)}{(1-\tau^2)(m-1)} \tag{12.30}$$

Үүнд: m-санамсаргүй хэмжигдэхүүний утгуудыг бүлэглэсэн интервалын тоо. F_τ нь $k_1 = m-1$ ба $k_2 = n-m$ чөлөөний зэргүүд бүхий Фишерийн F-тархалттай байна. Хэрэв өгөгдсөн итгэх түвшин α-ийн хувьд: $F > F_{\alpha,k_1,k_2}$ нөхцөл биелэгдэж байвал τ нь тэгээс мэдэгдэхүйцээр ялгагдана. Өөрөөр хэлбэл, τ-г түүврийн корреляцийн харьцааг илэрхийлэх итгэлтэй коэффициент гэж үзнэ. F_{α,k_1,k_2}-таблицаас олсон онолын утга.
Хоёр хувьсагчийн корреляцийн индекс R-ийн итгэлтэй чанарыг дараах шинжүүрийг ашиглан шалгана.

$$F_R = \frac{R^2(n-2)}{1-R^2} \tag{12.31}$$

F_R-хэмжигдэхүүн $k_1 = 1$, $k_2 = n-2$ чөлөөний зэргүүд бүхий Фишерийн F-тархалттай байна. Мөн $F_R > F_{\alpha,k_1,k_2}$ нөхцөл биелэгдвэл R-ийг итгэлтэй гэж тооцно.

Жишээ-12.4 12.3.-ын таблиц 12.2-ын өгөгдлөөр корреляцийн харьцаа $\tau_{\eta,\xi}$ ба корреляцийн индекс $R_{\eta\xi}$-г олж тэдгээрийн итгэлтэй эсэхийг шалга.

Эхлээд Жишээ-12.2-д олсон үр дүнгүүдээр $\overline{Y}, S^2_\eta$-г тус тус олъё.

$$\overline{Y} = \frac{846}{50} = 16.92(\text{т}), \quad S^2_\eta = \frac{15226}{50} - 16.92^2 \approx 18,23.$$

Бүлгийн дундаж ба интервалын давтамжуудыг оролцуулан тооцоог хийхдээ дараах хүснэгтийг ашиглах нь тохиромжтой.

x_i	n_i	$\overline{y}_i$	$(\overline{y}_i - \overline{Y})^2 n_i$	$\overline{y}_x$	$(\overline{y}_x - \overline{Y})^2 n_i$
22.5	3	10.3	131.5	10.4	127.5
27.5	13	13.3	170.4	13.8	126.5
32.5	21	17.8	16.3	17.2	1.6
37.5	11	20.3	125.7	20.6	149.0
42.5	2	23.0	73.9	23.9	97.4
$\sum$	-	-	517.8	-	502.0

(12.23) томьёо ёсоор: $\overline{\delta}^2_\eta = \dfrac{517.8}{50} = 10.36$ тул (12.25) томьёог ашиглавал

$$\tau_{\eta\xi} = \sqrt{\frac{10.36}{18.23}} = 0.754$$

Корреляцийн харьцаа $\tau_{\eta\xi}$-ийн утга нь өмнөх зүйлд олсон $r = 0.74$-тэй ойролцоо гарч байна. Энэ нь санамсаргүй хэмжигдэхүүнүүд шугаман корреляц

хамааралтай байх тухай дэвшүүлсэн таамаглалыг нотолж буй юм.
$R_{\eta\xi}$-индексийг олохын тулд: $\overline{y}_x = 0.6762x - 4.79$ (Энэхүү регрессийн тэгшитгэлийг 13.2-д гаргаж үзүүлнэ) гэсэн регрессийн тэгшитгэлээр $\overline{y}_{x_i}$ утгуудыг олно. Эдгээр утгуудыг өмнөх хүснэгтийн сүүлээсээ 2 дахь баганад байрлуулсан.

$$\delta^2_{\eta|\xi} = \frac{502.0}{50} = 10.04, \qquad R_{\eta\xi} = \sqrt{\frac{10.04}{18.23}} = 0.742$$

Завсрын үр дүнгүүдийг бодоход гарсан бүдүүвчлэлийн алдааг эс тооцвол $R_{\eta\xi}$ ба r нь бараг ижил гарч байна. Иймд хувьсагчууд шугаман корреляц хамааралтай үед $R_{\eta\xi}$ индексийг бодох шаардлагагүй, зөвхөн r коэффициентийг бодоход хангалттай болох нь харагдаж байна. Детерминацийн коэффициент $R^2_{\eta\xi} = 0.551$ байгаа нь, хоногт үйлдвэрлэх бүтээгдэхүүн (Y)-ийн өөрчлөлтийн 55.1% нь үйлдвэрлэлийн үндсэн фонд (X)-ийн өөрчлөлтөөр тайлбарлагдахыг харуулж байна. Бүлэглэсэн интервалын тоо $m = 5$ болохыг анхааран $\tau_{\eta\xi}$ коэффициентийн итгэлтэй эсэхийг шалгая.

$$F_\tau = \frac{0.754^2(50-5)}{(1-0.754)^2(5-1)} = 14.82$$

Итгэх түвшин $\alpha = 0.05$ ба $k_1 = 4,\ \ k_2 = 45$ үед фишерийн F-тархалтын таблицаас $F_{0.05;4;45} = 2.57$ гэж олдоно.
$F_\tau > F_{0.05;4;45}$ учир $\tau_{\eta\xi}$-г итгэлтэй гэж тооцно. Үүнчлэн $R_{\eta\xi}$ коэффициентийг шалгая.

$$F_R = \frac{0.742^2(50-2)}{(1-0.742)^2} = 58.8, \qquad F_R > F_{0.05,1,48} = 4.04$$

тул корреляцийн индекс $R_{\eta\xi} = 0.742$ нь итгэлтэйд тооцогдоно.

12.6 Олон хэмжээст корреляцийн шинжилгээ. Корреляцийн ерөнхий ба тухайн коэффициент

Амьдрал практикийн үзэгдлүүд ихэвчлэн олон хүчин зүйлт загвараар дүрслэгддэг. Ийм учраас өмнө авч үзсэн хоёр хэмжээст корреляцийн загварыг олон хэмжээст тохиолдолд өргөтгөх шаардлагатай. Олон хэмжээст корреляцийн шинжилгээний зорилго нь $X_1, X_2, \ldots, X_n$ гэсэн $(n > 2)$ хувьсах хэмжигдэхүүн хоорондын харилцан хамаарлыг судалж, эдгээрээс хамаарсан (Y)-хэмжигдэхүүнд хамгийн ихээр нөлөөлөх X_i хувьсагчдыг сонгон авахад оршино.
$X_1, X_2, \ldots, X_j, \ldots, X_m$-гэсэн олон хэмжээст хэвийн тархалттай санамсаргүй хэмжигдэхүүнүүд өгөгдсөн байг. Эдгээр санамсаргүй хэмжигдэхүүний харилцан хамаарлыг хос корреляцийн коэффициентүүдээр зохиогдсон, дараах

корреляцийн матрицаар илэрхийлж болно.

$$Q_m = \begin{pmatrix} 1 & \rho_{12} & \dots & \rho_{1m} \\ \rho_{21} & 1 & \dots & \rho_{2m} \\ \dots & \dots & \dots & \dots \\ \rho_{m1} & \rho_{m2} & \dots & 1 \end{pmatrix} \tag{12.32}$$

ρ_{jk} нь (12.6) томьёогоор тодорхойлогдох хос корреляцийн коэффициентүүд, $(j,k = 1,2,\dots,m)$ m-матрицын эрэмбэ.
Корреляцийн матриц Q_m нь тэгш хэмтэй бөгөөд ямар санамсаргүй хэмжигдэхүүнүүд хоорондоо корреляц хамааралтай болохыг үзүүлнэ. Корреляцийн матриц Q_m-ийг түүврийн утгуудаар үнэлэх нь олон хэмжээст корреляцийн шинжилгээний үндсэн бодлого юм. Энэ нь түүврийн корреляцийн коэффициентийн матриц байгуулснаар шийдэгдэнэ:

$$r_m = \begin{pmatrix} 1 & r_{12} & \dots & r_{1m} \\ r_{21} & 1 & \dots & r_{2m} \\ \dots & \dots & \dots & \dots \\ r_{m1} & r_{m2} & \dots & 1 \end{pmatrix} \tag{12.33}$$

Үүнд:

$$r_{jk} = \frac{\sum\limits_{i=1}^{n}(x_{ij} - \overline{X}_j)(x_{ik} - \overline{X}_k)}{nS_{x_j}S_{x_k}} \quad (j,k = 1,2,\dots,m) \tag{12.34}$$

Хоёр хэмжээст загварын адилаар, корреляцийн коэффициентийг үнэлэхийн тулд түүний математик дундаж ба дисперсийн үнэлэлтийг байгуулах шаардлагатай. Гэтэл олон хэмжээст корреляцийн шинжилгээнд, математик дундаж ба дисперсийн тус бүр m үнэлэлт, $\dfrac{m(m-1)}{2}$ тооны хос корреляцийн коэффициент , нийтдээ $2m + \dfrac{m(m-1)}{2}$ тоон параметрийн үнэлэлт байгуулах шаардлага гарч байна. Нөгөө талаас олон хэмжээст тохиолдолд хэмжигдэхүүн хоорондын хамаарал нийлмэл бөгөөд олон талтай учраас зөвхөн корреляцийн матрицын тусламжтайгаар тэдгээр хувьсагчдын харилцан хамаарлыг бүрэн дүүрэн илэрхийлж эс чадна. Ийм учраас олон хэмжээст корреляцийн шинжилгээнд дараах 2 төрлийн бодлогыг ихэвчлэн авч үздэг.

а). Аль нэг санамсаргүй хэмжигдэхүүн нь бусад үлдсэн $(m-1)$ санамсаргүй хэмжигдэхүүнээс хамаарах харилцан хамаарлыг тогтоох.

б). l тооны санамсаргүй хэмжигдэхүүний нөлөөг үл тооцож $(l \leq m-2)$ үлдсэн санамсаргүй хэмжигдэхүүн хоорондын хамаарлыг судлах.

Эдгээр бодлогууд нь корреляцийн ерөнхий ба тухайн коэффициентуудыг олсноор шийдэгдэнэ. Аль нэг X_i хэмжигдэхүүн бусад үлдсэн X_j санамсаргүй хэмжигдэхүүний ($m-1$ тооны) олонлогоос хамаарах шугаман корреляц хамаарлыг, корреляцийн коэффициент ρ_{ij}-ийн өргөтгөл болох $\rho_{i.12\dots m}$—коэффициентоор хэмжинэ.

$\rho_{i.12...m}$ коэффициентийг олонлог коэффициент буюу ерөнхий коэффициент гэнэ. i-р хувьсагч бусад ямар хувьсагчдын олонлогоос хамаарах үзүүлэлт болохыг тэмдэглэхдээ ρ-ийн индексэд i-ийн дараа цэг тавьж бусад хувьсагчдын дугаарыг бичнэ.

$\rho_{i.12...m}$-коэффициентийн үнэлэлт болох түүврийн, корреляцийн ерөнхий коэффициент $R_{i.12....m}$ дараах томъёогоор тодорхойлогдоно.

$$R_{i.12....m} = \sqrt{1 - \frac{|r_m|}{q_{ii}}} \tag{12.35}$$

Үүнд, $|r_m|$-түүврийн хос коэффициентуудаар зохиогдсон корреляцийн матрицын тодорхойлогч, $q_{ii} - r_{ii}$ элементийн алгебрийн гүйцээлт.

Тухайн тохиолдолд $m = 3$ буюу X_i, X_j, X_k-санамсаргүй хэмжигдэхүүнүүдийн хувьд корреляцийн ерөнхий коэффициент дараах хэлбэртэй болно.

$$R_{i.jk} = \sqrt{\frac{r_{ij}^2 + r_{ik}^2 - 2 \cdot r_{ij} \cdot r_{ik} \cdot r_{jk}}{1 - r_{jk}^2}} \tag{12.36}$$

Үүнд, r_{st} нь харгалзах индекс бүхий хувьсагчдын хос корреляцийн коэффициентууд $(s = 1, 2, 3, \quad t = 1, 2, 3)$.

Корреляцийн ерөнхий коэффициентийн квадрат $R^2_{i.12...m}$-ийг детерминацийн ерөнхий коэффициент гэх бөгөөд судалж буй хувьсагчийн өөрчлөлтийн хэдэн хувийг хүчин зүйлийн хувьсагчдын өөрчлөлт тайлбарлаж чадахыг үзүүлэх үзүүлэлт юм.

$R_{i.12...m}$ ба $R^2_{i.12...}$-коэффициентийн утга [0,1] хэрчимд харьяалагдана. Корреляцийн ерөнхий коэффициентийн итгэлтэй чанарыг шалгахад дараах шинжүүрийг ашиглана.

$$F = \frac{R^2(n-m)}{(1-R^2)(m-1)} \tag{12.37}$$

Энэхүү шинжүүр $k_1 = m-1, \ k_2 = n-m$ чөлөөний зэрэг бүхий Фишерийн F-тархалттай байна. $F > F_{\alpha;k_1;k_2}$ нөхцөл биелэгдэж байвал ерөнхий коэффициентийг тэгээс мэдэгдэхүйц хэмжээгээр ялгагдаж байна гэж үзнэ. Өөрөөр хэлбэл R-коэффициентийг итгэлтэйд тооцно.

Санамсаргүй хэмжигдэхүүнүүд бие биетэйгээ корреляц хамааралтай үед тэдгээрийн хос корреляцад бусад хувьсагчдын корреляц нөлөөлдөг. Үүнтэй уялдан нэг буюу хэд хэдэн хэмжигдэхүүний нөлөөг тооцохгүй (авч үзээгүй) нөхцөлд зарим хувьсагчдын хоорондох тухайн корреляцийг авч үзэх шаардлага гардаг.

$m-2$ санамсаргүй хэмжигдэхүүний нөлөөг тогтмолоор тооцох нөхцөлд X_i ба X_j хэмжигдэхүүний түүврийн, корреляцийн тухайн коэффициент дараах томъёогоор тодорхойлогдоно.

$$r_{ij.12...m} = \frac{q_{ij}}{\sqrt{q_{ii}q_{jj}}} \tag{12.38}$$

Үүнд q_{ij}, $q_{ii}-r_m$ матрицын харгалзах r_{ij}, r_{ii} элементийн алгебрийн гүйцээлтүүд. Тухайн тохиолдолд $m=3$ буюу X_i, X_j, X_k хувьсагчдын хувьд корреляцийн тухайн коэффициентүүд дараах томъёогоор бодогдоно.

$$r_{ij.k}=\frac{r_{ij}-r_{ik}-r_{jk}}{\sqrt{(1-r_{ik}^2)(1-r_{jk}^2)}} \tag{12.39}$$

r_{st}-харгалзах индекс бүхий хувьсагчдын хос корреляцийн коэффициентүүд.

Корреляцийн тухайн коэффициент нь хос корреляцийн коэффициентийн адилаар $-1\leq r_{ij.12\ldots m}\leq 1$ утга авна. Түүнчлэн n хэмжээт түүврийн утгаар олсон дээрх коэффициент нь $n-m+2$ тооны ажиглалтаар олсон түүврийн хос корреляцийн коэффициенттэй адил тархалттай байна. Ийм учраас корреляцийн тухайн коэффициентийн итгэлтэй чанар нь туршилтын тоо $n-m+2$ байх үеийн хос корреляцийн коэффициентийн адилаар үнэлэгдэнэ.

Жишээ-12.5 Хөдөлмөрийн бүтээмж (X_1), ажилчдын насны хэмжээ (X_2), үйлдвэрлэлд ажилласан жилийн (X_3) харилцан хамаарлыг судлахын тулд ижил мэргэжилтэй ажилчдаас $n=100$ хэмжээтэй түүвэр авч түүврийн хос корреляцийн коэффициентүүд $r_{12}=0.20$, $r_{13}=0.41$, $r_{23}=0.82$-ыг олж итгэлтэй коэффициентууд гэж тогтоожээ. Тэгвэл корреляцийн ерөнхий коэффициент $R_{1.23}$ ба тухайн коэффициентүүдийг олж тэдгээрийн итгэлтэй эсэхийг шалга.

(12.36) томъёог ашиглавал:

$$R_{1.23}=\sqrt{\frac{0.28^2+0.41^2-2\cdot 0.20\cdot 0.41\cdot 0.82}{1-0.82^2}}=0.47$$

Үүнээс үзвэл, нэг талаас хөдөлмөрийн бүтээмж, нөгөө талаас ажилчдын насны хэмжээ ба ажилласан жил хоёрын хооронд мэдэгдэхүйц корреляц хамаарал байгаа нь илэрч байна.

Детерминацийн ерөнхий коэффициент $R^2_{1.23}=0.225$ байгаа нь ажилчдын хөдөлмөрийн бүтээмжийн өөрчлөлтийн 22.5% нь тэдний нас ба ажилласан жилтэй холбоотойгоор тайлбарлагдана гэсэн үг. Одоо $R_{1.23}$-ийн итгэлтэй эсэхийг шалгая.

$$F=\frac{0.47^2(100-3)}{(1-0.47^2)(3-1)}=14.1$$

Таблицаас $F_{0.05;2;97}=3.09$ гэж олдох ба $F>F_{0.05;2;97}$ учир $R_{1.23}$ тэгээс мэдэгдэхүйц ялгагдана.

Одоо корреляцийн тухайн коэффициентуудыг бодьё.

$$r_{12.3}=\frac{r_{12}-r_{13}\cdot r_{23}}{\sqrt{(1-r_{13}^2)(1-r_{23}^2)}}=\frac{0.20-0.41\cdot 0.82}{\sqrt{(1-0.41^2)(1-0.82^2)}}=-0.26$$

Үүнчлэн $r_{13.3}=0.44$, $r_{23.1}=0,83$.

$r_{12.3}$-коэффициент итгэлтэй эсэхийг шалгая.

$n - m + 2 = 100 - 3 + 2 = 99$ учир (12.16) шинжүүр ёсоор

$$t = \frac{-0.26 \cdot \sqrt{99-2}}{\sqrt{1-0.26^2}} = -2.65$$

Стьюдентийн таблицаас $t_{0.05;97} = 1.99$ гэж олдоно. $|t| > t_{0.05;97}$ учир корреляцийн тухайн коэффициент $r_{12.3}$ нь итгэлтэй коэффициент болно. $r_{13.2} > r_{12.3} > r_{12.3}$ тул мөн $r_{13.2}$, $r_{23.1}$ коэффициентүүд итгэлтэй. Энэ бүхнээс үзвэл, дээр авч үзсэн 3 хэмжигдэхүүний аль ч хоёр нь гурав дахийн нөлөөллийг эс тооцох үед тодорхой корреляц холбоо хамааралтай болох нь харагдаж байна. Нөгөө талаас, хос корреляцийн коэффициентоос тухайн коэффициентэд шилжихэд түүний тоон хэмжээ мэдэгдэхүйц өөрчлөгдөхөөс гадна коэффициентийн тэмдэг хүртэл өөрчлөгдөж байгаа нь харагдаж байна
($r_{12} = 0.20$; $r_{12.3} = -0.26$).

Шинжлэх ухааны аливаа судалгаа учир шалтгааны холбоо хамаарлыг илрүүлэхэд чиглэгддэг гэж хэлж болно. Үзэгдэл юмсын жинхэнэ бодит шалтгааны тухай мэдлэг л судалж буй зүй тогтлыг илэрүүлхэд тустай. Гэтэл корреляцийн тухай ойлголт нь тодорхой хэмжээгээр статистикийн формаль ойлголт тул холбоо хамаарлын учир шалтгааны талыг тэр бүр илрүүлж эс чадна. Өөрөөр хэлбэл, корреляцийн шинжилгээний тусламжтайгаар ямар хэмжигдэхүүнийг шалтгаан болгон авах ямрыг нь үр дүн болгон авахыг тогтоож болохгүй. Тухайлбал, ХАА-н ургацын хэмжээ ба цаг уурын нөхцөл (t^0, чийг, хур тундас г.м) нь корреляц хамааралтай. Энэ 2 хэмжигдэхүүний хувьд аль нь шалтгаан аль нь үр дүн болох нь тодорхой байна. Гэтэл , хоногт үйлдвэрлэх бүтээгдэхүүний хэмжээ ба үйлдвэрлэлийн үндсэн фондын хувьд алийг нь ч шалтгаан болгон авч болно. Тэгвэл, хүний биеийн жин ба өндрийн хамаарлын хувьд алийг нь ч шалтгаан болгон сонгон авч болохгүй. Зарим тохиолдолд хувьсагч хоорондын тоон харьцаан дээр тулгуурласан цэвэр формаль хамаарал байгаа боловч бодит утгаараа яаж ч тайлбарлагдахгүй хуурмаг корреляц байдаг. Нэг үгээр хэлбэл, тоон үнэлэлтээр хэмжигдэхгүй хамаарал оршин байна. Хэсэг сурагчдын математик сэтгэлгээ ба хөгжмийн авьяасын хамаарлыг үнэлье гэвэл сурагч бүрд дүн тавих замаар дүнгүүдийн хоорондын корреляц хамаарлыг авч үзэж болох боловч энэ тоон үнэлэлт төдийлэн бодитой биш. Энэ тохиолдолд корреляцийн рангийн коэффициентийг хэрэглэдэг болохыг тэмдэглэе.
Ийм учраас корреляц хамаарлын тухай тоон үнэлэлтээс логик буюу практик дүгнэлт хийхдээ судалж байгаа хэмжигдэхүүний мөн чанар, учир шалтгааны харилцан хамаарал дээр тулгуурласан үндэслэлтэй мэдээллүүдийг ашиглах хэрэгтэй.

Нөгөө талаас корреляцийн шинжилгээний урьдчилсан нөхцөл болох хувьсагчид нь олон хэмжээст хэвийн тархалттай эсэхийг шалгах ерөнхий шинжүүр практикт байхгүй. Иймд онолын загварынхаа шинж чанар, хувьсагчдын нэг хэмжээст тархалтууд хэвийн тархалтанд харшлахгүй байх чанар, хоёр хэмжээст корреляцийн орны цэгүүд нь мөн хоёр хэмжээст хэвийн тархалттай байж болох тухай шалгасан таамаглалууд дээрээ тулгуурлан олон хэмжээст хэвийн тархалттай гэж тооцож болохыг тэмдэглэе.

Бүлэг 13

Регрессийн шинжилгээ

13.1 Регрессийн шинжилгээний үндсэн бодлого

Практикийн аливаа судалгаанд тоон өгөгдлүүдийг олон хэмжээст хэвийн тархалттай олонлогоос зохиосон түүвэр мэтээр ямагт авч үзэж болдоггүй. Тухайлбал, авч үзэж буй хувьсах хэмжигдэхүүний аль нэг нь санамсаргүй биш эсвэл, регрессийн муруй шугаман биш гэх мэт байж болно. Эдгээр тохиолдын үед тоон өгөгдлүүдэд хамгийн сайн дөхөх муруйг (гадаргууг) байгуулах шаардлагатай бөгөөд ийм аргуудыг регрессийн шинжлэлд авч үзнэ. Регрессийн шинжилгээний үндсэн бодлого нь:
Хувьсагч хоорондын хамаарлын хэлбэрийг тогтоох, регрессийн загварын функцийн (12.1-ийг үз) үнэлэлтийг байгуулах, хамааран хувьсагчийн мэдэгдэхгүй байгаа утгуудыг үнэлэхэд тус тус оршино.
Регрессийн шинжилгээнд нэг буюу хэд хэдэн хувьсагчдаас ($\boldsymbol{X}$) хамаарсан, хамааран хувьсагч $\boldsymbol{Y}$-ийн нэг талын хамааралыг авч үздэг. Дараах урьдчилсан нөгцөл биелэгдсэн нөхцөлд регрессийн шинжлэл хийнэ.

1. Хамааран хувьсагч $\boldsymbol{Y}$-нь тогтмол дисперстэй, хамааралгүй санамсаргүй хэмжигдэхүүнүүд байх.
2. Үл хамаарах хувьсагч $\boldsymbol{X}$ нь санамсаргүй биш хэмжигдэхүүнүүд байх.
3. Нөхцөлт математик дундаж $M(\boldsymbol{Y}/\boldsymbol{X}=\boldsymbol{x})$-ийг $M(\boldsymbol{Y}/\boldsymbol{X}=\boldsymbol{x})=\varphi(\boldsymbol{x})$ хэлбэртэй бичиж болдог байх.

Энэ бүлэгт зөвхөн шугаман регрессийн шинжилгээг авч үзнэ. Өөрөөр хэлбэл, үнэлж буй параметрүүд нь шугаман байх, харин үл хамаарах хувьсагчид нь шугаман бишээр оролцож болох $\varphi(\boldsymbol{x},\boldsymbol{\beta})$ хэлбэрийн хамааралыг авч үзнэ. Тухайлбал:

$$M(Y/X=x)=\beta_0+\beta_1 x,\quad M(Y/X=x)=\beta_0+\beta_1 x+\beta_2 x^2,$$

$$M(Y/X=x)=\beta_0+\frac{\beta_1}{x},\quad M(Y/X=x)=\beta_0+\beta_1 x+\beta_2 x^2+\beta_3 x^3$$

гэх мэт. $\varphi(x, \beta_0, \beta_1, \ldots, \beta_q)$ функцийн хэлбэрийг сонгож авахдаа корреляцийн орны цэгүүдийн байрлалын ерөнхий характер, урьд өмнө хийсэн судалгааны туршлага, үзэгдэл процессийн физик шинж чанар дээр тулгуурласан мэргэжлийн хүний бодол дүгнэлт дээр үндэслэнэ. Шугаман регрессийн онолд нормаль регресс чухал үүрэг гүйцэтгэдэг. Регрессийн шинжилгээ явуулах урьдчилсан нөхцлүүд дээр $\boldsymbol{Y}$ хувьсагчийг хэвийн тархалттай гэж нэмж шаардвал нормаль регрессийн нөхцөл болно. Энэ тохиолдолд регрессийн коэффициентуудын үнэлэлт хэвийн тархалттай, хамгийн бага дисперстэй, хазайлтгүй болно. Эдгээр чанарууд нь регрессийн коэффициентуудын үнэлэлтийн итгэлттэй чанарыг шалгах, нөхцөлт математик дундаж
$M(\boldsymbol{Y}/\boldsymbol{X} = \boldsymbol{x})$ ба регрессийн коэффициентуудын итгэх завсрыг байгуулах боломжийг олгоно.

Санамсаргүй хэмжигдэхүүнүүдийн хосын статистик хамаарлыг нөхцөлт математик дунджаар нь $M(Y/X = x) = \varphi(x)$ хэлбэртэй илэрхийлдэг болохыг 12.1-д дурьдсан билээ. Гэвч санамсаргүй шалтгаан ба тооцогдоогүй хүчин зүйлийн нөлөөнөөс хамаарч ажиглалтын y_i утгууд загварын функцийн $\varphi(x_i)$ утгаас тодорхой хэмжээгээр хазайдаг. Иймучраас хос регрессийн загварыг (2 хэмжигдэхүүний харилцан хамаарлын тэгшитгэл) дараах хэлбэртэй дүрсэлж болно.

$$y = \varphi(x) + \varepsilon \tag{13.1}$$

Үүнд, ε-регрессийн загварын функцийн утгаас хазайх хазайлтыг илэрхийлсэн санамсаргүй хэмжигдэхүүн. Энэ хэмжигдэхүүнийг зөрөө гэж нэрлэнэ.

$\varphi(x)$ функцийн үнэлэлтийг байгуулахын тулд (X, Y)-санамсаргүй хэмжигдэхүүний системээс n-хэмжээст түүвэр авсан гэе. i-р туршилтын үр дүнг (x_i, y_i) гэе. $(i = \overline{1, n})$
Энэ тохиолдолд регрессийн дээрх загвар

$$y_i = \varphi(x_i) + \varepsilon_i \tag{13.2}$$

хэлбэртэй болно. Энэ загварын хувьд регрессийн шинжилгээний урьдчилсан нөхцөлүүдийг дэлгэрүүлэн томьёолбол:

1. y_i (эсвэл ε_i)-нь санамсаргүй хэмжигдэхүүн байх. x_i-нь санамсаргүй биш хэмжигдэхүүн байх.

2. ε_i зөрөөний математик дундаж тэгтэй тэнцүү:

$$M(\varepsilon_i) = 0. \tag{13.3}$$

3. Дурын i-ийн хувьд , хамааран хувьсагч y_i-ийн дисперс (эсвэл ε_i зөрөөний) тогтмол:

$$D(\varepsilon_i) = \sigma_\varepsilon^2. \tag{13.4}$$

4. Ялгаатай $i,\ j$-ийн хувьд $y_i,\ y_j$ хувьсагчид (эсвэл $\varepsilon_i,\ \varepsilon_j$) корреляц хамааралгүй:

$$M(\varepsilon_i \cdot \varepsilon_j) = 0. \tag{13.5}$$

5. Хамааран хувьсагч y_i (эсвэл ε_i)-нь хэвийн тархалттай.
Сүүлчийн шаардлага нь нормаль регрессийн нөхцөл болно.

13.2 Шугаман регресс

$\varphi(x, \beta_0, \beta_1, \ldots, \beta_q)$ хамаарал, үнэлж буй параметрүүд ба хувьсагчдынхаа хувьд нэгэн зэрэг шугаман байвал регрессийн шинжлэлийн хамгийн хялбар загвар болно. Хэрэв регрессийн загварын тэгшитгэл

$$\varphi(X) = \beta_0 + \beta_1 x \tag{13.6}$$

хэлбэртэй байвал Y-ийг X дээр шугаман регресстэй байна гэж ярих ба β_0, β_1-ийг регрессийн коэффициентууд гэнэ. β_0, β_1 параметрүүдийн тусламжтайгаар X-хувьсагчийн Y-д нөлөөлөх нөлөөг тооцох бөгөөд харин тооцогдоогүй хүчин зүйл болон туршилтын санамсаргүй алдааны нөлөөг үлдэгдэл дисперсийн ($\sigma^2_{\text{үлд}}$) тусламжтай тодорхойлно.
β_0, β_1 параметрүүдийн үнэлэлтийг b_0, b_1, $\sigma^2_{\text{үлд}}$ параметрийн үнэлэлтийг $S^2_{\text{үлд}}$ гэж тус тус тэмдэглэе. Регрессийн онолд b_0, b_1-үнэлэлтийг хамгийн бага квадратын аргаар олох боловч (13.3-д үз) шугаман регресс нь корреляцийн онолд чухал үүрэг гүйцэтгэх учраас дээрх 2 коэффициентийг корреляцийн ерөнхий коэффициентын тусламжтай зөвхөн ганц коэффициент байдлаар илэрхийлж болдог. ((13.9)-томъёог үз)

Эхлээд шугаман регрессийн коэффициент корреляцийн коэффициенттэй хэрхэн холбогдохыг авч үзье. Корреляцийн онолд, X, Y-санамсаргүй хэмжигдэхүүнүүд хамтын нормаль тархалттай үед нөхцөлт математик дунджийг ашиглан шугаман регрессийн тэгшитгэлүүдийг дараах хэлбэртэй авч үздэг. (12.2-ыг үз).

$$\overline{y}_x - \overline{Y} = r \cdot \frac{S_y}{S_x}(x - \overline{X}) \tag{13.7}$$

$$\overline{x}_y - \overline{X} = r \cdot \frac{S_x}{S_y}(y - \overline{Y}) \tag{13.8}$$

$$b_{yx} = r \cdot \frac{S_y}{S_x} \tag{13.9}$$

коэффициентийг Y-ийн X дээрх шугаман регрессийн (ерөнхий) коэффициент гэнэ.
b_{yx}-коэффициент нь, X хувьсагч нэг нэгжээр өөрчлөгдөхөд Y хувьсагч дунджаар хэдэн нэгжээр өөрчлөгдөхийг илэрхийлнэ. Үүний адилаар

$$b_{xy} = r \cdot \frac{S_x}{S_y} \tag{13.10}$$

коэффициентийг X-ийн Y дээрх шугаман регрессийн коэффициент гэнэ. (13.9) ба (13.10) ёсоор (13.7), (13.8) тэгшитгэлүүд дараах хэлбэртэй болно.

$$\overline{y}_x - \overline{Y} = b_{yx} \cdot (x - \overline{X}) \tag{13.11}$$

$$\overline{x}_y - \overline{X} = b_{yx} \cdot (y - \overline{Y}) \tag{13.12}$$

Санамсаргүй хэмжигдэхүүний ажиглалтын утгууд (x_i, y_i)-ээр b_{yx}, b_{yx} параметрүүдийг олох томъёог бичвэл $(i = \overline{1,k},\ j = \overline{1,s})$:

$$b_{yx} = \frac{\overline{X \cdot Y} - \overline{X} \cdot \overline{Y}}{S_x^2} = \frac{K_{xy}}{S_x^2} \tag{13.13}$$

$$K_{xy} = \frac{1}{n} \sum_{i=1}^{k} \sum_{j=1}^{s} x_i y_j n_{ij} - \overline{X} \cdot \overline{Y} \tag{13.14}$$

$$S_x^2 = \frac{1}{n} \sum_{i=1}^{k} x_i^2 n_i - \overline{X}^2 \tag{13.15}$$

$$b_{xy} = \frac{K_{xy}}{S_y^2} \tag{13.16}$$

$$S_y^2 = \frac{1}{n} \sum_{j=1}^{s} y_j^2 n_j - \overline{Y}^2 \tag{13.17}$$

Үүнд, $K_{xy} - X, Y$-ийн корреляцийн момент, $\overline{X}, \overline{Y}$−түүврийн дунджууд, S_x^2, S_y^2−түүврийн дисперсүүд. (13.9), (13.10) томьёоноос $r^2 = b_{yx} \cdot b_{xy}$ буюу

$$r = \pm\sqrt{b_{yx} \cdot b_{xy}} \tag{13.18}$$

томъёо мөрдөнө. Иймд түүврийн корреляцийн коэффициент нь түүврийн регрессийн коэффициентуудын геометр дундаж болж байна.

Хэрэв корреляцийн коэффициент r-итгэлтэй бол b_{yx}, b_{xy} үнэлэлтүүд итгэлтэй байх ба эх олонлогийн шугаман регрессийн коэффициентууд β_{yx}, β_{xy}-ийн итгэх завсрыг дараах томъёогоор олно.

$$b_{yx} - t_{\alpha;\ n-2} \cdot \frac{S_y\sqrt{1-r^2}}{S_x\sqrt{n-2}} \le \beta_{yx} \le b_{yx} + t_{\alpha;\ n-2} \cdot \frac{S_y\sqrt{1-r^2}}{S_x\sqrt{n-2}} \tag{13.19}$$

$$b_{xy} - t_{\alpha;\ n-2} \cdot \frac{S_x\sqrt{1-r^2}}{S_y\sqrt{n-2}} \le \beta_{xy} \le b_{xy} + t_{\alpha;n-2} \cdot \frac{S_x\sqrt{1-r^2}}{S_y\sqrt{n-2}} \tag{13.20}$$

$t_{\alpha;n-2}$−Стьюдентийн тархалтын таблицын утга.

Жишээ-13.1 Жишээ-12.2-д авч үзсэн, үйлдвэрлэлийн үндсэн фонд-X, хоногт үйлдвэрлэх бүтээгдэхүүний хэмжээ Y-ийн өгөгдлөөр тэдгээрийн регрессийн тэгшитгэл ба регрессийн коэффициентуудыг олж үнэлэлтийг байгуул.

Таблиц-12.2-д олж үзүүлсэн бүлгийн дунджууд дээр үндэслэн X, Y хэмжигдэхүүний хоорондын хамаарал шугаман байж болох тухай дүгнэлт хийсэн болохоор регрессийн тэгшитгэлийг $\overline{y}_x = b_0 + b_1 x$ хэлбэртэй эрж болно.

Дурдсан жишээнд (13.13)÷(13.17) томъёоны нийлбэрүүдийг нэгэнт олсон учраас

$$K_{yx} = \frac{27895}{50} - 32.1 \cdot 16.92 = 14.768,$$

$$b_{yx} = \frac{14.768}{21.84} = 0.6762, \quad b_{xy} = \frac{14.768}{18.2336} = 0.8099.$$

Иймд регрессийн тэгшитгэлүүд $\overline{y}_x = 0.6762x - 4.79$, $\overline{x}_y = 0.8099y + 16.70$ хэлбэртэй болно. Эхний тэгшитгэлээс, хэрэв үйлдвэрлэлийн үндсэн фондыг 1 сая төгрөгөөр нэмэгдүүлбэл хоногт үйлдвэрлэх бүтээгдэхүүний хэмжээ дунджаар 0.6762 тонноор нэмэгдэх нь харагдаж байна. Хоёр дахь тэгшитгэл хэрэв, хоногт үйлдвэрлэх бүтээгдэхүүний хэмжээг 1 тонноор нэмэгдүүлье гэвэл, үйлдвэрлэлийн үндсэн фондоо дунджаар 0.8099 сая төгрөгөөр нэмэгдүүлэх шаардлагатай болохыг илэрхийлнэ.
Мөн Жишээ-12.3-д корреляцийн коэффициент r-ийг итгэлтэй гэж харуулсан учраас b_{xy}, b_{yx}-үнэлэлтүүд итгэлтэй тул β_{xy}, β_{yx} коэффициентуудын итгэх завсрыг байгуулъя.

$$S_x = \sqrt{S_x^2} = \sqrt{21.84}, \quad S_y = \sqrt{S_y^2} = \sqrt{18.2336} = 4.270.$$

$\alpha = 0.05$, $k = n = 48$ үед Стьюдентийн тархалтын таблицаас $t_{\alpha;n-2}$-ыг олбол $t_{0.05;48} = 2.01$ (13.19), (13.20) томъёо ёсоор, итгэх завсрууд
$0.4979 \leq \beta_{yx} \leq 0.8545$, $\quad 0.5963 \leq \beta_{xy} \leq 1.0235$ болно.

Шугаман регрессийн хувьд (13.2) загвар дараах хэлбэртэй болно.

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \tag{13.21}$$

β_0, β_1 коэффициентуудын үнэлэлт b_0, b_1-ийг байгуулахдаа түүврийн корреляцийн коэффициент ашиглахгүйгээр, хамгийн бага квадратын аргаар хэрхэн олохыг дараагийн дэд бүлэгт дэлгэрэнгүй өгүүлэх болно. Харин үлдэгдэл дисперс $\sigma^2_{\text{үлд}}$-ийн үнэлэлт $S^2_{\text{үлд}}$−ийг дараах томъёогоор олно.

$$S^2_{\text{үлд}} = \frac{1}{n-2}\sum_{i=1}^{n}(\overline{y}_{x_i} - y_i)^2 = \frac{1}{n-2}\sum_{i=1}^{n} e_i^2 \tag{13.22}$$

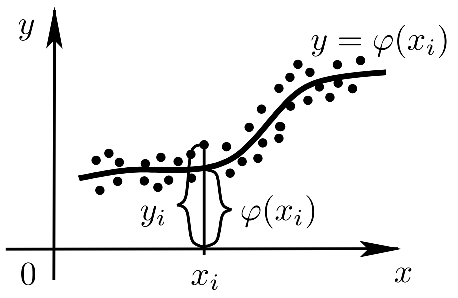
Үүнд, $\overline{y}_{x_i}$-регрессийн тэгшитгэлээр олсон нөхцөлт дунджийн утгууд,
e_i−регрессийн үлдэгдэл буюу ε_i−ийн түүврийн үнэлэлт.
Хэрэв шугаман загвар q тооны параметр агуулж байвал үлдэгдэл дисперс дараах хэлбэртэй болно.

$$S^2_{\text{үлд}} = \frac{1}{n-q}\sum_{i=1}^{n}(\overline{y}_{x_i} - y_i)^2 = \frac{1}{n-q}\sum_{i=1}^{n} e_i^2 \tag{13.23}$$

13.3 Регрессийн коэффициентуудыг олох хамгийн бага квадратын арга

Статистик хамааралтай 2 хувьсах хэмжигдэхүүний регрессийн тэгшитгэл болон энэхүү тэгшитгэлийн параметрүүдийн үнэлэлтийг олоходоо түгээмэл хэрэглэдэг арга нь хамгийн бага квадратын арга юм.
$y_i = \varphi(x_i) + \varepsilon_i$ (13.2) гэсэн загварын $\varphi(x_i)$ функц $b_0, b_1, \ldots, b_q$ параметрүүдийг агуулдаг ($\varphi(x_i) = \varphi(x_i, b_0, b_1, \ldots, b_q)$) бөгөөд эдгээр үл мэдэгдэх параметрүүдийн хамгийн оновчтой утгыг олох зорилго тавья. n удаагийн туршилтын үр дүнд X, Y хэмжигдэхүүний ажиглагдсан утгуудаар дараах статистик таблиц зохиоё.

туршилтын тоо	1	2	3	...	i	...	n
x_i	x_1	x_2	x_3	...	x_i	...	x_n
y_i	y_1	y_2	y_3	...	y_i	...	y_n



Зураг 13.1:

Туршилтын утга x_i-д харгалзах функцийн утга $\varphi(x_i)$ нь ажиглалтын утга y_i-тэй давхцах албагүй. Өөрөөр хэлбэл зарим нэгэн x_i утгууд ба эсвэл бүх утгуудын хувьд $y_i - \varphi(x_i) \neq 0$ байна.
Иймд $\varepsilon_i = y_i - \varphi(x_i)$ нь туршилтын утгууд ба регрессийн тэгшитгэлээр олсон утгуудын зөрөөг илэрхийлнэ.

Хамгийн бага квадратын аргын үндсэн санаа нь дээрх ε_i-зөрөөний квадратуудын нийлбэр хамгийн бага байхаар b_i параметрүүдийг сонгон авахад оршино. Өөрөөр хэлбэл:

$$U = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} \Big(y_i - \varphi(x_i, b_0, b_1, \ldots, b_q) \Big)^2 \longrightarrow \min \qquad (13.24)$$

нөхцлийг хангах $b_0, b_1, \ldots, b_q$-г олох бодлого болно. Хэрэв $U = U(b_0, b_1, \ldots, b_q)$ функц бүх хувьсагчуудаараа тасралтгүй тухайн уламжлалтай бол олон хувьсагчийн функцийн экстремум орших нөхцөл ёсоор

$$\frac{\partial U}{\partial b_0} = 0, \quad \frac{\partial U}{\partial b_1} = 0, \ldots, \frac{\partial U}{\partial b_q} = 0, \qquad (13.25)$$

нөхцөл биелэгдэнэ.
(13.25) нь $b_0, b_1, \ldots, b_q$ параметрүүдийн хувьд $(q+1)$ хувьсагч бүхий систем тэгшитгэл юм. Энэ системийг нормаль тэгшитгэлүүдийн систем буюу товчоор нормаль систем гэнэ. Ихэнх тохиолдолд $\varphi(x)$ функц

$$\varphi(x) = \sum_{k=0}^{q} b_k f_k(x) \qquad (13.26)$$

хэлбэртэй байдаг. Үүнд $f_k(x)$ $(k=\overline{0,q})$−өгөгдсөн функцүүд.
Энэ тохиолдолд U функц ба нормаль систем дараах хэлбэртэй болно.

$$U=\sum_{i=1}^{n}\Big(y_i-\sum_{k=0}^{q}b_kf_k(x_i)\Big)^2 \tag{13.27}$$

$$\left.\begin{array}{l}
\sum\limits_{i=1}^{n}\Big(y_i-\sum\limits_{k=0}^{q}b_kf_k(x_i)\Big)\cdot(-f_0(x_i))=0\\
\sum\limits_{i=1}^{n}\Big(y_i-\sum\limits_{k=0}^{q}b_kf_k(x_i)\Big)\cdot(-f_1(x_i))=0\\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
\sum\limits_{i=1}^{n}\Big(y_i-\sum\limits_{k=0}^{q}b_kf_k(x_i)\Big)\cdot(-f_q(x_i))=0
\end{array}\right\} \tag{13.28}$$

Тухайн тохиолдолд

$$\varphi(x,b_0,b_1,\dots,b_q)=b_0+b_1x+b_2x^2+\cdots+b_qx^q \tag{13.29}$$

бол нормаль систем дараах хэлбэртэй болно.

$$\left.\begin{array}{l}
b_0n+b_1\cdot\sum\limits_{i=1}^{n}x_i+\cdots+b_q\cdot\sum\limits_{i=1}^{n}x_i^q=\sum\limits_{i=1}^{n}y_i\\
b_0\sum\limits_{i=1}^{n}x_i+b_1\cdot\sum\limits_{i=1}^{n}x_i^2+\cdots+b_q\cdot\sum\limits_{i=1}^{n}x_i^{q+1}=\sum\limits_{i=1}^{n}y_ix_i\\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
b_0\sum\limits_{i=1}^{n}x_i^q+b_1\cdot\sum\limits_{i=1}^{n}x_i^{q+1}+\cdots+b_q\cdot\sum\limits_{i=1}^{n}x_i^{2q}=\sum\limits_{i=1}^{n}y_ix_i^q
\end{array}\right\} \tag{13.30}$$

Хэрэв $q=2$ бол $\varphi(x,b_0,b_1,b_2)=b_0+b_1x+b_2x^2$ болох ба регрессийн коэффициентуудыг дараах системээс олно.

$$\left.\begin{array}{l}
b_0n+b_1\cdot\sum\limits_{i=1}^{n}x_i+b_2\cdot\sum\limits_{i=1}^{n}x_i^2=\sum\limits_{i=1}^{n}y_i\\
b_0\sum\limits_{i=1}^{n}x_i+b_1\cdot\sum\limits_{i=1}^{n}x_i^2+b_2\cdot\sum\limits_{i=1}^{n}x_i^3=\sum\limits_{i=1}^{n}y_ix_i\\
b_0\sum\limits_{i=1}^{n}x_i^2+b_1\cdot\sum\limits_{i=1}^{n}x_i^3+b_2\cdot\sum\limits_{i=1}^{n}x_i^4=\sum\limits_{i=1}^{n}y_ix_i^2
\end{array}\right\} \tag{13.31}$$

$q=1$ тохиолдолд X, Y хэмжигдэхүүний хамаарал шугаман болох бөгөөд

$$y=b_0+b_1x \tag{13.32}$$

гэсэн шугаман регрессийн коэффициентуудыг

$$\left.\begin{array}{l}
b_0n+b_1\cdot\sum\limits_{i=1}^{n}x_i=\sum\limits_{i=1}^{n}y_i\\
b_0\sum\limits_{i=1}^{n}x_i+b_1\cdot\sum\limits_{i=1}^{n}x_i^2=\sum\limits_{i=1}^{n}y_ix_i
\end{array}\right\} \tag{13.33}$$

системээс олбол:

$$\left.\begin{aligned} b_1 &= \frac{n\cdot\sum x_i y_i - \sum x_i \cdot \sum y_i}{n\cdot\sum x_i^2 n_i - \left(\sum x_i\right)^2} \\ b_0 &= \frac{1}{n}\cdot\sum y_i - b_1\frac{1}{n}\sum x_i \end{aligned}\right\} \tag{13.34}$$

Хэрэв түүврийн хэмжээ их бол тооцоог хялбарчлахын тулд ажиглалтын утгуудыг бүлэглэж корреляцийн таблицыг зохиодог. Иймээс ажиглалтын утгууд нь корреляцийн таблицаар өгөгдсөн тохиолдолд (13.33) систем дараах хэлбэртэй болно.

$$\left.\begin{aligned} & b_0 n + b_1\cdot\sum_i x_i n_i = \sum_j y_j n_j \\ & b_0\sum_i x_i n_i + b_1\cdot\sum_i x_i^2 n_i = \sum_{i,j} x_i y_j n_{ij} \end{aligned}\right\} \tag{13.35}$$

Хамгийн бага квадратын аргаар олсон үнэлэлтүүд нь шугаман үнэлэлтийн ангид хамгийн бага дисперстэй байдгаараа давуу талтай.

Жишээ-13.2 Д.И.Менделеев азотот натрийн давсны ($NaNO_3$) уусах чанар ба усны температурын хамаарлыг судалжээ.

x_i	0	4	10	15	21	29	36	51	68
y_i	66.7	71.0	76.3	80.6	85.7	92.9	99.4	113.6	125.1

Судалгаагаар, усны t^0 ба уусах чанарын хооронд $y = 0,87x + 67,5$ гэсэн хамааралтай гэж дүгнэжээ. Энэ дүгнэлтийг нотолж үзүүл.

Дүгнэлтийг нотлохын тулд Y-ийн X дээрх шугаман регрессийн тэгшитгэлийг $y = b_0 + b_1 x$ хэлбэртэй эрье. Дараах хүснэгтийг ашиглан тооцоо хийе.

i	x_i	y_i	$x_i y_i$	x_i^2
1	0	66.7	0	0
2	4	71.0	284.0	16
3	10	76.3	763.0	100
4	15	80.6	1209.0	225
5	21	85.7	1799.7	441
6	29	92.9	2694.1	841
7	36	99.4	3578.4	1296
8	51	113.6	5793.6	2601
9	68	125.1	8506.8	4624
$\sum$	234	811.3	24628.6	10144

Нормаль системийг бичвэл $\left.\begin{aligned} 9b_0 + 234b_1 &= 811.3 \\ 234b_0 + 10144b_1 &= 24628.6 \end{aligned}\right\}$

болох ба шийд нь: $b_0 = 67.5$, $b_1 = 0.87$. Иймд регрессийн тэгшитгэл

$y = 67.5 + 0.87x$ хэлбэртэй нь батлагдав.
Ямарч туршилтын өгөгдлүүдийг хамгийн бага квадратын аргаар боловсруулж болох боловч зөвхөн санамсаргүй хэмжигдэхүүнүүд хэвийн тархалттай үед энэхүү арга оновчтой болохыг тэмдэглэе.

13.4 Шугаман биш регресс

Үйлдвэрлэл, шинжлэх ухаан, нийгэм, эдийн засгийн аливаа үзэгдэл процессуудын хоорондох харьцаа хамаарал ямагт шугаман хуулиар илэрхийлэгдэх албагүй учраас хувьсагчдынхаа хувьд шугаман биш регресс авч үзэх шаардлага гардаг.
Практикт нилээд элбэг тааралдах шугаман биш регрессүүд нь:
$\overline{y}_x = b_0 + b_1 x + \cdots + b_q x^q$ (олон гишүүнт регресс)
$\overline{y}_x = b_0 + \dfrac{b_1}{x}$ (гипербол регресс)
$\overline{y}_x = b_0 x_1^{b_1} \ldots x_p b_p$ (зэрэгт регресс) юм.
Олон гишүүнт регрессийн коэфициентуудыг хэрхэн олохыг өмнөх зүйлийн (13.29), (13.30) томъёонд дурдсан. Зэрэгт регрессийн бусдаасаа ялгагдах онцлог нь b_j параметрүүдийнхээ хувьд шугаман биш боловч логарифмчлах замаар түүнийг шугаман регрессд шилжүүлж болдогт оршино.

$$Ln\overline{y}_x = Lnb_0 + b_1 Lnx_1 + \cdots + b_p Lnx_p$$

$b_0, b_1, \ldots, b_p$-үл мэдэгдэх параметрүүдийг олохдоо хамгийн бага квадратын аргыг хэрэглэнэ. Тухайн тохиолдолд

$$\overline{y}_x = b_0 x^{b_1} \tag{13.36}$$

хэлбэрийн зэрэгт регрессийн коэффициентүүдийг хэрхэн олохыг үзүүлье. (13.36)-г логарифмчилбал $Lg\overline{y}_x = Lgb_0 + b_1 Lgx$. $Lgb_0 = B$, $Lgx = X$, $Lg\overline{y}_x = Y$ гэвэл $Y = b_1 X + B$ хэлбэрийн шугаман регресс болно. Манай тохиолдолд $U = U(B, b_1)$ функц

$$U = \sum_{i=1}^{n} [Lgy_i - (B + b_1 Lgx_i)]^2 \tag{13.37}$$

болох ба нормаль систем дараах хэлбэртэй болно.

$$\left.\begin{array}{l} B \cdot n + b_1 \cdot \sum\limits_{i=1}^{n} Lgx_i = \sum\limits_{i=1}^{n} Lgy_i \\ B \cdot \sum\limits_{i=1}^{n} Lnx_i + b_1 \cdot \sum\limits_{i=1}^{n} (Lgx_i)^2 = \sum\limits_{i=1}^{n} Lgx_i \cdot Lgy_i \end{array}\right\} \tag{13.38}$$

(13.38) системээс b_1 болон B-г олж $Lgb_0 = B$ нөхцөлөөс b_0-параметрийг олно.

Гипербол регрессийг $Z = \dfrac{1}{x}$ орлуугын тусламжтайгаар шугаман регрессд шилжүүлж болно.
Нормаль системийг бичвэл:

$$\left.\begin{aligned} nb_0 + b_1 \cdot \sum_{i=1}^{n} \frac{1}{x_i} &= \sum_{i=1}^{n} y_i \\ b_0 \cdot \sum_{i=1}^{n} \frac{1}{x_i} + b_1 \sum_{i=1}^{n} \frac{1}{x_i^2} &= \sum_{i=1}^{n} \frac{y_i}{x_i} \end{aligned}\right\} \tag{13.39}$$

Хэрэв түүвэр корреляцийн таблиц хэлбэрээр өгөгдсөн бол гипербол регрессийн нормаль систем дараах хэлбэртэй болно.

$$\left.\begin{aligned} nb_0 + b_1 \cdot \sum_{i} \frac{n_i}{x_i} &= \sum_{j} y_j n_j \\ b_0 \cdot \sum_{i} \frac{n_i}{x_i} + b_1 \sum_{i} \frac{n_i}{x_i^2} &= \sum_{i,j} \frac{y_i}{x_i} n_{ij} \end{aligned}\right\} \tag{13.40}$$

Жишээ-13.3 Ижил төрлийн 30 үйлдвэрийн нэг өдөрт үйлдвэрлэсэн бүтээгдэхүүний хэмжээ (X) ба нэгж бүтээгдэхүүний өөрийн өртөг (Y) дараах хүснэгтэнд өгөгджээ.

X \ Y	100	110	120	130	$\sum$
50	-	-	1	3	4
100	-	3	3	-	6
150	-	6	2	1	9
200	1	4	-	-	5
250	4	1	1	-	6
$\sum$	5	14	7	4	30

Өдөрт үйлдвэрлэх бүтээгдэхүүний хэмжээ ба өөрийн өртгийн корреляц хамаарлын хэлбэрийг олж, тэгшитгэлийг бич.

Y-санамсаргүй хэмжигдэхүүний бүлэглэсэн дунджуудыг олъё.

$$\overline{y}_1 = \frac{1}{4}(120 \cdot 1 + 130 \cdot 3) = 127.5, \quad \overline{y}_2 = \frac{1}{6}(110 \cdot 3 + 120 \cdot 3) = 115,$$

$$\overline{y}_3 = 114.44, \quad \overline{y}_4 = 108, \quad \overline{y}_5 = 1054.$$

x_i утгуудыг өсөхөд $\overline{y}_i$ бүлгийн дунджууд буурч, бууралт нь аажимдаа багасч буй тул Y-ийг X дээр гипербол регресстэй гэж үзэх үндэстэй. Иймд регрессийн тэгштгэлиийг $y = b_0 + \dfrac{b_1}{x}$-хэлбэртэй эрж нормаль системийн коэффициентүүдийг олохын тулд бодолтыг дараах хүснэгтээр гүйцэтгэе.

x_i	n_i	$\dfrac{1}{x_i}$	$\dfrac{1}{x_i}n_i$	$\dfrac{1}{x_i^2}n_i$	$\overline{y}_j$	$\overline{y}_j n_j$	$\overline{y}_j n_j / x_i$
50	4	0.02	0.08	0.0016	127.5	510	10.2
100	6	0.01	0.06	0.0006	115	690	6.9
150	9	0.0067	0.06	0.0004	114.44	1030	6.87
200	5	0.005	0.025	0.000125	108	540	2.7
250	6	0.004	0.024	0.000096	105	630	2.52
$\sum$	30	-	0.249	0.002821	-	3400	29.19

Дараах нормаль системээс b_0, b_1 параметрүүдийг олбол:

$$\left.\begin{array}{l} 30 \cdot b_0 + 0.249 \cdot b_1 = 3400 \\ 0.249 \cdot b_0 + 0.002821 b_1 = 29.19 \end{array}\right\} \qquad \begin{array}{l} b_0 = 102.66 \\ b_1 = 1256.96 \end{array}$$

Иймд бүтээгдэхүүний өөрийн өртөг бүтээгдэхүүний хэмжээнээс хамаарах регресс хамаарал $y = \dfrac{1256.96}{x} + 102.66$ тэгшитгэлээр илэрхийлэгдэнэ.

13.5 Шугаман регрессийн коэффициентуудын итгэлтэй чанарыг үнэлэх ба завсран үнэлэлт байгуулах

Регрессийн тэгшитгэлийг олсны дараа гаргаж авсан үр дүндээ статистик шинжилгээ хийх шаардлагатай. Энэ нь регрессийн бүх коэффициентуудын итгэлтэй эсэхийг шалгах, регрессиийн тэгшитгэлийн илэрхийлэх чадварыг тогтооход оршино. Коэффициентуудын үнэлэлт итгэлтэй эсэхийг шалгана гэдэг нь шугаман регрессийн коэффициент тэгээс ялгаатай байх тухай үндэслэлтэй статистик дүгнэлт гаргахад эдгээр үнэлэлтүүдийн утга хангалттай юу? гэдгийг тогтооно гэсэн үг. Үүний тулд нормаль регрессийн нөхцөлд регрессийн коэффициент тэгтэй тэнцүү байх тухай $\boldsymbol{H}_0 : \boldsymbol{\beta} = \boldsymbol{0}$ гэсэн анхны таамаглалыг шалгана. Энэ таамаглалыг шалгахад дараах шинжүүрийг ашиглах ба $k = n - 2$ чөлөөний зэрэг бүхий Стьюдентийн тархалттай байна.

$$\boldsymbol{t} = \frac{|\boldsymbol{b}|}{S_{\boldsymbol{b}}} \tag{13.41}$$

Үүнд: $\boldsymbol{b}$-регрессийн коэффициентийн үнэлэлт. $S_{\boldsymbol{b}}$-регрессийн коэффициентийн дундаж квадрат хазайлтын үнэлэлт (өөрөөр хэлбэл үнэлэлтийн стандарт алдаа).
Өгөгдсөн итгэх түвшин α-ийн хувьд $|\boldsymbol{t}| \geq \boldsymbol{t}_{\alpha,k}$ нөхцөл биелэгдвэл $\boldsymbol{H}_0$ таамаглалыг хэрэгсэхгүй. Өөрөөр хэлбэл коэффициент итгэлтэй.
$\boldsymbol{t}_{\alpha,k}$-Стьюдентийн таблицын утга. Итгэлтэй биш коэффициентуудыг регрессийн тэгшитгэлээс хасна.

Хос регрессийн хувьд, коэффициентуудын дундаж квадрат хазайлтын үнэлэлтийг дараах томъёогоор олно.

$$S_{b_0} = \frac{S^2_{\text{үлд}}}{\sqrt{n-2}}, \quad S_{b_1} = \frac{S^2_{\text{үлд}}}{S_x\sqrt{n-2}} \tag{13.42}$$

Үүнд: $S^2_{\text{үлд}}$ нь (13.22) томъёогоор бодогдох үлдэгдэл дисперсийн үнэлэлт,

$$S_x = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{X})^2}$$

Итгэлтэй параметруудын итгэх завсрыг $P(-\boldsymbol{t}_{\alpha;\ k} < \dfrac{\boldsymbol{\beta} - \boldsymbol{b}}{S_{\boldsymbol{b}}} < \boldsymbol{t}_{\alpha;\ k}) = 1-\alpha$ нөхцлөөс олбол

$$\boldsymbol{b} - \boldsymbol{t}_{\alpha;\ k} \cdot S_{\boldsymbol{b}} < \boldsymbol{\beta} < \boldsymbol{b} + \boldsymbol{t}_{\alpha;\ k} \cdot S_{\boldsymbol{b}}\ . \tag{13.43}$$

13.6 Нөхцөлт математик дунджийн завсран үнэлэлт

Регрессийн тэгшитгэлээр олсон $\overline{y}_x$ утга нь нөхцөлт математик дундаж $M(Y/X = x)$-ийн цэгэн үнэлэлт юм. Одоо хос шугаман регрессийн тохиолдолд $M(Y/X = x)$ дунджийн завсран үнэлэлтийг байгуулья.

$\beta_0,\ \beta_1$ параметрүүд түүврийн утгуудаар үнэлэгдэх учраас тэдгээрийн үнэлэлт $b_0,\ b_1$ нь санамсаргүй алдааг агуулдаг. b_0 утгын алдаа нь регрессийн шугам босоо тэнхлэгийн дагуу шилжихэд хүргэнэ. Харин b_1 үнэлэлтийн алдаа нь $(\overline{X}, \overline{Y})$ цэгтэй харьцангуйгаар регрессийн муруйг хэлбэлзэхэд хүргэнэ. Ийм учраас $\overline{y}_x = \overline{Y} + b_1(x - \overline{X})$ гэсэн регрессийн тэгшитгэлээр олсон утга санамсаргүй алдааг агуулна.

Эхлээд нөхцөлт дундаж $\overline{y}_x$-ийн дисперс $S^2_{\overline{y}_x}$-ийг ольё. Түүнийг үл хамаарах 2 нэмэгдэхүүний дисперсийн нийлбэрт тавьж болно.

Түүврийн дунджийн дисперс:

$$S^2_{\overline{Y}} = \frac{S^2_{\text{үлд}}}{n} \tag{13.44}$$

Регрессийн коэффициент b_1-ийн дисперс:

$$S^2_{b_1} = \frac{S^2_{\text{үлд}}}{\sum(x_i - \overline{X})^2} = \frac{S^2_{\text{үлд}}}{nS^2_x} \tag{13.45}$$

томъёогоор тус тус тодорхойлогдох тул

$$S^2_{\overline{y}_x} = S^2_{\text{үлд}} \left[\frac{1}{n} + \frac{(x - \overline{X})^2}{nS^2_x}\right] \tag{13.46}$$

болно.

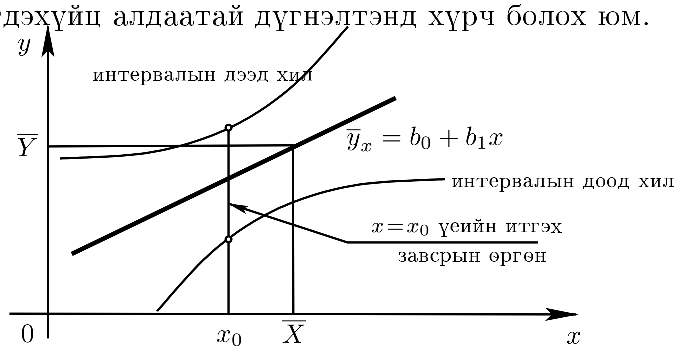
$$t = \frac{\overline{y} - M(Y/X = x)}{S_{\overline{y}_x}} \tag{13.47}$$

шинжүүр нь $k = n - 2$ чөлөөний зэрэг бүхий Стьюдентийн тархалттай учир энэхүү шинжүүрийг ашиглан, өгөгдсөн итгэх түвшин α-ийн хувьд нөхцөлт математик дунджийн итгэх завсрыг олбол

$$\overline{y}_x - S_{\overline{y}_x} \cdot t_{\alpha;k} \leq M(Y/X = x) \leq \overline{y}_x + S_{\overline{y}_x} \cdot t_{\alpha;k} \tag{13.48}$$

Үүнд $S_{\overline{y}_x}$ нь нөхцөлт дунджийн стандарт алдаа.

(13.46), (13.48) томъёоноос үзвэл итгэх завсрын өргөн нь x-хувьсагчийн утгаас хамаарч байна. $x = \overline{X}$ үед хамгийн бага өргөнтэй, x-ийн утга $\overline{X}$ утгаас холдох тутам өргөн болж байна. Иймд, регрессийн тэгшитгэлийг үндэслэн хийсэн урьдчилсан дүгнэлтээр, үл хамаарах хувьсагчийн утга түүврийн утгуудын хязгаараас гарахгүй нөхцөлд нарийвчлал сайтай байна. Эсрэг тохиолдолд мэдэгдэхүйц алдаатай дүгнэлтэнд хүрч болох юм.



Зураг 13.2:

(13.46) ба (13.48) үнэлэлтийг x-ийн тодорхой x_0 утгын хувьд байгуулахдаа $x = x_0$ утгыг орлуулна. Нөхцөлт математик дунджийн итгэх завсар нь (Зураг 13.2) регрессийн загварын муруйн орших мужийг тодорхойлж буй боловч хамааран хувьсагчийн ямар нэг ганцаарчилсан $\overline{y}_0$ утгын итгэх завсрыг үл тодорхойлно. Иймд хувьсагчийн ганцаарчилсан утга $\overline{y}_0$-ийн итгэх завсрыг олохдоо түүний регрессийн шугамын орчимд сарних сарнилыг тооцох шаардлагатай. Өөрөөр хэлбэл, нийлбэр дисперс $S^2_{y_0}$−д үлдэгдэл дисперс $S^2_{\text{үлд}}$−ийг оруулан тооцох ёстой. Тэгвэл $x = x_0$ үед ганцаарчилсан утга $\overline{y}_0$-ийн дисперс:

$$S^2_{\overline{y}_0} = S^2_{\text{үлд}}\left[1 + \frac{1}{n} + \frac{(x_0 - \overline{X})^2}{nS_x^2}\right] \tag{13.49}$$

Харин харгалзах итгэх завсар дараах хэлбэртэй болно:

$$\overline{y}_x - S_{\overline{y}_0} \cdot t_{\alpha;k} \le y_0^* \le \overline{y}_x + S_{\overline{y}_0} \cdot t_{\alpha;k} \tag{13.50}$$

Үүнд: $S_{\overline{y}_0} - \overline{y}_0$ утгын стандарт алдаа.

Жишээ-13.4 Жишээ-12.2-ын өгөгдлөөр үйлдвэрлэлийн үндсэн фонд нь 36 (сая төгрөг) байх үйлдвэрүүдийн хоногт үйлдвэрлэх бүтээгдэхүүний дундаж хэмжээ болон тусгайлан сонгон авсан ганцаарчилсан утгын итгэх завсрыг тус тус байгуул.

Бид үүнээс өмнөх жишээнүүдэд

$\overline{X} = 32.1,\quad S_x^2 = 21.84,\quad \overline{y}_x = 0.6762 \cdot x - 4.79,\quad \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (\overline{y}_{x_i} - y_i)^2 = 502.0$

гэсэн үр дүнгүүдийг нэгэнт гарган авсан билээ. Одоо үлдэгдэл дисперсийг олъё. $S^2_{\text{үлд}} = \dfrac{502.0}{50-2} = 10.46$. (13.46) томъёо ёсоор нөхцөлт дунджийн дисперсийг олбол:

$$S^2_{\overline{y}_0} = 10.46\left[\frac{1}{50} + \frac{(36-32)^2}{50 \cdot 21.84}\right] = 0.355, \quad S_{\overline{y}_{x_0}} = 0.596.$$

Регрессийн тэгшитгэл ёсоор $y_{x_0} = 0.6762 - 36 - 4.79 = 19.55$.
(13.48) томъёогоор нөхцөлт дунджийн итгэх завсрыг олбол:

$18.35 \leq M(Y/X = 36) \leq 20.75$, ($t_{0.05;48} = 2.01$) болно.

Өөрөөр хэлбэл, үйлдвэрлэлийн үндсэн фонд нь 36 (сая төгрөг) байх үйлдвэрүүдийн хоногт үйлдвэрлэх бүтээгдэхүүний дундаж хэмжээ 13.35÷20.75(т) байхыг 0.95 гэсэн итгэх магадлалтайгаар хэлж болно. Одоо тусгайлан сонгосон зөвхөн нэг үйлдвэрийн хувьд дээрх завсрыг байгуулъя. (y^*_{36})

$$S^2_{y^*_{36}} = 10.46\left[1 + \frac{1}{50} + \frac{(36-32)^2}{50 \cdot 21.84}\right] = 10.81, \quad S_{y^*_{36}} = 3.29.$$

Иймд $12.94 \leq y^*_{36} \leq 26.16$(т)
Эндээс үзвэл нийт үйлдвэрийн хувьд итгэх завсрын өргөн бага, харин зөвхөн нэг үйлдвэрийн хувьд итгэх завсар илүү өргөн байна. Энэ чанар ямар ч түүврийн хувьд биелэгдэх бөгөөд бодит байдалтай таарч буй нь ойлгомжтой.

13.7 Регрессийн тэгшитгэлийн илэрхийлэх чадварыг шалгах

Регрессийн тэгшитгэлийн илэрхийлэх чадварыг шалгана гэдэг нь X ба Y-ийн хамаарлыг илэрхийлж байгаа гарган авсан математик загвар туршилтын өгөгдлүүдтэй таарч буй эсэх, хамааран хувьсагчийг дүрслэхэд тэгшитгэлд оруулсан үл хамаарах хувьсагчид (нэг буюу хэд хэдэн) хангалттай эсэхийг тогтооно гэсэн үг юм. Үүнийг тогтоохдоо нормаль регрессийн нөхцөлд $H_0 : \beta_1 = 0$ гэсэн таамаглалыг шалгадаг. Энэ таамаглалыг дисперсийн шинжилгээний үндсэн дээр шалгана. Өөрөөр хэлбэл, хамааран хувьсагч Y-ийн $\overline{Y}$ дунджаасаа хазайх хазайлтын квадратуудын ерөнхий нийлбэр $Q-$г хоёр үл хамаарах $Q_1,\ Q_2$ нэмэгдэхүүнд задалдаг. ($Q = Q_1 + Q_2$)

$$\sum_{i=1}^{n}(y_i - \overline{Y})^2 = \sum_{i=1}^{n}(y_{x_i} - \overline{Y})^2 + \sum_{i=1}^{n}(y_i - y_{x_i})^2 \tag{13.51}$$

y_{x_i}-бүлгийн дунджууд.
$Q_1 = \sum\limits_{i=1}^{n}(y_{x_i} - \overline{Y})^2$ нийлбэр үл хамаарах хувьсагч X-ийн нөлөөг илэрхийлнэ.
$Q_2 = \sum\limits_{i=1}^{n}(y_i - y_{x_i})^2$ нь тооцогдоогүй үлдсэн санамсаргүй хүчин зүйлүүдийн нөлөөг илэрхийлсэн үлдэгдэл нийлбэр.
Иймд санамсаргүй хүчин зүйлийн нөлөө хичнээн бага байх тутам, математик загвар туршилтын өгөгдлүүдтэй төдийчинээн сайн таарна. H_0 таамаглалыг шалгахын тулд:

$$F = \frac{Q_1(n-2)}{Q_2} \tag{13.52}$$

шинжүүрийг ашиглана. Энэ шинжүүр $k_1 = 1, \quad k_2 = n - 2$ чөлөөний зэргүүд бүхий Фишерийн F-тархалттай байна. Хэрэв $F > F_{\alpha;k_1;k_2}$ нөхцөл биелэгдвэл H_0 таамаглалыг үгүйсгэж регрессийн тэгшитгэлийн илэрхийлэх чадварыг үнэмшилтэйд тооцно. $F_{\alpha;k_1;k_2}$-Фишерийн тархалтын таблицын утга.
Хэрэв регрессийн тэгшитгэл k параметрийг агуулсан байвал:

$$F = \frac{Q_1(n-k)}{Q_2(k-1)} \tag{13.53}$$

шинжүүрийг ашиглана. Энэ шинжүүр (13.52)-ын адил тархалттай байна.

Жишээ-13.5 Жишээ-12.2-т авч үзсэн хоногт үйлдвэрлэх бүтээгдэхүүний хэмжээ (Y) үйлдвэрлэлийн үндсэн фонд (X)-оос хамаарах регрессийн тэгшитгэлийн илэрхийлэх чадварыг шалга.

$$\begin{aligned}
&Q_1 = \sum_{i=1}^{n}(y_{x_i} - \overline{Y})^2 = 502.0\\
&Q = \sum_{i=1}^{n}(y_i - \overline{Y})^2 = \sum_{i=1}^{n} y_i^2 - \frac{1}{n}\cdot\left(\sum_{i=1}^{n} y_i\right)^2 = 911.68\\
&Q_2 = Q - Q_1 = 911.68 - 502.0 = 409.68
\end{aligned}$$

(13.52) шинжүүрийн утгыг бодвол $F = \dfrac{502(50-2)}{409.68} = 58.8$ Итгэх түвшин $\alpha = 0.05$-ын хувьд $F_{0.05;48} = 4.04$ ба $F > F_{0.05;1;48}$ учир Жишээ-13.1-д гарган авсан регрессийн тэгшитгэл илэрхийлэх чадвартай. Өөрөөр хэлбэл, 0.95 гэсэн үнэний хувьтайгаар бодит байдлыг илэрхийлж чадна гэсэн үг.

13.8 Олон хэмжээст регрессийн шинжилгээ

Нэгэн зэрэг хамтран үйлчлэх олон тооны хүчин зүйлийн нөлөөн дор явагдах үзэгдэл процессын хамаарлыг судлах шаардлага гардаг. Ийм учраас олон хувьсагчийн регрессийн шинжилгээнд, хамааран хувьсагч Y нь $X_1, X_2, \ldots, X_q$ гэсэн хэд хэдэн үл хамаарах хувьсагчдаас хамаарах регресс хамаарлыг судална.
i-р ажиглалтын хамааран хувьсагчийн утгыг y_i, үл хамаарах хувьсагчдын утгыг $x_{i1}, x_{i2}, \ldots, x_{iq}$ гэе. Тэгвэл олон хэмжээст шугаман регрессийн загварыг дараах хэлбэртэй дүрсэлж болно.

$$y_i = \beta_0 x_{i0} + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_q x_{iq} + \varepsilon_i \tag{13.54}$$

Үүнд: $i = \overline{1,q}$, $x_{i0} = 1$-хуурмаг хувьсагч, ε_i-13.1-д авч үзсэн регрессийн шинжилгээний урьдчилсан нөхцлүүдийг хангах санамсаргүй хэмжигдэхүүн (зөрөө).

Регрессийн загварт олон үл хамаарах хувьсагчид оролцох учраас томьёо болон бодолтуудын бичиглэл нүсэр түвэгтэй болдог. Иймд вектор ба матрицан алгебрийн бичиглэлийг хэрэглэх нь тохиромжтой. Дараах матрицан тэмдэглэгээ оруулъя:

$$\boldsymbol{X}=\begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1q}\\ 1 & x_{21} & x_{22} & \dots & x_{2q}\\ \dots & \dots & \dots & \dots & \dots\\ 1 & x_{n1} & x_{n2} & \dots & x_{nq}\end{pmatrix}, \qquad \begin{matrix}\boldsymbol{\beta}=(\beta_0,\beta_1,\dots,\beta_q)^T,\\ \boldsymbol{\varepsilon}=(\varepsilon_1,\varepsilon_2,\dots,\varepsilon_n)^T.\end{matrix}$$

$\boldsymbol{X}$ −үл хамаарах хувьсагчдын $n\cdot(q+1)$ хэмжээст матриц,
$\boldsymbol{Y}$ −ажиглалтын багана матриц,
$\boldsymbol{\beta}$−үл мэдэгдэх параметрүүдийн багана матриц,
$\boldsymbol{\varepsilon}$−санамсаргүй алдаануудын (зөрөө) вектор багана.
(13.54) загварыг матриц хэлбэрээр бичвэл:

$$\boldsymbol{Y}=\boldsymbol{X}\cdot\boldsymbol{\beta}+\boldsymbol{\varepsilon} \tag{13.55}$$

Энэ загварын түүврийн үнэлэлт нь дараах хэлбэртэй болно.

$$\boldsymbol{Y}=\boldsymbol{X}\cdot\boldsymbol{b}+\boldsymbol{e} \tag{13.56}$$

үүнд: $\boldsymbol{b}=(b_0,b_1,\dots,b_q)^T,\quad \boldsymbol{e}=(e_1,e_2,\dots,e_n)^T$
Үл мэдэгдэх параметруудын вектор $\boldsymbol{\beta}$-ийн үнэлэлтийг олохдоо хамгийн бага квадратын аргыг хэрэглэнэ.

$$\boldsymbol{e}^T\boldsymbol{e}=(e_1,e_2,\dots,e_n)\begin{pmatrix} e_1\\ e_2\\ \vdots\\ e_n\end{pmatrix}=e_1^2+e_2^2+\dots e_n^2=\sum_{i=1}^{n}e_i^2$$

адилтгалыг ашиглан үлдэгдэл нийлбэрийг минимальчилах нөхцөлийг бичвэл:

$$S=\sum_{i=1}^{n}(y_{x_i}-y_i)^2=\sum_{i=1}^{n}e_i^2=(\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{b})^T\cdot(\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{b})\longrightarrow\min \tag{13.57}$$

Нормаль систем дараах матрицан хэлбэрээр бичигдэнэ.

$$\boldsymbol{X}^T\cdot\boldsymbol{X}\cdot\boldsymbol{b}=\boldsymbol{X}^T\cdot\boldsymbol{Y} \tag{13.58}$$

Үүнд:

$$\boldsymbol{X}^T\boldsymbol{X}=\begin{pmatrix} n & \sum x_{i1} & \sum x_{i2} & \dots & \sum x_{iq}\\ \sum x_{i1} & \sum x_{i1}^2 & \sum x_{i1}x_{i2} & \dots & \sum x_{i1}x_{iq}\\ \dots & \dots & \dots & \dots & \dots\\ \sum x_{iq} & \sum x_{iq}x_{i1} & \sum x_{iq}x_{i2} & \dots & \sum x_{iq}^2\end{pmatrix} \tag{13.59}$$

$$\boldsymbol{X}^T\boldsymbol{Y}=\begin{pmatrix} 1 & 1 & \dots & 1\\ x_{11} & x_{21} & \dots & x_{n1}\\ \dots & \dots & \dots & \dots\\ x_{1q} & x_{2q} & \dots & x_{nq}\end{pmatrix}\cdot\begin{pmatrix} y_1\\ y_2\\ \dots\\ y_n\end{pmatrix}=\begin{pmatrix}\sum y_i\\ \sum y_i x_{i1}\\ \dots\\ \sum y_i x_{iq}\end{pmatrix} \tag{13.60}$$

(13.58) системийг координатан хэлбэрээр нь (13.30) томъёогоор, $q=2$, $q=1$ үед (13.31)÷(13.33) томъёогоор тус тус бичсэн билээ. (13.58) тэгшитгэлд $(\boldsymbol{X}^T \cdot \boldsymbol{X})$ матрицыг урвуутай гэсэн нөхцөлийг нэмж шаардвал түүний шийд

$$\boldsymbol{b} = (\boldsymbol{X}^T \cdot \boldsymbol{X})^{-1} \cdot \boldsymbol{X}^T \cdot \boldsymbol{Y} \tag{13.61}$$

хэлбэртэй болно.
Хамгийн бага квадратын аргаар олсон (13.61) үнэлэлт нь хазайлтгүй, зохимжтой, эрчимтэй байдаг. Эдгээр үнэлэлтүүд түүврийн утгуудаар бодогдох учраас санамсаргүй алдаа агуулдаг. Энэ нь, эцэстээ олон хэмжээст регрессийн тэгшитгэлийн нарийвчлалд нөлөөлнө. Олон хэмжээст регрессийн шинжилгээнд энэ алдааг үнэлэх зорилгоор ковариацийн матриц $\boldsymbol{K}$-г авч үздэг.

$$\boldsymbol{K} = \begin{pmatrix} k_{00} & k_{01} & \dots & k_{0q} \\ k_{10} & k_{11} & \dots & k_{1q} \\ \dots & \dots & \dots & \dots \\ k_{q0} & k_{q1} & \dots & k_{qq} \end{pmatrix} \tag{13.62}$$

Үүнд:

$$k_{ij} = M[(b_i - \beta_i) \cdot (b_j - \beta_j)] \tag{13.63}$$

Ковариацийн матрицыг хураангуй хэлбэрээр бичвэл

$$\boldsymbol{K} = M[(\boldsymbol{b} - \boldsymbol{\beta}) \cdot (\boldsymbol{b} - \boldsymbol{\beta})^T] \tag{13.64}$$

Ковариацийн матрицын диагоналын элементүүд нь регрессийн параметрүүдын үнэлэлтийн дисперс, диагоналын биш элементүүд нь параметр хоорондын статистик хамаарлыг илэрхийлсэн харгалзах элементүүдийн ковариац байна. Регрессийн шинжилгээний урьдчилсан нөхцөлүүдийн 3; 4 чанаруудыг тооцвол:(13.1-ийг үз)

$$\boldsymbol{K} = (\boldsymbol{X}^T \cdot \boldsymbol{X}) \cdot \sigma_\varepsilon^2 \tag{13.65}$$

σ_ε^2-санамсаргүй алдааны дисперс.
σ_ε^2-дисперсийн үнэлэлтийг бичвэл:

$$S^2_{\text{үлд}} = \frac{\sum\limits_{i=1}^{n} e_i^2}{n-(q+1)} = \frac{\boldsymbol{e}^T \cdot \boldsymbol{e}}{n-q-1}. \tag{13.66}$$

Регрессийн b_j-коэффициентуудын дисперсийг дараах томъёогоор олно.

$$S^2_{b_j} = S^2_{\text{үлд}} \cdot C_{jj} \tag{13.67}$$

Үүнд: $C_{jj} - (\boldsymbol{X}^T \cdot \boldsymbol{X})^{-1}$ матрицын диагоналын элемент.
b_j-коэффициентийн стандарт хазайлт:

$$S_{b_j} = S_{\text{үлд}} \cdot \sqrt{C_{jj}} \tag{13.68}$$

Регрессийн коэффициентийн итгэлтэй чанарыг $k = n - q - 1$ чөлөөний зэрэг бүхий Стьюдентийн тархалттай дараах шинжүүрээр шалгана.

$$t = \frac{|b_j|}{S_{\text{үлд}}\sqrt{C_{jj}}} \tag{13.69}$$

Хэрэв өгөгдсөн итгэх түвшин α-ийн хувьд $t > t_{\alpha;k}$ нөхцөл биелэгдвэл b_j-г итгэлтэй гэнэ.
(13.68)-аас β_j-коэффициентийн итгэх завсар мөрдөнө:

$$b_j - t_{\alpha;k} \cdot S_{b_j} \leq \beta_j \leq b_j + t_{\alpha;k} \cdot S_{b_j} \tag{13.70}$$

Практикт, үл хамаарах хувьсагчид нь хэмжилтийн янз бүрийн нэгжээр илэрхийлэгдсэн байх тохиолдолд тэдгээрийн Y хувьсагчид нөлөөлөх нөлөөг тооцохдоо регрессийн стандартчилагдсан коэффициент(стандартизованный коэффициент) b'_j ба мэдрэмжийн коэффициент (коэффициент эластичности) ϑ_j-г хэрэглэдэг.

$$b'_j = b_j \cdot \frac{S_{x_j}}{S_y} \tag{13.71}$$

$$\vartheta_j = b_j \cdot \frac{\overline{x}_j}{\overline{Y}} \tag{13.72}$$

Регрессийн стандартчилагдсан коэффициент нь үл хамаарах j-р хувьсагчийн утга S_{x_j} хэмжээгээр нэмэгдэхэд хамааран хувьсагч Y дунджаар хичнээн (S_y) хэмжээгээр өөрчлөгдөхийг илэрхийлнэ. Харин мэдрэмжийн коэффициент ϑ_j нь зөвхөн X_j 1 процентоор нэмэгдэхэд Y дунджаар хэдэн процентоор өөрчлөгдөхийг (дунджаасаа) илэрхийлнэ.

Хамааран хувьсагчийн тухай урьдчилан дүгнэлт гаргахад түүний итгэх завсрыг байгуулах шаардлагатай. Хос регрессийн тэгшитгэлд гаргасны адилаар, үл хамаарах хувьсагч нь $\boldsymbol{X}_0^T = (1, x_1, x_2, \ldots, x_q)$ утга авч байх үед $M_{x_0}(\boldsymbol{Y})$-нөхцөлт математик дунджиийн (регрессийн функцийн) итгэх завсар дараах хэлбэртэй байна.

$$\boldsymbol{y}_{x_0} - S_{\boldsymbol{y}_{x_0}} \cdot t_{\alpha,k} \leq M_{x_0}(\boldsymbol{Y}) \leq \boldsymbol{y}_{x_0} + S_{\boldsymbol{y}_{x_0}} \cdot t_{\alpha,k} \tag{13.73}$$

Үүнд: $\boldsymbol{Y}_{x_0} = \boldsymbol{X}_0^T \cdot \boldsymbol{b}$, $k = n - q - 1$ чөлөөний зэрэг, $S_{\boldsymbol{y}_{x_0}}$ нь $\boldsymbol{y}_{x_0}$ утгын стандарт алдаа.

$$S_{\boldsymbol{y}_{x_0}} = S_{\text{үлд}}\sqrt{\boldsymbol{X}_0^T \cdot (\boldsymbol{X}^T \cdot \boldsymbol{X})^{-1} \cdot \boldsymbol{X}_0} \tag{13.74}$$

Хамааран хувьсагчийн ганцаарчилсан утгын итгэх завсар буюу (13.50) томъёоны өргөтгөл нь:

$$\boldsymbol{y}_{x_0} - S_{\boldsymbol{y}_{x_0}} \cdot t_{\alpha,k} \leq \boldsymbol{y}_0^* \leq \boldsymbol{y}_{x_0} + S_{\boldsymbol{y}_{x_0}} \cdot t_{\alpha,k} \tag{13.75}$$

$$S_{\boldsymbol{y}_0} = S_{\text{үлд}}\sqrt{1 + \boldsymbol{X}_0^T \cdot (\boldsymbol{X}^T \cdot \boldsymbol{X})^{-1} \cdot \boldsymbol{X}_0} \tag{13.76}$$

хэлбэртэй болно.

Олон хэмжээст регрессийн тэгшитгэлийн илэрхийлэх чадварыг шалгахдаа Фишерийн F-тархалт бүхий дараах шинжүүрийг ашиглана.

$$F = \frac{Q_1(n-q-1)}{Q_2 \cdot q} \tag{13.77}$$

$Q_1,\ Q_2$ нь (13.51) томъёогоор тодорхойлогдох ба $F > F_{\alpha;k_1;k_2}$ $(k_1 = q,\ k_2 = n{-}q{-}1)$ нөхцөл биелэгдэж байвал регрессийн тэгшитгэлийн илэрхийлэх чадварыг үнэмшилтэйд тооцно.

Жишээ-13.6 Дараах хүснэгтэнд 20 уурхайн нүүрс олборлох процессын зарим үзүүлэлтүүд өгчээ. Үүнд: давхаргын зузаан-X_1 (м); ажлын механикжилтын түвшин-X_2 (%); ээлжийн хугацаанд нэг ажилчины олборлох нүүрсний хэмжээ-Y (т).

i	x_{i1}	x_{i2}	y_i	i	x_{i1}	x_{i2}	y_i
1	76	49	5.0	11	84	84	5.5
2	108	82	10.3	12	94	64	6.1
3	116	85	9.7	13	113	81	8.7
4	98	83	7.5	14	141	100	11.6
5	91	47	6.5	15	88	37	4.5
6	125	100	9.7	16	98	82	9.6
7	85	35	6.2	17	82	49	5.5
8	113	82	10.9	18	83	37	6.8
9	76	67	5.2	19	115	67	7.8
10	101	65	7.0	20	105	84	10.5

а) $\boldsymbol{Y} = \boldsymbol{Y}(X_1, X_2)$ хамаарлын аналитик илэрхийллийг ол.
б) Регрессийн коэффициентуудын итгэлтэй чанар ба регрессийн тэгшитгэлийн илэрхийлэх чадварыг шалга.
в) Давхаргын зузаан нь 1000 м, механикжилтын түвшин нь 60% байх уурхайнуудын хувьд 1 ажилчны ээлжиндээ олборлох нүүрсний хэмжээний дундаж ба ганцаарчилсан утгын 95%-ийн итгэх завсрыг тус тус байгуул.
Дараах матрицан тэмдэглэгээг оруулъя.

$$\begin{pmatrix} 1 & 76 & 49 \\ 1 & 108 & 82 \\ 1 & 116 & 85 \\ \dots & \dots & \dots \\ 1 & 115 & 67 \\ 1 & 105 & 84 \end{pmatrix}, \quad \boldsymbol{Y} = \begin{pmatrix} 5.0 \\ 10.3 \\ 9.7 \\ \vdots \\ 7.8 \\ 10.5 \end{pmatrix}. \quad \text{Тэгвэл:}$$

$$\boldsymbol{X}^T \cdot \boldsymbol{X} = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ 76 & 108 & 116 & \dots & 115 & 105 \\ 49 & 82 & 85 & \dots & 67 & 84 \end{pmatrix} \cdot \begin{pmatrix} 1 & 76 & 49 \\ 1 & 108 & 82 \\ 1 & 116 & 85 \\ \dots & \dots & \dots \\ 1 & 115 & 67 \\ 1 & 105 & 84 \end{pmatrix} =$$

$$\begin{pmatrix} 20 & 1992 & 1380 \\ 1992 & 204130 & 142480 \\ 1380 & 142480 & 103232 \end{pmatrix},$$

$$\boldsymbol{X}^T \cdot \boldsymbol{Y} = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ 76 & 108 & 116 & \dots & 115 & 105 \\ 49 & 82 & 85 & \dots & 67 & 84 \end{pmatrix} \cdot \begin{pmatrix} 5.0 \\ 10.3 \\ 9.7 \\ \vdots \\ 7.8 \\ 10.5 \end{pmatrix} = \begin{pmatrix} 154.6 \\ 16023.3 \\ 11322.8 \end{pmatrix}.$$

$det(\boldsymbol{X}^T\boldsymbol{X}) = 411214952$ учир $(\boldsymbol{X}^T\boldsymbol{X})^{-1}$ матриц оршино.

$$(\boldsymbol{X}^T\boldsymbol{X})^{-1} = \begin{pmatrix} 1.877 & -2.192 \cdot 10^{-2} & 5.157 \cdot 10^{-3} \\ -2.192 \cdot 10^{-2} & 3.896 \cdot 10^{-4} & -2.447 \cdot 10^{-4} \\ 5.157 \cdot 10^{-3} & -2.447 \cdot 10^{-4} & 2.785 \cdot 10^{-4} \end{pmatrix}.$$

(13.61) ёсоор: $\boldsymbol{b} = \begin{pmatrix} -2.595 \\ 0.08319 \\ 0.02955 \end{pmatrix}.$

Тэгвэл олон хэмжээст регрессийн тэгшитгэл

$$y = -2.595 + 0.08319 \cdot x_1 + 0.02955 \cdot x_2 \quad \text{хэлбэртэй болно.}$$

Энэ тэгшитгэл, зөвхөн давхаргын зузаан 1 м-ээр нэмэгдэхэд нэг ажилчны ээлжиндээ олборлох нүүрсний хэмжээ дунджаар 0.08319 тонноор, зөвхөн механикжилтын түвшин 1 процентоор нэмэгдэхэд олборлолт 0.02955 тонноор нэмэгдэхийг тус тус үзүүлнэ.

(13.71) ба (13.72) томъёогоор регрессийн стандартчилагдсан коэффициентүүд ба мэдрэмжийн коэффициентийг олъё.

$\overline{X}_1 = 99.6, \quad \overline{X}_2 = 69.0, \quad y = 7.73, \quad S_{x_1} = 75.68, \quad S_{x_2} = 89.51, \quad S_y = 9.68,$

$b'_1 = 0.08319\dfrac{75.68}{9.68} = 0.65, \quad b'_2 = 0.02955\dfrac{89.51}{9.68} = 0.273$

$$\vartheta_1 = 0.08319\frac{99.6}{7.73} = 1.071, \quad \vartheta_2 = 0.02955\frac{69.0}{7.73} = 0.264.$$

Эдгээр коэффициентүүдийн утгыг тайлбарлавал: давхаргын зузаан, механикжилтын түвшин харгалзан нэг нэг S_{x_1}, S_{x_2} нэгжээр нэмэгдэхэд нэг ажилчны ээлжиндээ олборлох нүүрсний дундаж хэмжээ харгалзан $0.65 \cdot S_y$, $0.273 \cdot S_y$ хэмжээгээр, харин дээрх хоёр үзүүлэлт (дундаж утгаасаа) 1% нэмэгдэхэд нүүрс олборлолт 1.071% ба 0.264%-иар тус тус нэмэгдэнэ гэсэн үг юм. Эндээс үзвэл "давхаргын зузаан" гэсэн үзүүлэлт "механикжилтын түвшин" гэсэн үзүүлэлтийг бодвол нүүрс олборлох хэмжээнд илүү нөлөө үзүүлж байна.

b_0, b_1, b_2 параметрүүдийг олохдоо $(\boldsymbol{X}^T\boldsymbol{X})^{-1}$ матрицын урвууг олохгүйгээр дараах системээс олж болохыг тэмдэглэе.

$$\left.\begin{array}{l} 20b_0 + 1992b_1 + 1380b_2 = 154.6 \\ 1992b_0 + 204130b_1 + 142480b_2 = 16023.3 \\ 1380b_0 + 142480b_1 + 103232b_2 = 11322.8 \end{array}\right\}$$

б) Дараах туслах таблицыг зохиож тооцоо хийе.

i	x_{i1}	x_{i2}	y_i	$(y_i - \overline{Y})^2$	y_{x_i}	$e_i^2 = (y_{x_i} - y_i)^2$
1	76	49	5.0	7.4529	5.17	0.0308
2	108	82	10.3	6.6049	8.81	2.2123
3	116	85	9.7	3.8809	9.57	0.0177
...	...	...	...	...	...	...
19	115	67	7.8	0.0049	8.95	1.3264
20	105	84	10.5	7.6729	8.62	3.5263
$\sum$	-	-	-	93.7021	-	22.3947

Эхлээд регрессийн тэгшитгэлийн илэрхийлэх чадварыг шалгая.
Таблицаас: $Q=93.7021$, $Q_2=22.3947$, $Q_1=Q-Q_2=93.7021-22.3947=71.3074$ (13.77) ёсоор:

$$F = \frac{71.3074(20-2-1)}{22.3947.2} = 27.1, \quad F_{0.05;2;17} = 3.59, \quad F > F_{0.05;2;17}$$

учир гарган авсан регрессийн тэгшитгэл судалж буй процессийг үнэмшилтэйгээр илэрхийлж байна.
Одоо регрессийн b_1, b_2 коэффициентуудын итгэлтэй чанарыг шалгая. (13.66) ёсоор:

$$S^2_{\text{үлд}} = \frac{22.3947}{20-2-1} = 1.399, \quad S_{\text{үлд}} = 1.182,$$

$C_{11} = 3.896 \cdot 10^{-4}$, $C_{22} = 2.785 \cdot 10^{-4}$ тул (13.69) томъёог ашиглавал:

$$t_{b_1} = \frac{0.08319}{1.182\sqrt{3.896 \cdot 10^{-4}}} = 3.57\ , \quad t_{b_2} = \frac{0.02955}{1.182\sqrt{2.785 \cdot 10^{-4}}} = 1.5$$

$t_{0.05;17} = 2.11$. Иймд 5%-ын итгэх түвшний хувьд b_2-коэффициент итгэлтэй биш гарч байна. Энэ нь өгөгдсөн түүврийн хувьд нүүрс олборлолтын хэмжээнд механикжилтын түвшинг мэдрэхүйц нөлөөтэй гэж тооцох үндэсгүй гэсэн үг.
Итгэлтэй биш коэффициентийг регрессийн тэгшитгэлээс хасаж регрессийн параметруудыг цөөн тооны хувьсагчдын хувьд дахин боддог. Манай тохиолдолд X_2 хувьсагчийг хасаж дахин тооцоо хийвэл регрессийн тэгшитгэл

$y = -3.142 + 0.1091x_1$ хэлбэртэй болно. Нөлөө багатай хувьсагчийг хасахдаа үндэслэл сайтай чанарын судалгаан дээр тулгуурлавал зохино. Нөлөө багатай ч тодорхой үндэслэлтэйгээр түүнийг регрессийн тэгшитгэлд үлдээж болно.

в) $x_1 = 100$, $x_2 = 60$ үед (өөрөөр хэлбэл, $\boldsymbol{X}_0^T = (1,\ 100,\ 60)$) хамааран хувьсагчийн утга y_{x_0}-ийг бодъё.

$y_{x_0} = -2.595 + 0.008319 \cdot 100 + 0.02955 \cdot 60 = 7,497$ (т). (13.74) ёсоор:

$$\boldsymbol{X}_0^T(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}_0 = (1, 100, 60)\cdot$$

$$\cdot \begin{pmatrix} 1.877 & -2.192 \cdot 10^{-2} & 5.157 \cdot 10^{-3} \\ -2.192 \cdot 10^{-2} & 3.896 \cdot 10^{-4} & -2.447 \cdot 10^{-4} \\ 5.157 \cdot 10^{-3} & -2.447 \cdot 10^{-4} & 2.785 \cdot 10^{-4} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 100 \\ 60 \end{pmatrix} = 0.07404$$

$S_{\text{үлд}} = 1.82\sqrt{0.07404} = 0.322$ (т). (13.73) томъёогоор итгэх завсрыг байгуулбал:

$7.497 - 2.11 \cdot 0.322 \leq M_{x_0}(\boldsymbol{Y}) \leq 7.497 + 2.11 \cdot 0.322$ буюу

$$6.82 \leq M_{x_0}(\boldsymbol{Y}) \leq 8.18 \text{ (т)}$$

(13.76), (13.75) томъёогоор ганцаарчилсан утгын итгэх завсрыг байгуулбал:

$$S_{y_0} = 1.182\sqrt{1 + 0.7404} = 1.225 \quad \text{(т)}$$

$$4.92 \leq y_0^* \leq 10.08 \quad \text{(т)}.$$

Бүлэг 14

Паскаль программууд

Энэ бүлэгт санамсаргүй хэмжигдэхүүний тоон үзүүлэлтүүд, зарим магадлалын утга, корреляцийн коэффициент, шугаман ба шугаман биш регрессийн коэффициент зэрэг параметрүүдийг хэрхэн олохыг хялбар жишээн дээр үзүүлж тэдгээрийг бодох программыг **паскаль** хэл дээр бичиж үзүүлсэн болно.

14.1 Математик дундаж (дискрет тохиолдол)

$$M(\xi) = \sum_{i=1}^{N} x_i p_i.$$

Жишээ-14.1 $N = 4$

N	1	2	3	4
x_i	0	1	2	3
p_i	0.2	0.4	0.3	0.1

Үр дүн: $M(\xi) = 1.3$

```
Program "МАТЕМАТИК ДУНДАЖ";
        Const w = 50;
        Type Mat=array[1..w] of real;
        Var n,i: integer;
           x, p:Mat;sum:real;
   Begin
       Write('n =');Read(n);
       For i:=1 To n Do
            begin
               Write('x(',i,')=');Read(x[i]);
               Write('p(',i,')=');Read(p[i])
            end;
       sum:=0;
```

```
    For i:=1 To n Do sum:=sum+x[i]*p[i];
    Writeln('Mat Dy=',sum:5:3)
End.
```

14.2 Дисперс (дискрет тохиолдол)

$$D(\xi) = \sum_{i=1}^{N} (x_i - M(\xi))^2 p_i.$$

Жишээ-14.2 $N = 4$

N	1	2	3	4
x_i	0	1	2	3
p_i	0.2	0.4	0.3	0.1

Үр дүн: $D(\xi) = 0.81$

Үр дүн: $D(\xi) = 0.81$

```
Program "ДИСПЕРС";
    Const w=50;
    Type Mat=array[1..w] of real;
    Var x,p:Mat;
        n,i: integer;
       a,sum:real;
   Function M(x,p:Mat):real;
     Var s:real;
   begin
    s:=0;
     for i:=1 to n do s:=s+x[i]*p[i];
     M:=s
    End;
Begin
    Write('n =');Read(n);
    For i:=1 To n Do
       begin
         Write('x(',i,')=');Read(x[i]);
         Write('p(',i,')=');Read(p[i])
       end;
    a:=M(x,p);sum:=0;
    For i:=1 To n Do sum:=sum+sqr(x[i]-a)*p[i];
    Writeln('Disperc=',sum:7:3)
End.
```

14.3 к-эрэмбийн анхны момент (дискрет тохиолдол)

$$\nu_k = \sum_{i=1}^{N} x_i^k p_i.$$

Жишээ-14.3 $N = 4, \quad k = 2.$

N	1	2	3	4
x_i	0	1	2	3
p_i	0.2	0.4	0.3	0.1

Үр дүн: $\nu_k = 2.5$

```
Program "АНХНЫ МОМЕНТ";
        Const w=50;
        Type Mat=array[1..w] of real;
        Var k,n,i,j: integer;
             f, sum: real;
              x,p:Mat;
    Begin
         Write('erembe:k=');Read(k);
         Write('n=');Read(n);
        For i:=1 To n Do
           begin
             Write('x(',i,')=');Read(x[i]);
             Write('p(',i,')=');Read(p[i])
           end;
        sum:=0;
        For i:=1 To n Do
           begin
             f:=1;
             for j:=1 to k do f:=f*x[i];
             sum:=sum+f*p[i]
           end;
        Writeln('First- moment=',sum:7:3)
    End.
```

14.4 к-эрэмбийн төвийн момент (дискрет тохиолдол)

$$\mu_k = \sum_{i=1}^{N} (x_i - M(\xi))^k p_i.$$

Жишээ-14.4 $N = 4, \quad k = 3$

N	1	2	3	4
x_i	0	1	2	3
p_i	0.2	0.4	0.3	0.1

Үр дүн: $\mu_k = 0.144$

Үр дүн: $\mu_k = 0.144$

```
Program "ТӨВИЙН МОМЕНТ";
        Const w=50;
        Type Mat=array[1..w] of real;
        Var k,n,i,j:integer;
              x,p:Mat;
        f, a,sum:real;
      Function M(x,p:Mat):real;
        Var s:real;
      Begin
        s:=0;
         for i:=1 to n do s:=s+x[i]*p[i];
         M:=s;
      End;
Begin
        Write('erembe:k=');Read(k);
        Write('n=');Read(n);
        For i: =1 To n Do
           begin
             Write('x(',i,')=');Read(x[i]);
             Write('p(',i,')=');Read(p[i])
           end;
        a:=M(x,p);
        sum:=0;
        For i:=1 To n Do
           begin
             f:=1;
             for j:=1 to k do f:=f*(x[i]−a);
             sum:=sum+f*p[i]
           end;
         Writeln('Central moment=', sum:7:3)
    End.
```

14.5 Корреляцийн коэффициент

$$R = \frac{\sum_{i=1}^{N} x_i y_i - \left(\sum_{i=1}^{N} x_i \cdot \sum_{i=1}^{N} y_i\right)/N}{\sqrt{\sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i\right)^2/N} \cdot \sqrt{\sum_{i=1}^{N} y_i^2 - \left(\sum_{i=1}^{N} y_i\right)^2/N}}$$

Жишээ-14.5 $N = 5$

N	1	2	3	4	5
x_i	0.95	2.1	3	4.1	4.9
y_i	2	4.05	5.8	8.1	9.2

Үр дүн: $R = 0.9987280191$

```
Program "КОРРЕЛЯЦИЙН КОЭТТИЦИЕНТ";
     Const w=50;
     Type Mat=array[1..w] of real;
     Var x,y:Mat;
       i,j,N:integer;
      r,rx,ry:real;
    Function DS(x,y:Mat):real;
     Var s:real;
    begin s:=0;
     for i:=1 to N do s:=s+x[i]*y[i];
     DS:=s
    End;
    Function SS(x:Mat):real;
     Var s:real;
    begin
     s:=0;
     for i: = l to N do s:=s+x[i];
     SS:=s
    End;
  Begin
    Write('N=');Read(N);
    For i:=1 To N Do
     begin
       Write('x(',i,')=');Read(x[i]);
       Write('y(',i,')=');Read(y[i])
     end;
    rx:=sqrt(DS(x,x)-sqr(SS(x))/N);
    ry:=sqrt(DS(y,y)-sqr(SS(y))/N);
    R:=(DS(x,y)-(SS(x)*SS(y))/N)/(rx*ry);
    Writeln('R=',r:7:10)
  End.
```

14.6 к-эрэмбийн төвийн момент (тасралтгүй тохиолдол)

$$\mu_k = \int\limits_a^b (x - M(\xi))^k f(x)dx$$

Жишээ-14.6 $k = 2, \quad a = 0, \quad b = 1, \quad f(x) = \dfrac{1}{b-a}.$

$$\mu_2 = \int\limits_a^b (x - M(\xi))^2 \frac{dx}{b-a} = \frac{(b-a)^2}{12} \ , \quad \text{Үр дүн:} \ \ \mu_2 = D(\xi) = 0.08333$$

```
Program "ТӨВИЙН МОМЕНТ";
    Const e=0.0001;
    Var al,be,M:real;
        n,k,T,L: integer;
   Function F(x:real):real;
    Var q:real;
        j: integer;
   Begin
    If T > 0 Then F:=x/(be-al)
         Else
           begin
             q:=1;
             For j:=1 To k Do q:=q*(x-M);
             F:=q/(be-al)
           end;
   End;
   Function Integ(n:integer):real;
    Var J1,J2:real;
    Function Cim(n:integer):real;
      Var h,s,s1,s2,x0,ff:real;
                        i:integer;
   Begin
     h:=(be-al)/(2*n);
     sl:=0; s2:=0; s:=0;
     For i:=0 to 2*n do
        begin
          x0:=al+i*h;ff:=f(x0);
          If (i <> 0)and(i <> 2*n)
            Then If odd(i)
```

```
                Then s2:=s2+ff
                Else sl:=sl+ff;
            s:=((2*sl+4*s2+f(al)+f(be))*h)/3;
            Cim:=s
          end;
    End;
   Begin
      Repeat
      Jl:=Cim(n);
      n:=n*2;
      J2:=Cim(n)
      Until abs(J1−J2)<e;
      Integ:=J1
    End;
 Begin
       Write('al be=');Read(al,be);
       Write('xybaalt:n=');Read(n);L:=n;
       Write('erembe :k=');Read(k);
       T:=1;
       M:=Integ(n);
       T:=-T;
       Writeln(Integ(L):20:6);
 End.
```

14.7 к-эрэмбийн анхны момент (тасралтгүй тохиолдол)

$$\nu_k = \int\limits_a^b x^k f(x)dx$$

Жишээ-14.7 $k = l, \quad a = 0, \quad b = 1, \quad f(x) = \dfrac{1}{b-a}.$

$$\nu_k = \int\limits_a^b x\frac{dx}{b-a} = \frac{a+b}{2}\ , \quad \text{Үр дүн:}\ \ \nu_1 = M(\xi) = 0.5$$

Program "**АНХНЫ МОМЕНТ**";

```
      Const e=0.0001;
      Var a,b:real;
          n,k:integer;
```

```
Function F(x:real):real;
  Var q:real;
      j:integer;
Begin
  q:=1;
  For j:=1 To k Do q:=q*x;
  F:=q/(b-a);
End;
Function Integ(n:integer):real;
  Var J1,J2:real;
 Function Cim(n:integer):real;
   Var h,s,sl,s2,x0,ff:real;
                  i:integer;
 Begin
   h:=(b-a)/(2*n);
   s1:=0;s2:=0;s:=0;
   For i:=0 to 2*n do
     begin
       x0:=a+i*h;ff:=f(x0);
       If (i<>0)and(i<>2*n)
         Then If odd(i)
           Then s2:=s2+ff
           Else s1:=s1+ff;
       s:=((2*s1+4*s2+f(a)+f(b))*h)/3;
       Cim:=s
     end;
 End;
Begin
  Repeat
  J1:=Cim(n);
  n:=n*2;
  J2:=Cim(n)
  Until abs(J1−J2)<e;
  Integ:=J1
End;
Begin
  Write('a b=');Read(a,b);
  Write('erembe:k=');Read(k);
  Write('xybaalt:n=');Read(n);
  Writeln(Integ(n):20:5);
End.
```

14.8 Санамсаргүй хэмжигдэхүүн өгөгдсөн завсарт утгаа авах магадлал

$$P(\alpha \le \xi \le \beta) = \int_{\alpha}^{\beta} f(x)dx.$$

Жишээ-14.8 $f(x) = \dfrac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-(x-a)^2/2\sigma^2}$

$\sigma = 0.1$, $\alpha = 6.7$, $a = 7$, $\beta = 7.3$, Үр дүн: 0.9973

```
Program "МАГАДЛАЛ";
     Const e=0.0001;
     Var al,be,a,sg:real;
            n:integer;
   Function F(x:real):real;
   Begin
     F:=exp((-sqr(x-a)/(2*sqr(sg))))/(sqrt(2*pi)*sg)
   End;
   Function Integ(n:integer):real;
     Var J1,J2:real;
    Function Cim(n:integer):real;
      Var h,s,s1,s2,x0,ff:real;
                 i:integer;
    Begin
     h:=(be-al)/(2*n);
     s1:=0;s2:=0;s:=0;
     For i:=0 to 2*n do
      begin
       x0:=a1+ i*h;ff:=f(x0);
       If (i<>0)and(i<>2*n)
        Then If odd(i)
         Then s2:=s2+ff
         Else s1:=s1+ff;
         s:=((2*s1+4*s2+f(al)+f(be))*h)/3;
         Cim:=s
      end;
     End;
   Begin
    Repeat J1:=Cim(n);
    n:=n*2;
    J2:=Cim(n)
    Until abs(J1-J2)<e;
```

```
    Integ:=J1
  End;
Begin
   Write('al be=');Read(al,be);
   Write('a,sg=');Read(a,sg);
   Write('n=');Read(n);
   Writeln(Integ(n):20:5);
End.
```

14.9 Шугаман регрессийн коэффициент

$$y(x) = b_1 x + b_0$$

$$b_1 = \frac{\sum_{i=1}^{N} x_i \cdot \sum_{i=1}^{N} y_i - N \cdot \sum_{i=1}^{N} x_i y_i}{\left(\sum_{i=1}^{N} x_i\right)^2 - N \cdot \sum_{i=1}^{N} x_i^2}$$

$$b_0 = \frac{1}{N}\left(\sum_{i=1}^{N} y_i - b_1 \cdot \sum_{i=1}^{N} x_i\right)$$

Жишээ-14.9 $N = 5$

N	1	2	3	4	5
x_i	2	4	6	8	10
y_i	5.5	6.3	7.2	8	8.6

Үр дүн: $b_0 = 4.75, \quad b_1 = 0.395$

```
Program "ШУГАМАН РЕГРЕСС";
     Uses Graph;
     Const w=50;
     Var A,B,C,D,B0,B1:Real;
            N,i:Integer;
     Gd,Gm,XX,YY,X0,Y0:integer;
          SX,SY,J:Real;
            x,y:array[1..w] of real;
   Function F(x:real):real;
   begin
     F:=B0+B1*x
   end;
 Begin
    A:=0;B:=0;C:=0;D:=0;
    Write('N=');Read(N);
    For i:=1 To N Do
```

```
    begin
      Write('x(',i,')=');Read(x[i]);
      Write('y(',i,')=');Read(y[i]);
      A:=A+x[i];B:=B+y[i];C:=C+sqr(x[i]);
      D:=D+x[i]*y[i];
    end;
  B1:=(A*B-N*D)/(sqr(A)-N*C);
  B0:=(B-B1*A)/N;
  Gd:=0;
  InitGraph(Gd,Gm,'');
  ClearDevice;
  X0:=100;Y0:=350;
  Line(X0,16,X0,199);
  Line(0,Round(Y0/2),600,Round(Y0/2));
  OutTextXY(90,16,'Y');
  OutTextXY(590,Round(Y0/2-10),'X');
  OutTextXY(150,20, 'Shygaman Regrecc');
  XX:=400; YY:=400;
  SX:=XX/(600-X0);SY:=YY/(Y0-16);
  For i:=1 to N Do
    begin
      PutPixel(Round(x[i]/sx+X0),Round(Y0/2-y[i]/SY/2),1);
      GraphDefaults;
    end;
  For I:=1 To XX Do
    begin
      J:=F(I);
      PutPixel(Round(I/SX+X0),Round(Y0/2-J/SY/2),1);
      GraphDefaults
     end;
End.
```

14.10 Гипербол регрессийн коэффициент

$$y(x) = b_0 + \frac{b_1}{x}$$

$$b_1 = \frac{\sum\limits_{i=1}^{N} \frac{1}{x_i} \cdot \sum\limits_{i=1}^{N} y_i - N \cdot \sum\limits_{i=1}^{N} \frac{y_i}{x_i}}{\left(\sum\limits_{i=1}^{N} \frac{1}{x_i}\right)^2 - N \cdot \sum\limits_{i=1}^{N} \frac{1}{x_i^2}}$$

$$b_0 = \frac{1}{N}\left(\sum_{i=1}^{N} y_i - b_1 \cdot \sum_{i=1}^{N} \frac{1}{x_i}\right)$$

Жишээ-14.10 $N = 8.$

N	1	2	3	4	5	6	7	8
x_i	1	2	3	4	5	6	7	8
y_i	12.2	6.8	5.2	4.6	3.9	3.7	3.5	3.2

Үр дүн: $b_0 = 1.9357$, $b_1 = 10.160$ Program "**ГИПЕРБОЛ РЕГРЕСС**";

```
     Uses Graph;
     Const w=50;
     Var A,B,C,D,B0,B1:Real;
            N,i:Integer;
     Gd,Gm,XX,YY,X0,Y0:integer;
          SX,SY,J:Real;
            x,y:array[1..w]of real;
    Function F(x:real):real;
    begin
     F:=B0+B1/x
    end;
  Begin
    A:=0;B:=0;C:=0;D:=0;
    Write('N=');Read(N);
    For i:=1 To N Do
     begin
      Write('x(',i,')=');Read(x[i]);
      Write('y(',i,')=');Read(y[i]);
      A:=A+1/x[i];B:=B+y[i];C:=C+sqr(1/x[i]);
      D:=D+y[i]/x[i]
     end;
    B1:=(A*B-N*D)/(sqr(A)-N*C);
    B0:=(B-B1*A)/N;
    Gd:=0;
    InitGraph(Gd,Gm,'');
    ClearDevice;
    X0:=300;Y0:=350;
    Line(X0,16,X0,250);
    Line(0,Round(Y0/2),600,Round(Y0/2));
    OutTextXY(X0-10,16,'Y');
    OutTextXY(590,Round(Y0/2+10),'X');
    OutTextXY(350,20,'Giperbol Regrecc');
    XX:=100;YY:=100;
    SX:=XX/(600-X0);SY:=YY/(Y0-16);
```

```
    For i:=1 to N Do
      begin
       PutPixel(Round(x[i]/sx+X0),Round(Y0/2-y[i]/SY/2),1);
       GraphDefaults;
      end;
    For I:=-XX To XX Do
       If I<>0 Then
      begin
        J:=F(I);
        PutPixel(Round(I/SX+X0),Round(Y0/2-J/SY/2),1);
        GraphDefaults
       end;
 End
```

14.11 Зэрэгт регрессийн коэффициент

$$y(x) = b_0 \cdot x^{b_1}$$

$$b_1 = \frac{\sum_{i=1}^{N} \ln(x_i) \cdot \sum_{i=1}^{N} \ln(y_i) - N \cdot \sum_{i=1}^{N} \ln(x_i) \cdot \ln(y_i)}{\left(\sum_{i=1}^{N} \ln(x_i)\right)^2 - N \cdot \sum_{i=1}^{N} \ln(x_i)^2}$$

$$b_0 = exp\left(\frac{1}{N}\left(\sum_{i=1}^{N} \ln(y_i) - b_1 \sum_{i=1}^{N} \ln(x_i)\right)\right)$$

Жишээ-14.11 $N = 6$.

N	1	2	3	4	5	6
x_i	1	2	3	4	5	6
y_i	3	12	27	48	75	108

Үр дүн: $b_0 = 2.999$, $b_1 = 2.000$

Program "**ЗЭРЭГТ РЕГРЕСС**";

```
      Uses Graph;
      Const w=50;
      Var A,B,C,D,B0,B1:Real;
                N,i:Integer;
      Gd,Gm,XX,YY,X0,Y0:integer;
            SX,SY,J:Real;
                x,y:array[1..w] of real;
    Function F(x:real):real;
      Var P:real;
          k:integer;
```

```
  begin
   P:=1;
   For k:=1 To Round(B1) Do P:=P*x;
   F:=B0*P
  end;
Begin
   A:=0;B:=0;C:=0;D:=0;
   Write('N=');Read(N);
   For i:=1 To N Do
    begin
     Write('x(',i,')=');Read(x[i]);
     Write('y(',i,')=');Read(y[i]);
     A:=A+ln(x[i]);B:=B+ln(y[i]);C:=C+sqr(ln(x[i]));
     D:=D+ln(x[i])*ln(y[i]);
    end;
  B1:=(A*B-N*D)/(sqr(A)-N*C);
  B0:=exp((B-B1*A)/N);
  Gd:=0;
  InitGraph(Gd,Gm,'');
  ClearDevice;X0:=300;Y0:=350;
  Line(X0,16,X0,199);
  Line(0,Round(Y0/2),600,Round(Y0/2));
  OutTextXY(350,20,'Zeregt Regrecc');
  OutTextXY(X0−10,16,'Y');
  OutTextXY(590,Round(Y0/2-10),'X');
  XX:=10;YY:=300;
  SX:=XX/(600-X0);SY:=YY/(Y0-16);
  For i:=1 to N Do
   begin
    PutPixel(Round(x[i]/sx+X0),Round(Y0/2-y[i]/SY/2),1);
    GraphDefaults;
   end;
  For I:=-XX To XX Do
   begin
    J:=F(I);
    PutPixel(Round(I/SX+X0),Round(Y0/2-J/SY/2),1)
   end;
End.
```

14.12 Экспоненциал регрессийн коэффициент

$$y(x) = b_0 \cdot exp(b_1 x)$$

$$b_1 = \frac{\sum_{i=1}^{N} x_i \cdot \sum_{i=1}^{N} \ln(y_i) - N \cdot \sum_{i=1}^{N} x_i \cdot \ln(y_i)}{\left(\sum_{i=1}^{N} x_i\right)^2 - N \cdot \sum_{i=1}^{N} x_i^2}$$

$$b_0 = exp\left(\frac{1}{N}\Big(\sum_{i=1}^{N} \ln(y_i) - b_1 \sum_{i=1}^{N} x_i\Big)\right).$$

Жишээ-14.12 $N = 9$.

N	1	2	3	4	5	6	7	8	9
x_i	2	3	4	5	6	7	8	9	10
y_i	3.5	5	6.2	9	13	16	23	30	40

Үр дүн:
$b_0 = 1.939,\ b_1 = 0.305$

```
Program "ЭКСПОНЕНЦИАЛ РЕГРЕСС";
     Uses Graph;
     Const w=50;
     Var A,B,C,D,B0,B1:Real;
            N,i:Integer;
     Gd,Gm,XX,YY,XO,YO:integer;
          SX,SY,J:Real;
            x,y:array[1..w] of real;
    Function F(x:real):real;
    begin
     F:=B0*exp(B1*x)
    end;
  Begin
    A:=0;B:=0;C:=0;D:=0;
    Write('N=');Read(N);
    For i:=1 To N Do
      begin
        Write('x(',i,')=');Read(x[i]);
        Write('y(',i,')=');Read(y[i]);
        A:=A+x[i];B:=B+ln(y[i]);C:=C+sqr(x[i]);     D:=D+x[i]*ln(y[i]);
      end;
    B1:=(A*B-N*D)/(sqr(A)-N*C);
    B0:=exp((B-B1*A)/N);
    Gd:=0;
    InitGraph(Gd,Gm,'');
    ClearDevice;
```

```
    X0:=200;Y0:=350;
    XX:=50;YY:=400;
    Line(X0,16,X0,199);
    Line(0,Round(Y0/2),600,Round(Y0/2));
    OutTextXY(250,20,'Exponenzial Regrecc');
    OutTextXY(X0-10,16,'Y');
    OutTextXY(580,Round(Y0/2-10),'X');
    SX:=XX/(600-X0);SY:=YY/(Y0-16);
    For i:=1 to N Do
      begin
        PutPixel(Round(x[i]/sx+X0),Round(Y0/2-y[i]/SY/2),1);
        GraphDefaults;
      end;
    For I:=-XX To XX Do
      begin
        J:=F(I);
        PutPixel(Round(I/SX+X0),Round(Y0/2-J/SY/2),1);
        GraphDefaults
      end;
  End.
```

14.13 Парабол регрессийн коэффициент

$$y(x) = b_0 + b_1 x + b_2 x^2$$

$$b_0 N + b_1 \cdot \sum_{i=1}^{N} x_i + b_2 \cdot \sum_{i=1}^{N} x_i^2 = \sum_{i=1}^{N} y_i$$

$$b_0 \sum_{i=1}^{N} x_i + b_1 \sum_{i=1}^{N} x_i^2 + b_2 \sum_{i=1}^{N} x_i^3 = \sum_{i=1}^{N} x_i y_i$$

$$b_0 \sum_{i=1}^{N} x_i^2 + b_1 \sum_{i=1}^{N} x_i^3 + b_2 \sum_{i=1}^{N} x_i^4 = \sum_{i=1}^{N} x_i^2 y_i$$

Жишээ-14.13 $N = 7$.

N	1	2	3	4	5	6	7
x_i	2	4	6	8	10	12	14
y_i	3.76	4.44	5.04	5.56	6	6.36	6.64

Үр дүн: $b_0 = 3$, $b_1 = 0.4$, $b_2 = -0.01$

Program "**ПАРАБОЛ РЕГРЕСС**";

```
    Uses Graph;
    Var B,C,F,M,P,R,S:Real;
```

```
3       D,E,K,L,Q,A:real;
      B2,B1,B0,SX,SY,J:real;
          N,i,Gd,Gm:Integer;
        XX,YY,X0,Y0:integer;
              x,y:array[1..50] of real;
   Function FF(x:real):real;
   begin
     FF:=B0+B1*x+B2*sqr(x)
   end;
Begin
    B:=0;C:=0;F:=0;M:=0;P:=0;R:=0;S:=0;
    Write('N=');Read(N);A:=N;
    For i:=1 To N Do
      begin
        Write('x(',i,')=');Read(x[i]);
        Write('y(',i,')=');Read(y[i]);
        B:=B+x[i];C:=C+sqr(x[i]);F:=F+x[i]*sqr(x[i]);
        M:=M+sqr(sqr(x[i]));P:=P+y[i];R:=R+x[i]*y[i];
        S:=S+y[i]*sqr(x[i]);
      end;
    D:=B;E:=C;K:=C;L:=F;
    Q:=D/A;E: =E−Q*B;F:=F−Q*C;
    R:=R−Q*P;Q:=K/A;L:=L-Q*B;
    M:=M-Q*C;S:=S-Q*P;Q:=L/E;
    B2:=(S-R*Q)/(M-F*Q);
    B1:=(R-F*B2)/E;
    B0:=(P-B*B1-C*B2)/A;
    Gd:=0;
    InitGraph(Gd,Gm,");
    ClearDevice
    OutTextXY(300,20,'Parabollog Regrecc');
    X0:=250;Y0:=350;
    Line(X0,16,X0,GetMaxY);
    Line(0,Round(Y0/4),600,Round(Y0/4));
    OutTextXY(X0-10,16,'Y');
    OutTextXY(590,Round(Y0/4)+8,'X');
    XX:=500;YY:=400;
    SX:=XX/(600-X0);SY:=YY/(Y0-16);
    For i:=1 To N Do
    PutPixel(Round(x[i]/SX+X0),Round(Y0/4-y[i]/SY/4),1);
    For I:= −XX To XX Do
      begin
        J:=FF(I);
```

```
        PutPixel(Round(I/SX+X0),Round(Y0/4-J/SY/4),1);
        GraphDefaults
      end
End.
```

Хавсралт

Таблиц N1

Хэвийн тархалтын магадлалын нягтын функц $\varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-x^2/2}$

x	0	1	2	3	4	5	6	7	8	9
0.0	39894	39892	39886	39876	39862	39884	39882	39797	39767	39733
0.1	39695	39654	39608	39559	39505	39448	39387	39322	39253	39181
0.2	39104	39024	38940	38853	38762	38667	38568	38466	38361	38251
0.3	38139	38023	37903	37780	37654	37524	37391	37255	37115	36973
0.4	36827	36678	36526	36371	36213	36053	35889	35723	35553	35381
0.5	35207	35029	34849	34667	34482	34294	34105	33912	33718	33521
0.6	33322	33121	32918	32713	32506	32297	32086	31874	31659	31443
0.7	31225	31006	30785	30563	30339	30114	29887	29659	29430	29200
0.8	28969	28737	28504	28269	28034	27798	27562	27324	27086	26848
0.9	26609	26369	26129	25888	25647	25406	25164	24923	24681	24439
1.0	24197	23955	23713	23471	23230	22988	22747	22506	22265	22025
1.1	21785	21546	21307	21069	20831	20594	20327	20121	19886	19652
1.2	19419	19186	18954	18724	18494	18265	18037	17810	17585	17360
1.3	17137	16915	16694	16474	16256	16038	15822	15608	15395	15183
1.4	14973	14764	14556	14350	14146	13943	13742	13542	13344	13147
1.5	12952	12758	12566	12376	12188	12001	11816	11632	11450	11270
1.6	11092	10915	10741	10567	10396	10226	10059	09893	09728	09566
1.7	09405	09246	09089	08933	08780	08628	08478	08329	08183	08038
1.8	07895	07754	07614	07477	07341	07206	07074	06943	06814	06687
1.9	06562	06438	06316	06195	06077	05959	05844	05730	05618	05508
2.0	05399	05292	05186	05082	04980	04879	04780	04682	04586	04491
2.1	04398	04307	04217	04128	04041	03955	03871	03788	03706	03626
2.2	03547	03470	03394	03319	03246	03174	03103	03034	02965	02898
2.3	02833	02768	02705	02643	02582	02522	02463	02406	02349	02294
2.4	02239	02186	02134	02083	02033	01984	01936	01888	01842	01797
2.5	01753	01709	01667	01625	01585	01545	01506	01468	01431	01394
2.6	01358	01323	01289	01256	01223	01191	01160	01130	01100	01071
2.7	01042	01014	00987	00961	00935	00909	00885	00861	00837	00814
2.8	00792	00770	00748	00727	00707	00687	00668	00649	00631	00613
2.9	00595	00578	00562	00545	00530	00514	00499	00485	00470	00457
3.0	00443	00430	00417	00405	00393	00381	00370	00358	00348	00337
3.1	00327	00317	00307	00298	00288	00279	00271	00262	00254	00246
3.2	00238	00231	00224	00216	00210	00203	00196	00190	00184	00178
3.3	00172	00167	00161	00156	00151	00146	00141	00136	00132	00127
3.4	00123	00119	00115	00111	00107	00104	00100	00097	00094	00090
3.5	00087	00084	00081	00079	00076	00073	00071	00068	00066	00063
3.6	00061	00059	00057	00055	00053	00051	00049	00047	00046	00044
3.7	00042	00041	00039	00038	00037	00035	00034	00034	00031	00030
3.8	00029	00028	00027	00026	00025	00024	00023	00022	00021	00021
3.9	00020	00019	00018	00018	00017	00016	00016	00015	00014	00014
4.0	00013	00009	00006	00004	00002	00002	00001	00001	00000	00000

Таблиц N2

Лапласын функцийн утга $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-t^2/2}$, $\Phi(-x) = 1 - \Phi(x)$.

x	0	1	2	3	4	5	6	7	8	9
0.0	50000	50399	50798	51197	51595	51994	52392	52790	53188	53586
0.1	53983	54380	54776	55172	55567	55962	56356	56749	57142	57535
0.2	57926	58317	58706	59095	59483	59871	60256	60642	61026	61409
0.3	61791	62172	62552	62930	63307	63683	64054	64431	64803	65173
0.4	65542	65910	66276	66640	67003	67364	67724	68082	68439	68793
0.5	69146	69497	69847	70194	70540	70884	71226	71566	71904	72240
0.6	72575	72907	73237	73565	73891	74215	74537	74857	75175	75490
0.7	75804	76115	76424	76730	77035	77337	77637	77935	78230	78524
0.8	78814	79103	79389	79673	79955	80234	80511	80785	81057	81327
0.9	81594	81859	82121	82381	82639	82894	83147	83398	83646	83891
1.0	84134	84375	84614	84850	85083	85314	85543	85769	85993	86214
1.1	86433	86650	86864	87076	87286	87493	87698	87900	88100	88298
1.2	88493	88686	88877	89065	89251	89435	89617	89796	89973	90147
1.3	90320	90490	90658	90824	90988	91149	91308	91466	91621	91774
1.4	91924	92073	92220	92364	92507	92647	92786	92922	93056	93189
1.5	93319	93448	93574	93699	93822	93943	94062	94179	94295	94408
1.6	94520	94630	94738	94845	94950	95053	95154	95254	95352	95449
1.7	95543	85637	95728	95814	95907	95994	96080	96164	96246	96327
1.8	96407	96485	96562	96638	96714	96784	96856	96926	96995	97062
1.9	97128	97193	97257	97320	97381	97441	97500	97558	97615	97670
2.0	97725	97778	97831	97882	97932	97982	98030	98077	98124	98169
2.1	98214	98257	98300	98341	98382	98422	98461	98500	98537	98574
2.2	98610	98645	98679	98713	98745	98778	98809	98840	98870	98899
2.3	98928	98956	98983	99010	99036	99061	99086	99111	99134	99158
2.4	99180	99202	99224	99245	99266	99286	99305	99324	99343	99361
2.5	99376	99396	99413	99430	99446	99461	99477	99492	99506	99520
2.6	99534	99547	99560	99573	99585	99598	99609	99621	99632	99643
2.7	99653	99664	99674	99683	99693	99702	99711	99720	99728	99736
2.8	99744	99752	99760	99767	99774	99781	99788	99795	99801	99807
2.9	99813	99819	99825	99831	99836	99841	99846	99851	99856	99861
3.0	99865	99869	99874	99878	99882	99886	99889	99893	99896	99900
3.1	99903	99906	99910	99913	99916	99918	99921	99924	99926	99929
3.2	99931	99934	99936	99938	99940	99942	99944	99946	99948	99950
3.3	99952	99953	99955	99957	99958	99960	99961	99962	99964	99965
3.4	99966	99968	99969	99970	99971	99972	99973	99974	99975	99976
3.5	99977	99978	99978	99979	99980	99981	99981	99982	99983	99983
3.6	99984	99985	99985	99986	99986	99987	99987	99988	99988	99989
3.7	99989	99990	99990	99990	99991	99991	99992	99992	99992	99992
3.8	99993	99993	99993	99994	99994	99994	99994	99995	99995	99995
3.9	99995	99995	99996	99996	99996	99996	99996	99996	99997	99997
4.0	99997	99998	99999	99999	99999	1.000	1.000	1.000	1.000	1.000

Таблиц N3

Пуассоны тархалтын магадлалын утга $P_n(k) = \frac{\lambda^k}{k!}e^{-\lambda}$. $(\lambda = n\cdot p)$

λ / k	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0	9048	8187	7408	6703	6065	5488	4966	4493	4066
1	0905	1638	2222	2681	3033	3293	3476	3595	3659
2	0045	0164	0333	0536	0758	0988	1217	1438	1647
3	0002	0019	0033	0072	0126	0198	0248	0383	0494
4		0001	0002	0007	0016	0030	0050	0077	0111
5				0001	0002	0004	0007	0012	0020
							0001	0002	0003

λ / k	1	2	3	4	5	6	7	8	9	10
0	3679	1353	0498	0183	0067	0025	0009	0003	0001	0000
1	3679	2707	1494	0733	0337	0149	0064	0027	0011	0005
2	1839	2707	2240	1465	0842	0446	0223	0107	0050	0023
3	0613	1804	2240	1954	1404	0892	0521	0286	0150	0076
4	0153	0902	1680	1954	1766	1339	0912	0572	0337	0189
5	0031	0361	1008	1563	1755	1606	1277	0916	0607	0378
6	0005	0120	0504	1042	1462	1606	1490	1221	0911	0631
7	0001	0037	0216	0595	1044	1377	1490	1396	1171	0901
8		0009	0081	0298	0653	1033	1304	1396	1318	1126
9		0002	0027	0132	0363	0688	1014	1241	1318	1251
10			0008	0053	0181	0413	0710	0993	1186	1251
11			0002	0019	0082	0225	0452	0722	0970	1137
12			0001	0006	0034	0126	0263	0481	0728	0948
13				0002	0013	0052	0142	0296	0504	0729
14				0001	0005	0022	0071	0169	0324	0521
15					0002	0009	0033	0090	0194	0347
16						0003	0014	0045	0109	0217
17						0001	0006	0021	0058	0128
18							0002	0009	0029	0071
19							0001	0004	0014	0037
20								0002	0006	0019
21								0001	0003	0009
22									0001	0004
23										0002
24										0001

Ноормчлогдсон хэвийн тархалтын зарим квантилууд. $(\Phi(u_{\alpha/2})=1-\alpha)$

итгэх магадлал $p=1-\alpha$	0.90	0.95	0.99	0.9973	0.999
квантил $u_{\alpha/2}$	1.64	1.96	2.58	3.00	3.37

Таблиц N4

Пирсоны χ^2−тархалт $P(\chi^2 < t_p) = p$

n	$\chi^2_{0.995}$	$\chi^2_{0.99}$	$\chi^2_{0.975}$	$\chi^2_{0.95}$	$\chi^2_{0.90}$	$\chi^2_{0.75}$	$\chi^2_{0.50}$	$\chi^2_{0.25}$	$\chi^2_{0.10}$	$\chi^2_{0.05}$	$\chi^2_{0.025}$	$\chi^2_{0.01}$	$\chi^2_{0.005}$
1	7.88	6.63	5.02	3.84	2.71	1.32	0.455	0.102	0.0158	0.0039	0.001	0.0002	0.0000
2	10.6	9.21	7.38	5.99	4.61	2.77	1.39	0.575	0.211	0.103	0.0506	0.0201	0.0100
3	12.8	11.3	9.35	7.81	6.25	4.11	2.37	1.21	0.584	0.352	0.216	0.115	0.072
4	14.9	13.3	11.1	9.49	7.78	5.39	3.36	1.92	1.06	0.711	0.484	0.297	0.207
5	16.7	15.1	12.8	11.1	9.24	6.63	4.35	2.67	1.61	1.15	0.831	0.554	0.412
6	18.5	16.8	14.4	12.6	10.6	7.84	5.35	3.45	2.20	1.64	1.24	0.872	0.676
7	20.3	18.5	16.0	14.1	12.0	9.04	6.35	4.25	2.83	2.17	1.69	1.24	0.989
8	22.0	20.1	17.5	15.5	13.4	10.2	7.34	5.07	3.49	2.73	2.18	1.65	1.34
9	23.6	21.7	19.0	16.9	14.7	11.4	8.34	5.90	4.17	3.33	2.70	2.09	1.73
10	25.2	23.2	20.5	18.3	16.0	12.5	9.34	6.74	4.87	3.94	3.25	2.56	2.16
11	26.8	24.7	21.9	19.7	17.3	13.7	10.3	7.58	5.58	4.57	3.82	3.05	2.60
12	28.3	26.2	23.3	21.0	18.5	14.8	11.3	8.44	6.30	5.23	4.40	3.57	3.07
13	29.8	27.7	24.7	22.4	19.8	16.0	12.3	9.30	7.04	5.89	5.01	4.11	3.57
14	31.3	29.1	26.1	23.7	21.1	17.1	13.3	10.2	7.79	6.57	5.63	4.66	4.07
15	32.8	30.6	27.5	25.0	22.3	18.2	14.3	11.0	8.55	7.26	6.26	5.23	4.60
16	34.3	32.0	28.8	26.3	23.5	19.4	15.3	11.9	9.31	7.96	6.91	5.81	5.14
17	35.7	33.4	30.2	27.6	24.8	20.5	16.3	12.8	10.1	8.67	7.56	6.41	5.70
18	37.2	34.8	31.5	28.9	26.0	21.6	17.3	13.7	10.9	9.39	8.23	7.01	6.26
19	38.6	36.2	32.9	30.1	27.2	22.7	18.3	14.6	11.7	10.1	8.91	7.63	6.84
20	40.0	37.6	34.2	31.4	28.4	23.8	19.3	15.5	12.4	10.9	9.59	8.26	7.43
21	41.4	38.9	35.5	32.7	29.6	24.9	20.3	16.3	13.2	11.6	10.3	8.90	8.03
22	42.8	40.3	36.8	33.9	30.8	26.0	21.3	17.2	14.0	12.3	11.0	9.54	8.64
23	44.2	41.6	38.1	35.2	32.0	27.1	22.3	18.1	14.8	13.1	11.7	10.2	9.26
24	45.6	43.0	39.4	36.4	33.2	28.2	23.3	19.0	15.7	13.8	12.4	10.9	9.89
25	46.9	44.3	40.6	37.7	34.4	29.3	24.3	19.9	16.5	14.6	13.1	11.5	10.5
26	48.3	45.6	41.9	38.9	35.6	30.4	25.3	20.8	17.3	15.4	13.8	12.2	11.2
27	49.6	47.0	43.2	40.1	36.7	31.5	26.3	21.7	18.1	16.2	14.6	12.9	11.8
28	51.0	48.3	44.5	41.3	37.9	32.6	27.3	22.7	18.9	16.9	15.3	13.6	12.5
29	52.3	49.6	45.7	42.6	39.1	33.7	28.3	23.6	19.8	17.7	16.0	14.3	13.1
30	53.7	50.9	47.0	43.8	40.3	34.8	29.3	24.5	20.6	18.5	16.8	15.0	13.8
40	66.8	63.7	59.3	55.8	51.8	45.6	39.3	33.7	29.1	26.5	24.4	22.2	20.7
50	79.5	76.2	71.4	67.5	63.2	56.3	49.3	42.9	37.7	34.8	32.4	29.7	28.0
60	92.0	88.4	83.3	79.1	74.4	67.0	59.3	52.3	46.5	43.2	40.5	37.5	35.5
70	104.2	100.4	95.0	90.5	85.5	77.6	69.3	61.7	55.3	51.7	48.8	45.4	43.3
80	116.3	112.3	106.6	101.9	96.6	88.1	79.3	71.1	64.3	60.4	57.2	53.5	51.2
90	128.3	124.1	118.1	113.1	107.6	98.6	89.3	80.6	73.3	69.1	65.6	61.8	59.2
100	140.2	135.8	129.6	124.3	118.5	109.1	99.3	90.1	82.4	77.9	74.2	70.1	67.3

Таблиц N5

Колмогоровын функцийн утга $P(\lambda)=\sum\limits_{k=-\infty}^{+\infty}(-1)^k\cdot e^{-2k^2}\lambda^2$

λ	$P(\lambda)$	λ	$P(\lambda)$	λ	$P(\lambda)$
0.34	0.00017	0.98	0.7079	1.62	0.9895
0.36	0.0005	1.00	0.7300	1.64	0.9908
0.38	0.0013	1.02	0.7500	1.66	0.9918
0.40	0.0028	1.04	0.7704	1.68	0.9929
0.42	0.0055	1.06	0.7889	1.70	0.9938
0.44	0.0097	1.08	0.8061	1.72	0.9946
0.46	0.0160	1.10	0.8223	1.74	0.9953
0.48	0.0247	1.12	0.8374	1.76	0.9956
0.50	0.0361	1.14	0.8514	1.78	0.9965
0.52	0.0503	1.16	0.8644	1.80	0.9969
0.54	0.0675	1.18	0.8765	1.82	0.9973
0.56	0.0876	1.20	0.8878	1.84	0.9977
0.58	0.1104	1.22	0.8981	1.86	0.9980
0.60	0.1357	1.24	0.9076	1.88	0.9983
0.62	0.1632	1.26	0.9164	1.90	0.9985
0.64	0.1927	1.28	0.9245	1.92	0.9987
0.66	0.2236	1.30	0.9319	1.94	0.9989
0.68	0.2558	1.32	0.9387	1.96	0.9991
0.70	0.2888	1.34	0.9449	1.98	0.9992
0.72	0.3223	1.36	0.9505	2.00	0.9993
0.74	0.3560	1.38	0.9557	2.02	0.9994
0.76	0.3896	1.40	0.9603	2.04	0.9995
0.78	0.4230	1.42	0.9646	2.06	0.9996
0.80	0.4559	1.44	0.9684	2.08	0.9997
0.82	0.4880	1.46	0.9718	2.10	0.9997
0.84	0.5194	1.48	0.9750	2.12	0.9998
0.86	0.5497	1.50	0.9778	2.14	0.9998
0.88	0.5791	1.52	0.9803	2.16	0.9998
0.90	0.6073	1.54	0.9826	2.18	0.9998
0.92	0.6343	1.56	0.9846	2.20	0.9998
0.96	0.6846	1.60	0.9880	2.40	0.9999

Таблиц N6

R/S−харьцааны доод ба дээд хилүүд

	Алдааны магадлалын доод хилүүд						Алдааны магадлалын дээд хилүүд					
n	0.000	0.005	0.001	0.025	0.05	0.10	0.10	0.05	0.025	0.001	0.005	0.000
3	1.732	1.735	1.737	1.745	1.758	1.782	1.997	1.999	2.000	2.000	2.000	2.000
4	1.732	1.83	1.87	1.93	1.98	2.04	2.709	2.429	2.439	2.445	2.447	2.449
5	1.826	1.98	2.02	2.09	2.15	2.22	2.712	2.753	2.782	2.803	2.813	2.828
6	1.826	2.11	2.15	2.22	2.28	2.37	2.949	3.012	3.056	3.095	3.115	3.162
7	1.871	2.22	2.26	2.33	2.40	2.49	3.143	3.222	3.282	3.338	3.369	3.465
8	1.871	2.31	2.35	2.43	2.50	2.59	3.308	3.399	3.471	3.543	3.585	3.742
9	1.897	2.39	2.44	2.51	2.59	2.68	3.449	3.552	3.634	3.720	3.772	4.000
10	1.897	2.46	2.51	2.59	2.67	2.76	3.57	3.685	3.777	3.875	3.935	4.243
11	1.915	2.53	2.58	2.66	2.74	2.84	3.68	3.80	3.903	4.012	4.079	4.472
12	1.915	2.59	2.64	2.72	2.80	2.90	3.78	3.91	4.02	4.134	4.208	4.690
13	1.927	2.64	2.70	2.78	2.86	2.96	3.87	4.00	4.12	4.244	4.325	4.899
14	1.927	2.70	2.75	2.83	2.92	3.02	3.95	4.09	4.21	4.34	4.431	5.099
15	1.936	2.74	2.80	2.88	2.97	3.07	4.02	4.17	4.29	4.44	4.53	5.292
16	1.936	2.79	2.84	2.98	3.01	3.12	1.09	4.24	4.37	4.52	4.62	5.477
17	1.944	2.83	2.88	2.97	3.06	3.17	4.15	4.31	4.44	4.60	4.70	5.657
18	1.944	2.87	2.92	3.01	3.10	3.21	4.21	4.37	4.51	4.67	4.78	5.831
19	1.949	2.90	2.96	3.05	3.14	3.25	4.27	4.43	4.57	4.74	4.85	6.000
20	1.949	2.94	2.99	3.09	3.18	3.29	4.32	4.49	4.63	4.80	4.91	6.164
25	1.961	3.09	3.15	3.24	3.34	3.45	4.53	4.71	4.87	5.06	5.19	6.93
30	1.966	3.21	3.27	3.37	3.47	3.59	4.70	4.89	5.06	5.26	5.40	7.62
35	1.972	3.32	3.38	3.48	3.58	3.70	4.84	5.04	5.21	5.42	5.57	8.25
40	1.975	3.41	3.47	3.37	3.67	3.79	4.96	5.16	5.34	5.56	5.71	8.83
45	1.978	3.49	3.55	3.66	3.75	3.88	5.06	5.26	5.45	5.67	5.83	9.38
50	1.980	3.56	3.62	3.73	3.83	3.95	5.14	5.35	5.54	5.77	5.93	9.90
55	1.982	3.62	3.69	3.80	3.90	4.02	5.22	5.43	5.63	5.86	6.02	10.39
60	1.983	3.68	3.75	3.86	3.96	4.08	5.29	5.51	5.70	5.94	6.10	10.86
65	1.985	3.74	3.80	3.91	4.01	4.14	5.35	5.57	5.77	6.01	6.17	11.31
70	1.986	3.79	3.85	3.96	4.06	4.19	5.41	5.63	5.83	6.07	6.24	11.75
75	1.987	3.83	3.90	4.01	4.11	4.24	5.46	5.68	5.88	6.13	6.30	12.17
80	1.987	3.88	3.94	4.05	4.16	4.28	5.51	5.73	5.93	6.18	6.35	12.57
85	1.988	3.92	3.99	4.09	4.20	4.33	5.56	5.78	5.98	6.23	6.40	12.96
90	1.989	3.96	4.02	4.13	4.24	4.36	5.60	5.82	6.03	6.27	6.45	13.34
95	1.990	3.99	4.06	4.17	4.27	4.40	5.64	5.86	6.07	6.32	6.49	13.71
100	1.990	4.03	4.10	4.21	4.31	4.44	5.68	5.90	6.11	6.36	6.53	14.07
150	1.993	4.32	4.38	4.48	4.59	4.72	5.96	6.18	6.39	6.64	6.82	17.26
200	1.995	4.55	4.59	4.68	4.78	4.90	6.15	6.39	6.60	6.84	7.01	19.95
500	1.998	5.06	5.13	5.25	5.37	5.49	6.72	6.94	7.15	7.42	7.60	31.59
1000	1.999	5.50	5.57	5.68	5.79	5.92	7.11	7.33	7.54	7.80	7.99	44.70

Таблиц N7

Стьюдентийн $t-$тархалт $P(t > t_\alpha) = \alpha$ ба $P(|t| > t_\alpha) = \alpha$

Хоёр талт критик муж								
t_α / k	$t_{0.2}$	$t_{0.1}$	$t_{0.05}$	$t_{0.02}$	$t_{0.01}$	$t_{0.005}$	$t_{0.002}$	$t_{0.001}$
1	3.08	6.31	12.71	31.82	63.66	127.32	318.30	636.61
2	1.89	2.92	4.30	6.96	9.92	14.09	22.33	31.60
3	1.64	2.35	3.18	4.54	5.84	7.45	10.21	12.92
4	1.53	2.13	2.78	3.75	4.60	5.60	7.17	8.61
5	1.48	2.02	2.57	3.36	4.03	4.77	5.89	6.87
6	1.44	1.94	2.45	3.14	3.71	4.32	5.21	5.96
7	1.41	1.89	2.36	3.00	3.50	4.03	4.79	5.41
8	1.40	1.86	2.31	2.90	3.36	3.83	4.50	5.04
9	1.38	1.83	2.26	2.82	3.25	3.69	4.30	4.78
10	1.37	1.81	2.23	2.76	3.17	3.58	4.14	4.59
11	1.36	1.80	2.20	2.72	3.11	3.50	4.02	4.44
12	1.36	1.78	2.18	2.68	3.05	3.43	3.93	4.32
13	1.35	1.77	2.16	2.65	3.01	3.37	3.85	4.22
14	1.34	1.76	2.14	2.62	2.98	3.33	3.79	4.14
15	1.34	1.75	2.13	2.60	2.95	3.29	3.73	4.07
16	1.34	1.75	2.12	2.58	2.92	3.25	3.69	4.02
17	1.33	1.74	2.11	2.57	2.90	3.22	3.65	3.97
18	1.33	1.73	2.10	2.55	2.88	3.20	3.61	3.92
19	1.33	1.73	2.09	2.54	2.86	3.17	3.58	3.88
20	1.33	1.72	2.09	2.53	2.85	3.15	3.55	3.85
21	1.32	1.72	2.08	2.52	2.83	3.14	3.53	3.82
22	1.32	1.72	2.07	2.51	2.82	3.12	3.51	3.79
23	1.32	1.71	2.07	2.50	2.81	3.10	3.48	3.77
24	1.32	1.71	2.06	2.49	2.80	3.09	3.47	3.75
25	1.32	1.71	2.06	2.49	2.79	3.08	3.45	3.73
26	1.32	1.71	2.06	2.48	2.78	3.07	3.44	3.71
27	1.31	1.70	2.05	2.47	2.77	3.06	3.42	3.69
28	1.31	1.70	2.05	2.47	2.76	3.05	3.41	3.67
29	1.31	1.70	2.05	2.46	2.76	3.04	3.40	3.66
30	1.31	1.70	2.04	2.46	2.75	3.03	3.39	3.65
40	1.30	1.68	2.02	2.42	2.70	2.97	3.31	3.55
60	1.30	1.67	2.00	2.39	2.66	2.91	3.23	3.46
120	1.29	1.66	1.98	2.36	2.62	2.85	3.16	3.37
∞	1.28	1.64	1.96	2.33	2.58	2.81	3.09	3.29
t_α	$t_{0.1}$	$t_{0.05}$	$t_{0.025}$	$t_{0.01}$	$t_{0.005}$	$t_{0.0025}$	$t_{0.001}$	$t_{0.0005}$
Нэг талт критик муж								

Таблиц N8

Фишерийн F−тархалт $P(F>t_\alpha)=\alpha, \quad \alpha=0.05$

k_1 / k_2	1	2	3	4	5	6	8	10	12	24	∞
1	161.4	199.5	215.7	224.6	230.2	234.0	238.9	242	243.9	249.0	254.3
2	18.51	19.00	19.16	19.25	19.30	19.33	19.37	19.4	19.41	19.45	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.84	8.79	8.74	8.64	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.04	5.96	5.91	5.77	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.82	4.74	4.68	4.53	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.15	4.06	4.00	3.84	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.73	3.64	3.57	3.41	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.44	3.35	3.28	3.12	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.23	3.14	3.07	2.90	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.07	2.98	2.91	2.74	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	2.95	2.85	2.79	2.61	2.40
12	4.75	3.88	3.49	3.26	3.11	3.00	2.85	2.75	2.69	2.50	2.30
13	4.67	3.80	3.41	3.18	3.02	2.92	2.77	2.67	2.60	2.42	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.70	2.60	2.53	2.35	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.64	2.54	2.48	2.29	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.59	2.49	2.42	2.24	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.55	2.45	2.38	2.19	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.51	2.41	2.34	2.15	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.48	2.38	2.31	2.11	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.45	2.35	2.28	2.08	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.42	2.32	2.25	2.05	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.40	2.30	2.23	2.03	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.38	2.27	2.20	2.00	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.36	2.25	2.18	1.98	1.73
25	4.24	3.38	2.99	2.76	2.60	2.49	2.34	2.24	2.16	1.96	1.71
26	4.22	3.37	2.98	2.74	2.59	2.47	2.32	2.22	2.15	1.95	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.30	2.20	2.13	1.93	1.67
28	4.20	3.34	2.95	2.71	2.56	2.44	2.29	2.19	2.12	1.91	1.65
29	4.18	3.33	2.93	2.70	2.54	2.43	2.28	2.18	2.10	1.90	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.27	2.16	2.09	1.89	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.18	2.08	2.00	1.79	1.52
60	4.00	3.15	2.76	2.52	2.37	2.25	2.10	1.99	1.92	1.70	1.39
120	3.92	3.07	2.68	2.45	2.29	2.17	2.02	1.91	1.83	1.61	1.25
∞	3.84	2.99	2.60	2.37	2.21	2.09	1.94	1.83	1.75	1.52	1.00

Фишерийн F−тархалт $P(F > t_\alpha) = \alpha, \quad \alpha = 0.01$

k_1 / k_2	1	2	3	4	5	6	8	10	12	24	∞
1	4052	4999	5403	5625	5764	5859	5981	6056	6106	6234	6366
2	98.49	99.00	99.17	99.25	99.30	99.33	99.36	99.40	99.42	99.46	99.50
3	34.12	30.81	29.46	28.71	28.24	27.91	27.49	27.23	27.05	26.60	26.12
4	21.20	18.00	16.69	15.98	15.52	15.21	14.80	14.55	14.37	13.93	13.46
5	16.26	13.27	12.06	11.39	10.97	10.67	10.29	10.05	9.89	9.47	9.02
6	13.74	10.92	9.78	9.15	8.75	8.47	8.10	7.87	7.72	7.31	6.83
7	12.25	9.55	8.45	7.85	7.46	7.19	6.84	6.62	6.47	6.07	5.65
8	11.26	8.65	7.59	7.01	6.63	6.37	6.03	5.81	5.67	5.28	4.86
9	10.56	8.02	6.99	6.42	6.06	5.80	5.47	5.26	5.11	4.73	4.31
10	10.04	7.56	6.55	5.99	5.64	5.39	5.06	4.85	4.71	4.33	3.91
11	9.65	7.20	6.22	5.67	5.32	5.07	4.74	4.54	4.40	4.02	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.50	4.30	4.16	3.78	3.36
13	9.07	6.70	5.74	5.20	4.86	4.62	4.30	4.10	3.96	3.59	3.16
14	8.86	6.51	5.56	5.03	4.69	4.46	4.14	3.94	3.80	3.43	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.00	3.80	3.67	3.29	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	3.89	3.69	3.55	3.18	2.75
17	8.40	6.11	5.18	4.67	4.34	4.10	3.79	3.59	3.45	3.08	2.65
18	8.28	6.01	5.09	4.58	4.25	4.01	3.71	3.51	3.37	3.00	2.57
19	8.18	5.93	5.01	4.50	4.17	3.94	3.63	3.43	3.30	2.92	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.56	3.37	3.23	2.86	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.51	3.31	3.17	2.80	2.36
22	7.94	5.72	4.86	4.31	3.99	3.76	3.45	3.26	3.12	2.75	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.41	3.21	3.07	2.70	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.36	3.17	3.03	2.66	2.21
25	7.77	5.57	4.68	4.18	3.86	3.63	3.32	3.13	2.99	2.62	2.17
26	7.72	5.53	4.64	4.14	3.82	3.59	3.29	3.09	2.96	2.58	2.13
27	7.68	5.49	4.60	4.11	3.78	3.56	3.26	3.06	2.93	2.55	2.10
28	7.64	5.45	4.57	4.07	3.75	3.53	3.23	3.03	2.90	2.52	2.06
29	7.60	5.42	4.54	4.04	3.73	3.50	3.20	3.00	2.87	2.49	2.03
30	7.56	5.39	4.51	4.02	3.70	3.47	3.17	2.98	2.84	2.47	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	2.99	2.80	2.66	2.29	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.82	2.63	2.50	2.12	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.66	2.47	2.34	1.95	1.38
∞	6.64	4.60	3.78	3.32	3.02	2.80	2.51	2.32	2.18	1.79	1.00

Таблиц N9

$r = thz$ функцийн утгын таблиц.

z	0.00	0.02	0.04	0.06	0.08
0.0	0.000	0.020	0.040	0.060	0.080
0.1	0.100	0.119	0.139	0.159	0.178
0.2	0.197	0.217	0.236	0.254	0.273
0.3	0.291	0.310	0.328	0.345	0.363
0.4	0.380	0.397	0.414	0.430	0.446
0.5	0.462	0.478	0.493	0.508	0.523
0.6	0.537	0.551	0.565	0.578	0.592
0.7	0.604	0.617	0.629	0.641	0.653
0.8	0.664	0.675	0.686	0.696	0.706
0.9	0.716	0.726	0.735	0.744	0.753
1.0	0.762	0.770	0.778	0.786	0.793
1.1	0.801	0.808	0.814	0.821	0.828
1.2	0.834	0.840	0.846	0.851	0.851
1.3	0.862	0.867	0.872	0.876	0.881
1.4	0.885	0.890	0.894	0.898	0.902
1.5	0.905	0.909	0.912	0.915	0.919
1.6	0.922	0.925	0.928	0.930	0.933
1.7	0.936	0.938	0.940	0.943	0.945
1.8	0.947	0.949	0.951	0.953	0.955
1.9	0.956	0.958	0.960	0.961	0.963

Ном зүй

[1] Douglas C. Montgomery, George C. Runger. *Applied Statistics and Probability for Engineers*, USA,Third edition, 2003, ISBN 0-471-20454-4.

[2] Murray R. Spiedel. *Theory and Problems of Statistics*, Singapore, First edition, 1972,ISBN 0-07-099008-5.

[3] Ахназарова.С.Л., Кафаров.В.В. *Методы оптимизации эксперимента в химической технологи* , -М. 1985.

[4] Боровков.А.А. *Теория вероятностей.* , -М. 1982.

[5] Вентцель.Е.С. *Теория вероятностей.* , -М. 1971.

[6] Гмурман.В.Е. *Теория вероятностей и математическая статистика.* , -М. 1972.

[7] Иванова.В.М., Калинина.В.Н., Нешумова.Л.А., Решетникова.И.О. *Математическая статистика.* ,-М. 1981.

[8] Карасев.А.И. и др. *Математические методы и модели в планировании.* ,-М. 1981.

[9] Карасев.А.И., Аксютина.З.М., Савельева.Т.И. *Курс высшей математики для экономических вузов. II том.* , -М. 1982.

[10] Колде.Я.К. *Практикум по теории вероятностей и математической статистике.* , -М. 1991.

[11] Львовский.Е.Н. *Статистические методы построения эмпирических формул.* , -М. 1988.

[12] Солодовников.А.С. *Теория вероятностей.* , -М. 1983.

[13] Теннант-Смит.ДЖ. *Бейсик для статистиков.* , -М. 1988.