Ковалішина І.В. Теорія ймовірностей: Конспект лекцій. Харків: УкрДАЗТ, 2004. – Ч. 2. - 69 с.

Конспект лекцій розглянуто та рекомендовано до друку на засіданні кафедри «Вища математика» 2 лютого 2004 р., протокол № 6.

Рецензент

проф. А.А.Янцевич (ХНУ)

ЗМІСТ

ТЕМА 1 ВАЖЛИВІШІ ЗАКОНИ РОЗПОДІЛУ ВИПАДКОВОЇ ВЕЛИЧИНИ

Розглянемо спочатку закони розподілу дискретних випадкових величин.

1.1 Біномний закон розподілу

Розглянемо випадкову величину **X**, яка дорівнює числу появ події **A** в „n" незалежних випробуваннях і запишемо для неї ряд розподілу

X	0	1	2	3	...	k	...	n
P	$P_n(0)$	$P_n(1)$	$P_n(2)$	$P_n(3)$	...	$P_n(k)$	...	$P_n(n)$

де $P_n(k)$ обчислюється за формулою Бернуллі

$$P_n(k)=C_n^k p^k q^{n-k}$$

де p - ймовірність появи події А в одному експерименті;
q – ймовірність протилежної події, тобто $q=1-p$.

Закон розподілу такої випадкової величини називається ***біномним***.

Назва пов'язана з тим, що ймовірності можливих значень випадкової величини X виявляються окремими елементами формули бінома Н'ютона, а саме:

$$(q+p)^n = q^n + C_n^1 q^{n-1} p + C_n^2 q^{n-2} p^2 + \ldots + C_n^n q^0 p^n = 1^n = 1,$$

тобто сума ймовірностей дійсно дорівнює одиниці.

Розглянемо далі числові характеристики біномного розподілу, тобто M_X, D_X, σ_X:

$$M_X = \sum_{k=0}^{n} k \cdot P_n(k) = \sum_{k=0}^{n} k \cdot C_n^k p^k q^{n-k}$$

$$D_X = M(X^2) - [M(X)]^2 = \sum_{k=0}^{n} k^2 P_n(k) - [M_X]^2$$

Суми у попередніх формулах можна безпосередньо підрахувати, але простіше зробити так.

Розглянемо сукупність випадкових подій $X_1, X_2, \ldots, X_n$, причому будемо вважати, що X_i – подія, яка полягає в тому, що подія А з'являється, або не з'являється в i – му випробуванні.

Очевидно, для кожної X_i має місце ряд розподілу

$$\begin{array}{|c|c|c|} X & 1 & 0 \\ P & p & q \end{array}$$

Обчислимо математичне сподівання $M(X_i)$ і дисперсію $D(X_i)$

$$M(X_i) = 1 \cdot p + 0 \cdot q = p,$$

$$D(X_i) = M(X_i^2) - [M(X_i)]^2 = 1^2 \cdot p + 0^2 \cdot q - p^2 = p - p^2 = p(1-p) = pq$$

Таким чином, усі випадкові величини X_i мають однакові математичні сподівання $M(X_i) = p$ і дисперсії $D(X_i) = pq$.

Випадкова величина X, яка дорівнює числу появ події A в „n" незалежних випробуваннях, дорівнює сумі $X = X_1 + X_2 + \ldots + X_n$, причому $X_1, X_2, \ldots, X_n$ попарно незалежні, тобто:

$$M(X) = M(X_1) + M(X_2) + \ldots + M(X_n) = np,$$

$$D(X) = D(X_1) + D(X_2) + \ldots + D(X_n) = npq.$$

Таким чином, числові характеристики біномного розподілу такі :

$$M(X) = np, \quad D(X) = npq, \quad \sigma(X) = \sqrt{npq}.$$

1.2 Теореми Муавра – Лапласа

При великих значеннях „n" можна наближено обчислювати біномні ймовірності за допомогою локальної теореми Муавра – Лапласа.

Вперше цю теорему отримав Муавр у 1730 році для окремого випадку схеми Бернуллі при p = q = 1/2 ; пізніше Лаплас узагальнив її на випадок довільного "p", відмінного від 0 і 1.

Локальна теорема Муавра – Лапласа

Нехай у кожному з „n" незалежних випробувань ймовірність настання події A однакова і дорівнює **p (0<p<1)** , $P_n(k)$ ймовірність того, що при „n" випробуваннях подія A настане "k" разів. Тоді:

$$P_n(k) \approx \frac{1}{\sqrt{2\pi} \cdot \sqrt{npq}} \ell^{-\frac{(x_k)^2}{2}}, \quad x_k = \frac{k - np}{\sqrt{npq}}$$

(Функція $\phi(x) = \frac{1}{\sqrt{2\pi}} \ell^{-\frac{x^2}{2}}$ - табульована).

При розв'язанні практичних задач часто виникає необхідність обчислення суми виду $\sum_{k=k_1}^{k=k_2} P_n(k)$, тобто ймовірності того, що подія A з'явиться в „n" незалежних випробуваннях не менше ніж "k_1" разів і не більше ніж „k_2" разів $(k_1 \le k \le k_2)$.

Безпосереднє обчислення таких сум навіть при застосуванні локальної теореми Муавра – Лапласа дуже громіздке. Крім того, при додаванні великої кількості наближених значень $P_n(k)$ можуть утворюватися значні похибки.

Такі суми доцільно обчислювати за допомогою теореми, відомої під назвою інтегральної теореми Муавра – Лапласа.

Інтегральна теорема Муавра – Лапласа

$$P_n(k_1 \le k \le k_2) \approx \frac{1}{\sqrt{2\pi}} \int_{x_1}^{x_2} \ell^{-\frac{x^2}{2}} dx, \quad x_1 = \frac{k_1 - np}{\sqrt{npq}}, \quad x_2 = \frac{k_2 - np}{\sqrt{npq}}.$$

Зауваження. Розглянемо інтеграл зі змінною верхньою межею

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \ell^{-\frac{t^2}{2}} dt,$$

який називають **інтегралом ймовірності**. Інтеграл ймовірності $\Phi(x)$ - табульований. За його допомогою, інтегральна теорема Муавра – Лапласа набуває вигляду

$$P_n(k_1 \le k \le k_2) \approx \Phi(x_2) - \Phi(x_1).$$

1.3 Розподіл Пуассона

Розглянемо випадкову величину, яка розподілена за біномним законом і для якої відомо, що число незалежних випробувань $n \to \infty$, а ймовірність „p" появи події A у кожному випробуванні прямує до нуля $(p \to 0)$.

В цьому випадку можна обчислювати біномні ймовірності $P_n(k)$ за наближеною формулою, яка є наслідком теореми Пуассона, доведеною при припущенні, що добуток $p \cdot n = \lambda$ (λ - стала величина).

Теорема Пуассона

Якщо $P_n(k)$ – ймовірність „к" успіхів у серії з „n" випробувань Бернуллі, в кожному із яких ймовірність успіху дорівнює $p = \frac{\lambda}{n}$, то

$$\lim_{n\to\infty} P_n(k) = \frac{\lambda^k}{k!}\ell^{-\lambda}.$$

Доведення. За формулою Бернуллі (q = 1-p)

$$P_n(k) = C_n^k p^k q^{n-k} = C_n^k \frac{\lambda^k}{n^k}\left(1-\frac{\lambda}{n}\right)^{n-k},$$

$$\lim_{n\to\infty} P_n(k) = \lim_{n\to\infty} C_n^k\left(\frac{\lambda}{n}\right)^k\cdot\left(1-\frac{\lambda}{n}\right)^{n-k}.$$

Обчислимо окремо границю кожного множника

$$\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{n-k} = \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{n}\cdot\left(1-\frac{\lambda}{n}\right)^{-k} = \ell^{-\lambda};$$

$$\lim_{n\to\infty} C_n^k\left(\frac{\lambda}{n}\right)^k = \lim_{n\to\infty}\frac{n\cdot(n-1)\cdot\ldots\cdot[n-(k-1)]}{k!}\cdot\frac{\lambda^k}{n^k} = \frac{\lambda^k}{k!}\cdot\lim_{n\to\infty}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right) = \frac{\lambda^k}{k!},$$

тобто дійсно

$$\lim_{n\to\infty} P_n(k) = \frac{\lambda^k}{k!}\cdot\ell^{-\lambda}.$$

Ряд розподілу Пуассона має вигляд:

X	0	1	2	3	…	к	…	n	…
p	$\ell^{-\lambda}$	$\frac{\lambda}{1!}\ell^{-\lambda}$	$\frac{\lambda^2}{2!}\ell^{-\lambda}$	$\frac{\lambda^3}{3!}\ell^{-\lambda}$	…	$\frac{\lambda^k}{k!}\ell^{-\lambda}$	…	$\frac{\lambda^n}{n!}\ell^{-\lambda}$	…

Перевіримо умову нормування, тобто підрахуємо суму усіх ймовірностей

$$\sum_{k=0}^{+\infty} P(k) = \ell^{-\lambda}\sum_{k=0}^{+\infty}\frac{\lambda^k}{k!} = \ell^{-\lambda}\left(1+\frac{\lambda}{1!}+\frac{\lambda^2}{2!}+\ldots+\frac{\lambda^k}{k!}+\ldots\right) = \ell^{-\lambda}\ell^{\lambda} = 1.$$

Числові характеристики розподілу Пуассона такі:

$$\mathrm{M}(X) = \sum_{k=0}^{+\infty} k\frac{\lambda^k}{k!}\ell^{-\lambda} = \sum_{k=1}^{+\infty} k\frac{\lambda^k}{k!}\ell^{-\lambda} = \ell^{-\lambda}\sum_{k=1}^{+\infty}\frac{\lambda^k}{(k-1)!} = \ell^{-\lambda}\cdot\lambda\sum_{k=1}^{+\infty}\frac{\lambda^{k-1}}{(k-1)!} = \ell^{-\lambda}\lambda\sum_{i=0}^{+\infty}\frac{\lambda^i}{i!} = \ell^{-\lambda}\lambda\ell^{\lambda} = \lambda.$$

Таким чином, математичне сподівання в розподілі Пуассона дорівнює

$$\mathrm{M}(X) = \lambda.$$

Зауваження. Можна дійти до попереднього значення математичного сподівання, якщо перейти до границі в математичному сподіванні в біномному розподілі, а саме:

$$\mathrm{M}(X) = \lim_{\substack{n\to\infty\\ p\to 0}} M(X)_b = \lim_{\substack{n\to\infty\\ p\to 0}}(np) = \lambda.$$

Дисперсію обчислюємо так:

$$D(X) = M(X^2) - [M(X)]^2 = \sum_{k=0}^{+\infty} k^2\frac{\lambda^k}{k!}\ell^{-\lambda} - \lambda^2 = \ell^{-\lambda}\lambda\sum_{k=1}^{+\infty}\frac{\lambda^{k-1}}{(k-1)!}k - \lambda^2 =$$

$$= \ell^{-\lambda}\lambda\sum_{k=1}^{+\infty}(k-1)\frac{\lambda^{k-1}}{(k-1)!} + \ell^{-\lambda}\lambda\sum_{k=1}^{+\infty}\frac{\lambda^{k-1}}{(k-1)!} - \lambda^2 = \ell^{-\lambda}\lambda^2\sum_{k=2}^{+\infty}\frac{\lambda^{k-2}}{(k-2)!} + \ell^{-\lambda}\lambda\sum_{i=0}^{+\infty}\frac{\lambda^i}{i!} - \lambda^2 =$$

$$= \ell^{-\lambda}\lambda^2\sum_{i=0}^{+\infty}\frac{\lambda^i}{i!} + \ell^{-\lambda}\lambda\,\ell^{\lambda} - \lambda^2 = \ell^{-\lambda}\lambda^2\ell^{\lambda} + \lambda - \lambda^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

Таким чином, дисперсія теж дорівнює:

$$D(X) = \lambda.$$

Можна зробити і так:

$$D(X) = \lim_{\substack{n\to\infty\\ p\to 0}} D(X)_b = \lim_{\substack{n\to\infty\\ p\to 0}}(npq) = \lim_{\substack{n\to\infty\\ p\to 0}}[np(1-p)] = \lambda.$$

Висновок: параметр λ розподілу Пуассона збігається з математичним сподіванням і дисперсією цього розподілу.

Зауваження. На практиці кількість значень, які приймає випадкова величина, буває великою, але не нескінченою. Тому формула Пуассона має асимптотичний характер і дає наближені значення ймовірності при великих (n>>1) і малих (p<<0,1).

1.4 Геометричний розподіл

В задачі Бернуллі (p=P(A), 0<p<1) розглянемо випадкову величину **X – число випробувань до першої появи події A**.

Очевидно, розподіл цієї випадкової величини має вигляд

X	0	1	2	3	…	n	…
P	p	qp	q^2p	q^3p	…	q^np	…

$$P(X = n) = q^n p.$$

Цей розподіл має властивість, яку називають **властивістю відсутності післядії.** Ця властивість полягає в тому, що ймовірність того, що подія A з'явиться у випробуванні з номером „n+m" не залежить від того, чи з'влялася подія A у випробуваннях з номерами до "n" включно, тобто

$$P(X=n+m/X\geq n) = P(X=m).$$

Дійсно, застосуємо формулу умовної ймовірності

$$P_B(A)=\frac{P(A\cdot B)}{P(B)},$$

тобто

$$P\left(X=n+m\middle/ X\geq n\right)=\frac{P((X=n+m)\cdot(X\geq n))}{P(X\geq n)}=\frac{P(X=n+m)}{P(X\geq n)}=\frac{P(X=n+m)}{\sum\limits_{k=n}^{+\infty}q^k p}=\frac{P(X=n+m)}{pq^n\sum\limits_{k=n}^{+\infty}q^{k-n}}=$$

$$=\frac{P(X=n+m)}{pq^n\sum\limits_{i=0}^{+\infty}q^i}=\frac{pq^{n+m}}{pq^n\cdot\frac{1}{1-q}}=q^m\cdot p=P(X=m).$$

Умова нормування виконується:

$$\sum_{k=0}^{+\infty}p_k=\sum_{k=0}^{+\infty}q^k p=p\cdot\frac{1}{1-q}=1.$$

Визначимо далі числові характеристики геометричного розподілу:

$$M(X)=\sum_{k=0}^{+\infty}kq^k p=\sum_{k=1}^{+\infty}kq^k p=pq\left(1+2q+3q^2+\cdots\right)=pq\left(q+q^2+q^3+\cdots\right)'=pq\left(\frac{q}{1-q}\right)'=pq\frac{1-q+q}{(1-q)^2},$$

тобто

$$M(X)=\frac{q}{p}$$

Дисперсію будемо обчислювати за формулою $D(X)=M(X^2)-[M(X)]^2$.
Обчислимо спочатку перший доданок

$$M(X^2)=\left[M(X^2)+M(X)\right]-M(X)=\left[\sum_{k=1}^{+\infty}k^2q^k p+\sum_{k=1}^{+\infty}kq^k p\right]-\frac{q}{p}=qp\sum_{k=1}^{+\infty}k(k+1)q^{k-1}-\frac{q}{p}=$$

$$=qp\left(1\cdot 2q^0+3\cdot 2q+4\cdot 3q^2+\cdots\right)-\frac{q}{p}=qp\left(q^2+q^3+q^4+\cdots\right)''-\frac{q}{p}=qp\left(\frac{q^2}{1-q}\right)''-\frac{q}{p}=$$

$$qp\left(-1-q+\frac{1}{1-q}\right)''=qp\frac{2}{(1-q)^3}-\frac{q}{p}=\frac{2q}{p^2}-\frac{q}{p}.$$

Далі,

$$D(X)=M(X^2)-[M(X)]^2=\frac{2q}{p^2}-\frac{q}{p}-\frac{q^2}{p^2}=q\,\frac{2-(p+q)}{p^2}=\frac{q}{p^2}.$$

Отже,

$$D(X)=\frac{q}{p^2}.$$

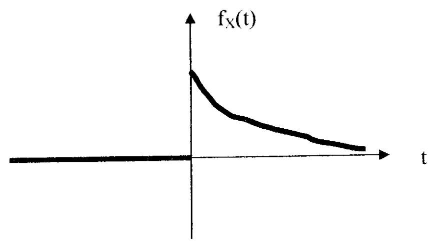
Розподіли неперервних випадкових величин

1.5 Показниковий (або експоненціальний) розподіл

Означення. Неперервна випадкова величина **X** розподілена за показниковим законом, якщо її щільність має вигляд

$$f_X(t)=\begin{cases}0, & t\leq 0\\ \mu e^{-\mu t}, & t>0,\ (\mu>0)\end{cases}.$$

Графік щільності $f_X(t)$ такий:



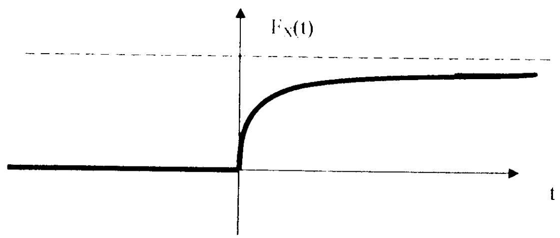
Інтегральна функція розподілу дорівнює

$$F_X(t)=\int_{-\infty}^{t}f(z)dz=\int_{0}^{t}f(z)dz=\int_{0}^{t}\mu\,\ell^{-\mu z}dz=\Big|_{0}^{t}\left(-\ell^{-\mu t}\right)=1-\ell^{-\mu t}.$$

Таким чином,

$$F_X(t)=\begin{cases}0, & t\le 0\\ 1-\ell^{-\mu t}, & t>0\end{cases}$$

і її графік такий:



Зауваження. Показниковий розподіл є неперервним аналогом геометричного розподілу і єдиний з усіх абсолютно неперервних випадкових величин задовольняє властивості відсутності післядії, тобто:

$$P\left(X\ge s+t\,\big/\,X\ge s\right)=P(X\ge t).$$

Дійсно, умовна ймовірність дорівнює

$$P\left(X\ge s+t\,\big/\,X\ge s\right)=\frac{P((X\ge s+t)\cdot(X\ge s))}{P(X\ge s)}=\frac{P(X\ge s+t)}{P(X\ge s)}=\frac{P(s+t\le X<+\infty)}{P(s\le X<+\infty)}=$$

$$=\frac{1-P(X<s+t)}{1-P(X<s)}=\frac{1-F_X(s+t)}{1-F_X(s)}=\frac{\ell^{-\mu(s+t)}}{\ell^{-\mu s}}=\ell^{-\mu t}=P(X\ge t).$$

Розглянемо, чи виконується умова нормування, тобто чи $\int_{-\infty}^{+\infty}f(t)dt=1$?

$$\int_{-\infty}^{+\infty}f(t)dt=\int_{0}^{+\infty}\mu\,\ell^{-\mu t}dt=\Big|_{0}^{+\infty}\left(-\ell^{-\mu t}\right)=1.$$

Обчислимо далі числові характеристики показникового розподілу:

$$M(X)=\int_{-\infty}^{+\infty}t\cdot f(t)dt=\int_{0}^{+\infty}t\cdot\mu\,\ell^{-\mu t}dt=\Big|_{0}^{+\infty}\left(t\left(-\frac{\mu\,\ell^{-\mu t}}{\mu}\right)+\int\ell^{-\mu t}dt\right)=\Big|_{0}^{+\infty}\left(-t\cdot\ell^{-\mu t}-\frac{\ell^{-\mu t}}{\mu}\right)=$$

$$=-\lim_{t\to+\infty}\left(\frac{t}{\ell^{\mu t}}\right)-\frac{1}{\mu}\left(\ell^{-\mu\infty}-\ell^{0}\right)=-\lim_{t\to+\infty}\left(\frac{1}{\mu\,\ell^{\mu t}}\right)+\frac{1}{\mu}=\frac{1}{\mu}.$$

Таким чином, математичне сподівання в показниковому розподілі таке:

$$M(X)=\frac{1}{\mu}.$$

Для обчислення дисперсії застосовуємо формулу D(X) = M(X^2)-[M(X)]^2:

$$M(X)=\int_{-\infty}^{+\infty}t^2\cdot f(t)dt=\int_{0}^{+\infty}t^2\cdot\mu\,\ell^{-\mu t}dt=\begin{vmatrix}u=t^2 & du=2tdt\\ dv=\mu\,\ell^{-\mu t}dt & v=-\ell^{-\mu t}\end{vmatrix}=\Big|_{0}^{+\infty}\left(-t^2\ell^{-\mu t}+2\int t\cdot\ell^{-\mu t}\right)=$$

$$=\begin{vmatrix}u=t & du=dt\\ dv=\ell^{-\mu t}dt & v=-\dfrac{\ell^{-\mu t}}{\mu}\end{vmatrix}=2\Big|_{0}^{+\infty}\left(-\frac{t\,\ell^{-\mu t}}{\mu}+\frac{1}{\mu}\int\ell^{-\mu t}dt\right)=\frac{2}{\mu}\cdot\Big|_{0}^{+\infty}\left(-\frac{\ell^{-\mu t}}{\mu}\right)=\frac{2}{\mu^2}.$$

Але тоді

$$D(X)=\frac{2}{\mu^2}-\frac{1}{\mu^2}=\frac{1}{\mu^2}.$$

Зауваження. Параметр μ показникового розподілу тісно пов'язаний з математичним сподіванням і дисперсією

$$M(X)=\frac{1}{\mu},\quad D(X)=\frac{1}{\mu^2},$$

тобто, якщо відомий параметр μ, тоді відомий і повністю показниковий розподіл абсолютно неперервної випадкової величини X.

Показниковий розподіл належить до **однопараметричних** розподілів.

1.6 Нормальний закон розподілу (закон Гаусса)

Означення. Нормально розподіленою випадковою величиною з параметрами a, σ називають абсолютно неперервну випадкову величину **X**, яка має щільність розподілу

$$f(x) = \frac{1}{\sqrt{2\pi}\cdot\sigma}\,\ell^{-\frac{(x-a)^2}{2\sigma^2}}.$$

З'ясуємо імовірнісний зміст параметрів a **і** σ, для чого обчислимо математичне сподівання M(X) і дисперсію D(X).

Але спочатку перевіримо,що виконується умова нормування

$$\int_{-\infty}^{+\infty} f(x)dx = 1.$$

$$\int_{-\infty}^{+\infty} f(x)dx = \frac{1}{\sqrt{2\pi}\cdot\sigma}\int_{-\infty}^{+\infty}\ell^{-\frac{(x-a)^2}{2\sigma^2}}dx = \left|\begin{array}{cc}\frac{x-a}{\sigma}=t & \frac{dx}{\sigma}=dt\\ \int_{-\infty}^{+\infty}\approx & \int_{-\infty}^{+\infty}\end{array}\right| = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\ell^{-\frac{t^2}{2}}dt.$$

Останній інтеграл обчислимо так: розглянемо подвійний інтеграл по всій площині XOY

$$\iint_{XOY}\ell^{-\frac{x^2+y^2}{2}}dxdy = \int_{-\infty}^{+\infty}\ell^{-\frac{x^2}{2}}dx\cdot\int_{-\infty}^{+\infty}\ell^{-\frac{y^2}{2}}dy = \left(\int_{-\infty}^{+\infty}\ell^{-\frac{x^2}{2}}dx\right)^2.$$

З іншого боку, переходячи до полярних координат, обчислюємо подвійний інтеграл, а саме:

$$\iint_{XOY}\ell^{-\frac{x^2+y^2}{2}}dxdy = \iint_{XOY}\ell^{-\frac{r^2}{2}}rdrd\varphi = \int_0^{2\pi}\left\{\int_0^{+\infty}\ell^{-\frac{r^2}{2}}rdr\right\}d\varphi = -\int_0^{2\pi}\left\{\left.\ell^{-\frac{r^2}{2}}\right|_0^{+\infty}\right\}d\varphi = \int_0^{2\pi}d\varphi = 2\pi.$$

Тобто

$$\left(\int_{-\infty}^{+\infty}\ell^{-\frac{x^2}{2}}dx\right)^2 = 2\pi \quad\Rightarrow\quad \int_{-\infty}^{+\infty}\ell^{-\frac{x^2}{2}}dx = \sqrt{2\pi}.$$

Таким чином, умова нормування виконується

$$\int_{-\infty}^{+\infty} f(x)dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\ell^{-\frac{t^2}{2}}dt = \frac{1}{\sqrt{2\pi}}\sqrt{2\pi} = 1.$$

Обчислюємо математичне сподівання

$$M(X) = \int_{-\infty}^{+\infty}xf(x)dx = \frac{1}{\sqrt{2\pi}\cdot\sigma}\int_{-\infty}^{+\infty}x\,\ell^{-\frac{(x-a)^2}{2\sigma^2}}dx = \frac{1}{\sqrt{2\pi}\cdot\sigma}\int_{-\infty}^{+\infty}[(x-a)+a]\,\ell^{-\frac{(x-a)^2}{2\sigma^2}}dx =$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\frac{x-a}{\sigma}\cdot\ell^{-\frac{1}{2}\left(\frac{x-a}{\sigma}\right)^2}\sigma d\left(\frac{x-a}{\sigma}\right) + \frac{a}{\sqrt{2\pi}\cdot\sigma}\int_{-\infty}^{+\infty}\ell^{-\frac{(x-a)^2}{2\sigma^2}}dx = a\int_{-\infty}^{+\infty}f(x)dx = a.$$

(Ми використали той факт, що в попередньому виразі перший доданок дорівнює нулю, як інтеграл в симетричних границях від непарної функції).

Таким чином, параметр a є математичним сподіванням

M(X) = a .

Далі обчислюємо дисперсію

$$D(X) = \int_{-\infty}^{+\infty}(x-a)^2 f(x)dx = \frac{1}{\sqrt{2\pi}\cdot\sigma}\cdot\sigma^2\int_{-\infty}^{+\infty}\left(\frac{x-a}{\sigma}\right)^2\ell^{-\frac{(x-a)^2}{2\sigma^2}}dx = \left|\begin{array}{cc}\frac{x-a}{\sigma}=t & \frac{dx}{\sigma}=dt\\ \int_{-\infty}^{+\infty}\approx & \int_{-\infty}^{+\infty}\end{array}\right| =$$

$$= \frac{1}{\sqrt{2\pi}}\sigma^2\int_{-\infty}^{+\infty}t^2\ell^{-\frac{t^2}{2}}dt.$$

Останній інтеграл інтегруємо за формулою інтегрування частинами, а саме:

$$\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}t^2\cdot\ell^{-\frac{t^2}{2}}dt = \left|\begin{array}{cc}u=t & du=dt\\ dv=t\ell^{-\frac{t^2}{2}}dt & v=-\ell^{-\frac{t^2}{2}}\end{array}\right| = \frac{\sigma^2}{\sqrt{2\pi}}\left(\left.\left(-t\ell^{-\frac{t^2}{2}}\right)\right|_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty}\ell^{-\frac{t^2}{2}}dt\right) = \sigma^2\cdot\frac{1}{\sqrt{2\pi}}\cdot\sqrt{2\pi} = \sigma^2,$$

тобто

$$D(X) = \sigma^2.$$

Отже, з'ясовано теоретико-ймовірнісний зміст параметрів a, σ : a є математичним сподіванням, σ^2 - дисперсією, а σ - середнім квадратичним відхиленням випадкової величини, розподіленої за нормальним законом.

Таким чином, **нормальний закон розподілу повністю визначається своїм математичним сподіванням і дисперсією.**

Інтегральна функція нормального закону дорівнює

$$F(x)=\int_{-\infty}^{x} f(t)dt=\frac{1}{\sqrt{2\pi}\cdot\sigma}\int_{-\infty}^{x} \ell^{-\frac{1}{2}\left(\frac{t-a}{\sigma}\right)^2}dt=\left|\begin{matrix}\frac{t-a}{\sigma}=z & \frac{dt}{\sigma}=dz\\ \int_{-\infty}^{x}\approx & \int_{-\infty}^{\frac{x-a}{\sigma}}\end{matrix}\right|=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{x-a}{\sigma}} \ell^{-\frac{z^2}{2}}dz=\Phi^*\left(\frac{x-a}{\sigma}\right),$$

де

$$\Phi^*(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\ell^{-\frac{z^2}{2}}dz$$ - **інтеграл ймовірності.**

Ймовірність попадання випадкової величини X в інтервал (α,β) можна обчислити за формулою

$$P(\alpha<X<\beta)=F(\beta)-F(\alpha),$$

або

$$\boxed{P(\alpha<X<\beta)=\Phi^*\left(\frac{\beta-a}{\sigma}\right)-\Phi^*\left(\frac{\alpha-a}{\sigma}\right)}$$

Нормальна крива (крива Гаусса)

Означення. Графік щільності ймовірності нормального закону розподілу називають **нормальною кривою (кривою Гаусса**).

Досліджуємо функцію

$$y=f(x)=\frac{1}{\sqrt{2\pi}\cdot\sigma}\ell^{-\frac{1}{2}\left(\frac{x-a}{\sigma}\right)^2}$$

методами диференціального числення:

1) функція „у" визначена на усій числовій осі і приймає тільки додатні значення, тобто нормальна крива розташована над віссю OX;

2) $\lim\limits_{x\to\pm\infty} y=0$, тобто вісь OX є горизонтальною асимптотою нормальної кривої;

3) знайдемо екстремум функції. Для цього обчислимо похідну

$$y'=\frac{1}{\sqrt{2\pi}\cdot\sigma}\ell^{-\frac{1}{2}\left(\frac{x-a}{\sigma}\right)^2}\cdot\left[-\frac{1}{2\sigma^2}\cdot 2(x-a)\right]=-\frac{(x-a)}{\sqrt{2\pi}\cdot\sigma^3}\ell^{-\frac{1}{2}\left(\frac{x-a}{\sigma}\right)^2},$$

тобто критичною точкою буде x = a. Зліва від цієї точки $y'>0$, тобто функція монотонно зростає, справа - $y'<0$ і функція монотонно спадає. Таким чином, в точці x = a функція „у" має максимум і

$$y(a)=\frac{1}{\sqrt{2\pi}\cdot\sigma},$$

4) далі, за допомогою похідної другого порядку знайдемо інтервали опуклості, угнутості і точки перегину нормальної кривої

$$y''=\frac{1}{\sqrt{2\pi}\cdot\sigma^3}\left\{-\ell^{-\frac{1}{2}\left(\frac{x-a}{\sigma}\right)^2}-(x-a)\ell^{-\frac{1}{2}\left(\frac{x-a}{\sigma}\right)^2}\cdot\left[-\frac{x-a}{\sigma^2}\right]\right\}=-\frac{1}{\sqrt{2\pi}\sigma^5}\ell^{-\frac{1}{2}\left(\frac{x-a}{\sigma}\right)^2}\cdot\left\{\sigma^2-(x-a)^2\right\}=$$

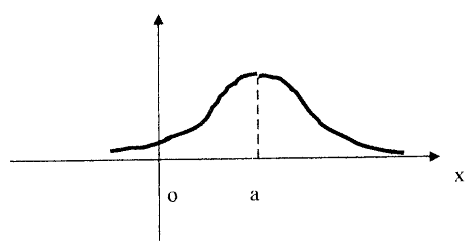
$$=\frac{1}{\sqrt{2\pi}\cdot\sigma^5}\left(x^2-2ax+a^2-\sigma^2\right)=\frac{1}{\sqrt{2\pi}\cdot\sigma^5}\left[x-(a+\sigma)\right]\cdot\left[x-(a-\sigma)\right],$$

тобто критичними точками другого роду є $x_1=a-\sigma$, $x_2=a+\sigma$. Саме в цих точках нормальна крива має перегин.

Розглянемо далі **вплив** параметрів розподілу на форму нормальної кривої:

а) нормальна крива симетрична відносно вертикальної прямої, яка проходить через точку x=a; крім того, графіки f(x) і f(x-a) мають однакову форму, тому математичне сподівання M(X) = a впливає тільки на місце центра симетрії, тобто на зсув нормальної кривої уздовж осі OX;

б) висота біля кола H = $\frac{1}{\sqrt{2\pi}\cdot\sigma}$ і тому зміна σ впливає на форму кривої, а саме графік спускається від максимуму тим крутіше, чим менше σ (загальна площа нескінченної трапеції завжди дорівнює одиниці).



Зокрема, ймовірність того, що X набуде значення з відрізка $[a-\varepsilon, a+\varepsilon]$, тим більше, чим менше σ, тобто чим менша дисперсія. Отже, дисперсія нормально розподіленої випадкової величини справді характеризує степінь розсіяння значень цієї величини.

Зауваження. Обчислимо ймовірність того, що X набуде значення з відрізка $[a-t\sigma, a+t\sigma]$, тобто $|X-a|\le t\sigma$ (t>0 – довільне). Ця ймовірність дорівнює

$$P(|X-a|\le t\sigma)=\int_{a-t\sigma}^{a+t\sigma} f(z)dz=\frac{1}{\sigma\sqrt{2\pi}}\int_{a-t\sigma}^{a+t\sigma}\ell^{-\frac{1}{2}\left(\frac{x-a}{\sigma}\right)^2}dx=\left|\begin{matrix} z=\frac{x-a}{\sigma} & dz=\frac{dx}{\sigma} \\ \int_{a-t\sigma}^{a+t\sigma}\approx & \int_{-t}^{t}\end{matrix}\right|=\frac{1}{\sqrt{2\pi}}\int_{-t}^{t}\ell^{-\frac{z^2}{2}}dz=$$

$$=\Phi^*(t)-\Phi^*(-t)=2\Phi(t),$$

де $\Phi(t)=\frac{1}{\sqrt{2\pi}}\int_0^t \ell^{-\frac{z^2}{2}}dz$ - інтеграл ймовірності (функція Лапласа).

Покладаючись в останній рівності послідовно. що t =1, 2, 3, за допомогою таблиці функції Лапласа отримаємо:

$$P(|X-a|\le\sigma)=2\Phi(1)\approx 0{,}68,$$

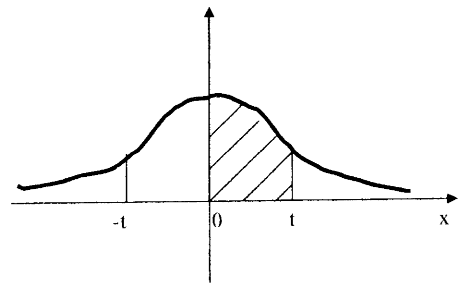
$$P(|X-a|\le 2\sigma)=2\Phi(2)\approx 0{,}95,$$

$$P(|X-a|\le 3\sigma)=2\Phi(3)\approx 0{,}997.$$

Число 0,997 мало відрізняється від одиниці, тому подію з ймовірністю 0,997 можна вважати практично вірогідною, і останню рівність часто формулюють у вигляді так званого ***„правила трьох сигм"***: **практично вірогідно, що нормально розподілена випадкова величина X може відхилятися від свого математичного сподівання M(X) = a не більше, ніж на потроєне середнє квадратичне відхилення.**

Властивості функції Лапласа

Властивості функції Лапласа (інтеграла ймовірності) можна дуже просто дослідити за допомогою графіка функції $y=\frac{1}{\sqrt{2\pi}}\ell^{-\frac{x^2}{2}}$



Функція $\Phi(t)=\frac{1}{\sqrt{2\pi}}\int_0^t \ell^{-\frac{x^2}{2}}dx$ - є площею заштрихованої криволінійної трапеції:

1) $\Phi(-t)=\frac{1}{\sqrt{2\pi}}\int_0^{-t}\ell^{-\frac{x^2}{2}}dx=-\frac{1}{\sqrt{2\pi}}\int_{-t}^{0}\ell^{-\frac{x^2}{2}}dx=-\Phi(t)$, тобто

$\Phi(-t)=-\Phi(t)$ - непарна функція ;

2) $\Phi(+\infty)=\frac{1}{2}$;

3) $\Phi^*(t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t}\ell^{-\frac{x^2}{2}}dx=\Phi(t)+\frac{1}{2}$;

4) $\Phi(t)$ - табульована.

1.7 Рівномірний розподіл

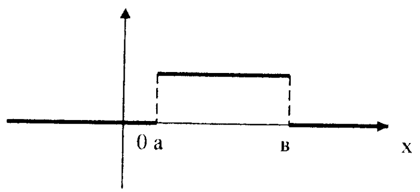
Означення. Абсолютно неперервна випадкова величина X розподілена за рівномірним розподілом, якщо її щільність має вигляд

$$f(x)=\begin{cases}\frac{1}{b-a} & x\in[a,b]\\ 0 & x\notin[a,b]\end{cases}$$

Очевидно, $f(x)\ge 0$ і виконується умова нормування

$$\int_{-\infty}^{+\infty} f(x)dx=\int_a^b\frac{1}{b-a}dx=\frac{1}{b-a}\cdot(b-a)=1.$$

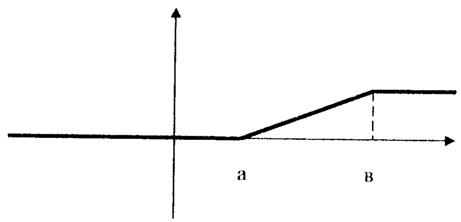
Графік щільності має вигляд



Інтегральна функція розподілу дорівнює

$$F(x)=\begin{cases}0 & -\infty<x<a\\ \dfrac{x-a}{b-a} & a\le x\le b\\ 1 & b<x<+\infty\end{cases}$$

і має графік



Обчислимо далі ймовірність попадання в інтервал (α, β); якщо цей інтервал належить сегменту [a , b] ,тоді

$$P(\alpha<X<\beta)=\int_\alpha^\beta \frac{1}{b-a}dx=\frac{\beta-\alpha}{b-a},$$

тобто ймовірність події $(\alpha<X<\beta)$ не залежить від розташування інтервалу $(\alpha, \beta)\in[a, b]$, а залежить лише від довжини цього інтервалу.

Знайдемо математичне сподівання M(X):

$$M(X)=\int xf(x)dx=\int_a^b x\frac{1}{b-a}dx=\left|_a^b \frac{x^2}{2}=\frac{1}{2}\cdot\frac{b^2-a^2}{b-a}=\frac{a+b}{2}.\right.$$

Дисперсія X дорівнює

$$D(X)=M(X^2)-M^2(X)=\int_{-\infty}^{+\infty}x^2f(x)dx-\left(\frac{a+b}{2}\right)^2=\int_a^b x^2\cdot\frac{1}{b-a}dx-\left(\frac{a+b}{2}\right)^2=$$

$$=\frac{1}{b-a}\cdot\left|_a^b\frac{x^3}{3}\right.-\left(\frac{a+b}{2}\right)^2=\frac{b^3-a^3}{3(b-a)}-\left(\frac{a+b}{2}\right)^2=\frac{b^2+ab+a^2}{3}-\frac{a^2+2ab+b^2}{4}=\frac{a^2+b^2-2ab}{12}=$$

$$=\frac{(b-a)^2}{12}.$$

Таким чином, числові характеристики рівномірного розподілу такі:

$$\boxed{M(X)=a+b/2,\ D(X)=(b-a)^2/12.}$$

1.8 Розподіл функції випадкової величини

Нехай X і Y – випадкові величини, пов'язані між собою функціональною залежністю Y = f(X) і нехай закон розподілу X відомий, а саме відомі $F_X(x)$, $f_X(x)=F_X'(x)$.

Виникає питання: як знайти закон розподілу функції Y = f(X) ?

1.8.1 Закон розподілу Коші

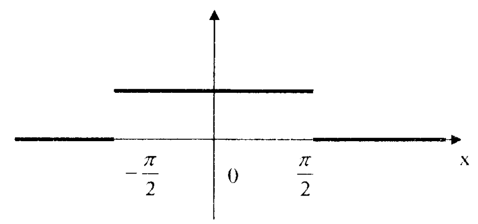
Нехай φ - випадкова величина, рівномірно розподілена на інтервалі $\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$:

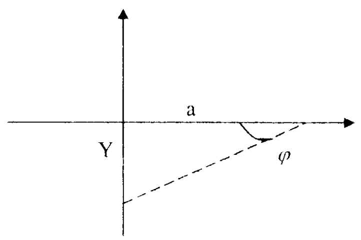
$$p_\varphi(x)=f_\varphi(x)=\begin{cases}\dfrac{1}{\pi}, & x\in\left(-\dfrac{\pi}{2},\dfrac{\pi}{2}\right)\\ 0, & x\notin\left(-\dfrac{\pi}{2},\dfrac{\pi}{2}\right)\end{cases}.$$

Розглянемо нову випадкову величину

$$Y=a\,tg\,\varphi\ .$$

Графік щільності випадкової величини φ і геометричний зв'язок між φ і Y наведені далі:





Очевидно,

$$F_Y(x) = P(Y < x) = P(atg\varphi < x) = P\left(\varphi < arctg\,\frac{x}{a}\right) = F_\varphi\left(arctg\,\frac{x}{a}\right).$$

Таким чином, інтегральна функція розподілу дорівнює

$$F_Y(x) = F_\varphi\left(arctg\,\frac{x}{a}\right).$$

Знайдемо далі щільність ймовірності

$$p_Y(x) = F_Y'(x) = \left\{F_\varphi\left(arctg\,\frac{x}{a}\right)\right\}'_x = F_\varphi'\left(arctg\,\frac{x}{a}\right)\cdot\frac{\frac{1}{a}}{1+\left(\frac{x}{a}\right)^2} = \frac{a}{a^2+x^2}\cdot F_\varphi'\left(arctg\,\frac{x}{a}\right) =$$

$$\frac{a}{a^2+x^2}f_\varphi\left(arctg\,\frac{x}{a}\right) = \frac{a}{a^2+x^2}\cdot\frac{1}{\pi},$$

тобто щільність (або диференціальна функція розподілу) в розподілі Коші така

$$p_Y(x) = \frac{a}{\pi\left(a^2+x^2\right)}.$$

Обчислимо математичне сподівання цього розподілу

$$M(Y) = \int_{-\infty}^{+\infty} xp_Y(x)dx = \int_{-\infty}^{+\infty}\frac{ax}{\pi\left(a^2+x^2\right)}dx = \left.\frac{a}{2\pi}\ln\left(x^2+a^2\right)\right|_{-\infty}^{0} + \left.\frac{a}{2\pi}\ln\left(x^2+a^2\right)\right|_{0}^{+\infty} \to \infty,$$

тобто числові характеристики в законі Коші такі:

$$\underline{M(Y) = +\infty,\quad D(Y) = +\infty.}$$

1.8.2 *Розподіл χ^2 з однією степінню волі*

Так називається випадкова величина, яка дорівнює

$$\chi^2 = X^2,$$

де X – стандартна нормальна випадкова величина з M(X) =0 і дисперсією D(X) = 1, тобто

$$p_X(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},\qquad f(X) = X^2.$$

Знайдемо

$$F_{\chi^2}(x),\quad p_{\chi^2}(x):$$

$$F_{\chi^2}(x) = P\left(\chi^2 < x\right) = P\left(X^2 < x\right) = \begin{cases} P\left(|X| < \sqrt{x}\right), & x > 0 \\ 0, & x < 0 \end{cases} = \begin{cases} P\left(-\sqrt{x} < X < \sqrt{x}\right), & x > 0 \\ 0, & x < 0 \end{cases} =$$

$$= \begin{cases} F_X\left(\sqrt{x}\right) - F_X\left(-\sqrt{x}\right), & x > 0 \\ 0, & x < 0 \end{cases},$$

а

$$p_{\chi^2}(x) = F'_{\chi^2}(x) = \begin{cases} \frac{1}{2\sqrt{x}}\left(F_X'\left(\sqrt{x}\right) + F_X'\left(-\sqrt{x}\right)\right), & x > 0 \\ 0, & x < 0 \end{cases} = \begin{cases} \frac{1}{2\sqrt{x}}\left(p_X\left(\sqrt{x}\right) + p_X\left(-\sqrt{x}\right)\right), & x > 0 \\ 0, & x < 0 \end{cases}$$

тобто

$$p_{\chi^2}(x) = \begin{cases} \frac{1}{\sqrt{2\pi}\cdot\sqrt{x}}e^{-\frac{x}{2}}, & x > 0 \\ 0, & x < 0 \end{cases}.$$

Обчислимо математичне сподівання та дисперсію:

$$M(\chi^2)=\int_{-\infty}^{+\infty} x p_{\chi^2}(x)dx=\int_0^{+\infty}\frac{x}{\sqrt{2\pi}\cdot\sqrt{x}}\ell^{-\frac{x}{2}}dx=\int_0^{+\infty}\frac{\sqrt{x}}{\sqrt{2\pi}}\ell^{-\frac{x}{2}}dx=\left|\begin{matrix} x=t^2 & dx=2tdt \\ \int\limits_0^{+\infty} & \int\limits_0^{+\infty}\end{matrix}\right| \int_0^{+\infty}\frac{2}{\sqrt{2\pi}}t^2\ell^{-\frac{t^2}{2}}dt=$$

$$=\frac{2}{\sqrt{2\pi}}\left(\Big|_0^{+\infty} t\left(-\ell^{-\frac{t^2}{2}}\right)+\int_0^{+\infty}\ell^{-\frac{t^2}{2}}dt\right)=2\Phi(+\infty)=2\cdot\frac{1}{2}=1;$$

$$D(\chi^2)=M([\chi^2]^2)-[M(\chi^2)]^2=\int_{-\infty}^{+\infty}x^2 p_{\chi^2}(x)dx-1=\int_0^{+\infty}\frac{x^2}{\sqrt{2\pi}\cdot\sqrt{x}}\ell^{-\frac{x}{2}}dx-1=$$

$$=\left|\begin{matrix} u=\frac{\sqrt{x^3}}{\sqrt{2\pi}} & du=\frac{1}{\sqrt{2\pi}}\frac{3\sqrt{x}}{2}dx \\ dv=\ell^{-\frac{x}{2}}dx & v=-2\ell^{-\frac{x}{2}}\end{matrix}\right|=\frac{1}{\sqrt{2\pi}}\left(\Big|_0^{+\infty}\sqrt{x^3}\left(-2\ell^{-\frac{x}{2}}\right)+\int_0^{+\infty}\frac{3}{2}\sqrt{x}\cdot 2\ell^{-\frac{x}{2}}dx\right)-1=$$

$$=3\cdot\frac{1}{\sqrt{2\pi}}\int_0^{+\infty}\sqrt{x}\cdot\ell^{-\frac{x}{2}}dx-1=3M(\chi^2)-1=3-1=2.$$

Таким чином, числові характеристики в цьому розподілі такі:

$$\underline{M(\chi^2)=1;\qquad D(\chi^2)=2.}$$

1.9 Розподіл Лапласа

Означення. Випадкова величина X розподілена за законом Лапласа, якщо її щільність така:

$$p_X(x)=a\ell^{-\lambda|x|},\quad (\lambda>0).$$

Задача дослідження цього розподілу формулюється так. Треба:

1) визначити коефіцієнт „a";
2) побудувати графік щільності ймовірності $p_X(x)$;
3) визначити інтегральну функцію розподілу $F_X(x)$ і побудувати її графік;
4) знайти математичне сподівання M(X) і дисперсію D(X).

Для визначення сталої „a" застосуємо умову нормування $\int_{-\infty}^{+\infty}p_X(x)dx=1$:

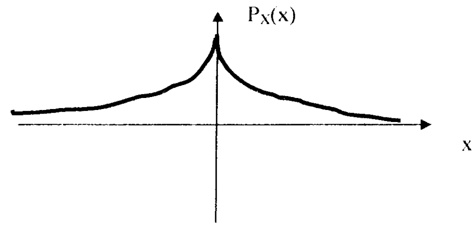
$$1=\int_{-\infty}^{+\infty}a\ell^{-\lambda|x|}dx=a\int_{-\infty}^{0}\ell^{\lambda x}dx+a\int_0^{+\infty}\ell^{-\lambda x}dx=a\Big|_{-\infty}^{0}\frac{\ell^{\lambda x}}{\lambda}-a\Big|_0^{+\infty}\frac{\ell^{-\lambda x}}{\lambda}=a\left(\frac{1}{\lambda}-\frac{\ell^{-\infty}}{\lambda}\right)-a\left(\frac{\ell^{-\infty}}{\lambda}-\frac{1}{\lambda}\right)=\frac{2a}{\lambda}\Rightarrow$$

$$\underline{a=\frac{\lambda}{2}.}$$

Таким чином, розподіл Лапласа має вигляд:

$$\underline{p_X(x)=\frac{\lambda}{2}\ell^{-\lambda|x|},\quad (\lambda>0).}$$

Графік щільності $p_X(x)=\begin{cases}\frac{\lambda}{2}\ell^{\lambda x}, & (x<0)\\ \frac{\lambda}{2}\ell^{-\lambda x}, & (x>0)\end{cases}$ такий:



Визначимо далі інтегральну функцію розподілу за формулою

$$F_X(x)=\int_{-\infty}^{x}p_X(t)dt.$$

Якщо $-\infty<x<0$

$$F_X(x)=\int_{-\infty}^{x}\frac{\lambda}{2}\ell^{\lambda t}dt=\frac{1}{2}\Big|_{-\infty}^{x}\ell^{\lambda t}=\frac{1}{2}\left(\ell^{\lambda x}-\ell^{-\infty}\right)=\frac{1}{2}\ell^{\lambda x};$$

якщо $0<x<+\infty$

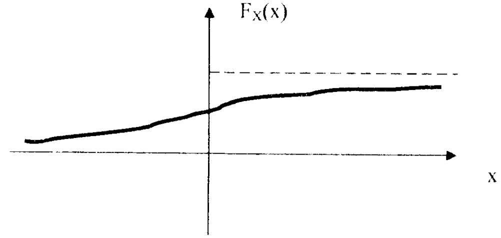
$$F_X(x)=\int_{-\infty}^{x}p_X(t)dt=\int_{-\infty}^{0}p_X(t)dt+\int_{0}^{x}p_X(t)dt=\frac{1}{2}+\int_{0}^{x}\frac{\lambda}{2}e^{-\lambda t}dt=\frac{1}{2}-\frac{1}{2}e^{-\lambda t}\Big|_{0}^{x}=\frac{1}{2}-\frac{1}{2}e^{-\lambda x}+\frac{1}{2}=$$

$$=1-e^{-\lambda x}.$$

Таким чином, інтегральна функція розподілу Лапласа така

$$F_X(x)=\begin{cases}\frac{1}{2}e^{\lambda x}, & (x<0),\\ 1-\frac{1}{2}e^{-\lambda x}, & (x>0).\end{cases}$$

Графік інтегральної функції розподілу має вигляд:



Обчислимо математичне сподівання і дисперсію:

$$M(X)=\int_{-\infty}^{+\infty}xp_X(x)dx=\int_{-\infty}^{0}x\frac{\lambda}{2}e^{\lambda x}dx+\int_{0}^{+\infty}x\frac{\lambda}{2}e^{-\lambda x}dx=\left|\begin{matrix}x=-t & dx=-dt\\ \int_{-\infty}^{0}\approx & \int_{+\infty}^{0}\end{matrix}\right|=$$

$$=-\int_{0}^{+\infty}t\frac{\lambda}{2}e^{-\lambda t}dt+\int_{0}^{+\infty}x\frac{\lambda}{2}e^{-\lambda x}dx=0,$$

тобто

<u>M(X) = 0.</u>

Для обчислення дисперсії застосуємо формулу

$D(X) = M(X^2) - [M(X)]^2$:

$$D(X)=M(X^2)=\int_{-\infty}^{+\infty}x^2p_X(x)dx=\int_{-\infty}^{0}x^2\frac{\lambda}{2}e^{\lambda x}dx+\int_{0}^{+\infty}x^2\frac{\lambda}{2}e^{-\lambda x}dx=\left|\begin{matrix}u=x^2 & du=2xdx\\ dv=e^{\pm\lambda x}dx & v=\pm\frac{e^{\pm\lambda x}}{\lambda}\end{matrix}\right|=$$

$$=\Big|_{-\infty}^{0}\left(x^2\frac{1}{2}e^{\lambda x}-\int xe^{\lambda x}dx\right)+\Big|_{0}^{+\infty}\left(-x^2\frac{1}{2}e^{-\lambda x}+\int xe^{-\lambda x}dx\right)=\left|\begin{matrix}u=x & du=dx\\ dv=e^{\pm\lambda x}dx & v=\pm\frac{e^{\pm\lambda x}}{\lambda}\end{matrix}\right|=$$

$$=-\Big|_{-\infty}^{0}\left(x\frac{e^{\lambda x}}{\lambda}-\int\frac{e^{\lambda x}}{\lambda}dx\right)+\Big|_{0}^{+\infty}\left(x\left(-\frac{e^{-\lambda x}}{\lambda}\right)+\int\frac{e^{-\lambda x}}{\lambda}dx\right)=\Big|_{-\infty}^{0}\frac{e^{\lambda x}}{\lambda^2}-\Big|_{0}^{+\infty}\frac{e^{-\lambda x}}{\lambda^2}=\frac{1}{\lambda^2}+\frac{1}{\lambda^2}=\frac{2}{\lambda^2}.$$

Отже,

$$\underline{M(X)=0,\quad D(X)=\frac{2}{\lambda^2}.}$$

1.10 Граничні теореми теорії ймовірностей

1.10.1 Означення збіжності за ймовірністю

Нехай X_1, X_2, ... , X_n, ... нескінченна послідовність випадкових величин.

Якщо для будь-якого додатного $\varepsilon>0$

$$P(|X_n-a|>\varepsilon)\to 0 \text{ при умові, що } n\to+\infty,$$

тоді кажуть, що послідовність $\{X_n\}$ збігається за ймовірністю до невипадкової величини a і позначають цей факт так

$$X_n\to a \text{ за ймовірністю,}$$

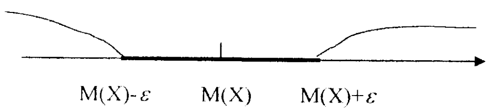
або

$$\lim_{n\to\infty}X_n=a \text{ за ймовірністю.}$$

1.10.2 Нерівність Чебишева

$$P(|X - m_X| > \varepsilon) \le \frac{D(X)}{\varepsilon^2}, \quad (\forall \varepsilon > 0).$$

Доведення: Нехай X – випадкова величина, M(X) – її математичне сподівання. Оберемо $\forall \varepsilon > 0$ і побудуємо на числовій осі OX інтервал $(M(X)-\varepsilon, M(X)+\varepsilon)$



Позначимо через D_X множину тих "x", для яких $|X - M(X)| > \varepsilon$. Обчислимо дисперсію

$$D(X) = \int_{-\infty}^{+\infty} (x - M(X))^2 p_X(x)dx \ge \int_{D_X} (x - M(X))^2 p_X(x)dx \ge \varepsilon^2 \int_{D_X} p_X(x)dx = \varepsilon^2 P(X \in D_X) =$$

$$= \varepsilon^2 P(|X - M(X)| > \varepsilon),$$

тобто

$$D(X) \ge \varepsilon^2 P(|X - M(X)| > 0),$$

або

$$P(|X - M(X)| > \varepsilon) \le \frac{D(X)}{\varepsilon^2}.$$

1.10.3 Закон великих чисел (ЗВЧ) (Чебишев, Хінчин)

Нехай $X_1, X_2, \dots, X_n, \dots$ - послідовність однаково розподілених випадкових величин, причому $M(X_k) = a$, $D(X_k) < +\infty$. Тоді середнє арифметичне збігається за ймовірністю до числа a, тобто

$$P\left(\left|\frac{X_1 + X_2 + \dots + X_n}{n} - a\right| > \varepsilon\right) \to 0, \quad n \to +\infty.$$

Доведення. Застосуємо нерівність Чебишева для середнього арифметичного

$$P\left(\left|\frac{X_1 + X_2 + \dots + X_n}{n} - a\right| > \varepsilon\right) \le \frac{D\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right)}{\varepsilon^2} = \frac{1}{n^2\varepsilon^2} \cdot nD(X_1) = \frac{D(X_1)}{n\varepsilon^2}$$

і, якщо

$$n \to +\infty \quad \Rightarrow \quad \frac{D(X)}{n\varepsilon^2} \to 0,$$

тобто

$$P\left(\left|\frac{X_1 + X_2 + \dots + X_n}{n} - a\right| > \varepsilon\right) \to 0, \quad n \to +\infty.$$

1.10.4 Застосування ЗВЧ

Нехай відбувається вимірювання деякої величини. Тоді, як <u>наближене значення цієї величини</u>, згідно з ЗВЧ, треба використовувати <u>середнє арифметичне спостережених значень.</u>

Дійсно, якщо розглядати спостережені значення як випадкові величини $X_1, X_2, \dots, X_n$ з рівними математичними сподіваннями $M(X_k) = a$ і дисперсіями $D(X_k) < +\infty$, причому $M(X_k) = a$ – дійсне значення вимірюваної величини, ми, за ЗВЧ, отримаємо, що середнє арифметичне збігається за ймовірністю до „a"

$$\frac{X_1 + X_2 + \dots + X_n}{n} \to a, \quad n \to +\infty \text{ за ймовірністю,}$$

тобто, якщо n>>1, то

$$\frac{X_1 + X_2 + \dots + X_n}{n} \approx a \text{ - за ймовірністю,}$$

якщо немає систематичних похибок.

1.10. 5 Центральна гранична теорема Ляпунова (ЦГТ)

Нехай $X_1, X_2, \dots, X_n$ – послідовність незалежних випадкових величин.

Розглянемо суму

$$S_n = X_1 + X_2 + \dots + X_n$$

і введемо нормовану випадкову величину

$$S_n^* = \frac{S_n - M(S_n)}{\sqrt{D(S_n)}}, \quad M(S_n^*) = 0, \quad D(S_n^*) = 1.$$

Тоді для будь-якого „x" функція розподілу

$$F_{S_n^*}(x) \to \frac{1}{2} + \Phi(x), \quad n \to +\infty,$$

або

$$F_{S_n^*}(x) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt.$$

Зауваження 1. ЦГТ стверджує, якщо деякий процес є сумарною дією множини незалежних випадкових величин, то його розподіл мало відрізняється від нормального, тобто теоретично доводить важливість і розповсюдженість у природі нормального розподілу.

Зауваження 2. Локальна і інтегральна теореми Мавра – Лапласа є прикладами застосування ЦГТ.

1.10. 6 Відхилення частоти від ймовірності. Теорема Бернуллі

Нехай μ_n - число появ події А в „n" випробуваннях Бернуллі; $\nu_n = \frac{\mu_n}{n}$ - частота появи події А; p = P(a) – ймовірність появи події А в одному випробуванні.

Якою буде ймовірність того, що відхилення частоти ν_n від ймовірності „p" не перевищує $\forall \varepsilon > 0$

$$P(|\nu_n - p| \le \varepsilon) = ?$$

Розглянемо цю ймовірність

$$P(|\nu_n - p| \le \varepsilon) = P\left(p - \varepsilon \le \frac{\mu_n}{n} \le p + \varepsilon\right) = P(np - n\varepsilon \le \mu_n \le np + n\varepsilon) = P(k_1 \le \mu_n \le k_2),$$

де $k_1 = np - n\varepsilon$, $k_2 = np + n\varepsilon$.

За інтегральною теоремою Муавра – Лапласа,

$$P(k_1 \le \mu_n \le k_2) = \Phi\left(\frac{k_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{k_1 - np}{\sqrt{npq}}\right) = \Phi\left(\frac{np + n\varepsilon - np}{\sqrt{npq}}\right) - \Phi\left(\frac{np - n\varepsilon - np}{\sqrt{npq}}\right) =$$

$$= \Phi\left(\frac{\sqrt{n}\varepsilon}{\sqrt{pq}}\right) - \Phi\left(-\frac{\sqrt{n}\varepsilon}{\sqrt{pq}}\right) = 2\Phi\left(\frac{\sqrt{n}\varepsilon}{\sqrt{pq}}\right)$$

Таким чином, має місце співвідношення

$$P(|\nu_n - p| \le \varepsilon) = 2\Phi\left(\frac{\sqrt{n}\varepsilon}{\sqrt{pq}}\right),$$

із якого випливає **формула і теорема Бернуллі :**

$$P(|\nu_n - p| \le \varepsilon) \to 1, \quad n \to +\infty, \quad \forall \varepsilon > 0.$$

Дійсно, якщо $n \to +\infty$, тоді $\Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right) \to \Phi(+\infty) = \frac{1}{2}$.

Зауваження. Теорема Бернуллі підтверджує визначення ймовірності за Мізесом

$$P(A) = \lim_{n \to \infty} \frac{\mu_n}{n}.$$

ТЕМА 2. ВИПАДКОВІ ВЕКТОРИ

2.1 Системи випадкових величин

В теорії ймовірностей і математичній статистиці часто зустрічається ситуація, коли декілька випадкових величин пов'язані між собою.

Означення. n – вимірним випадковим вектором (або системою „n" випадкових величин) називається вектор

$$\bar{\xi}(\xi_1, \xi_2, \cdots, \xi_n),$$

всі координати якого є випадковими величинами, визначеними на **загальному просторі елементарних подій деякого експерименту** (усі ξ_i з'являються в одному експерименті)

$$\xi_i = \xi_i(\omega), \quad \omega \in \Omega.$$

Приклад. Підкидаються разом 3 гральних кубики. Розглянемо випадкові величини:

ξ_1 - число очок, які випали на першому кубику;
ξ_2 - сума числа очок на усіх кубиках ;
ξ_3 - найбільше число очок.

$$\vec{\xi}(\xi_1, \xi_2, \xi_3)$$

тривимірний випадковий вектор заданий на просторі елементарних подій Ω, який складається з n = 6^3 елементарних подій.
Розглянемо усе більш детально.

2.2 Двовимірний випадковий вектор (ξ, η)

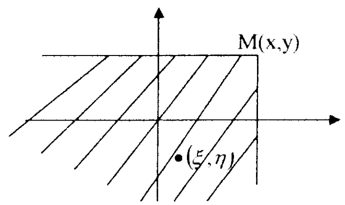
Означення. Розглянемо довільну точку M(x, y) на площині і подію, яка полягає в тому, що

$$(\xi < x; \eta < y).$$

Функцією розподілу випадкового вектора (ξ, η) називається ймовірність такої події:

$$\boxed{F_{\xi\eta}(x, y) = P(\xi < x; \eta < y)}$$

Геометричний зміст функції розподілу полягає в тому, що $F_{\xi\eta}(x, y)$ - ймовірність того, що випадкова точка (ξ, η) міститься у нескінченному прямокутнику з вершиною в точці M(x, y)



2.3 Властивості функції розподілу $F_{\xi\eta}(x, y)$

1) $0 \le F_{\xi\eta}(x, y) \le 1$;

2) $F_{\xi\eta}(-\infty, y) = 0$; $F_{\xi\eta}(x, -\infty) = 0$, тому що події $\xi < -\infty$, $\eta < -\infty$ неможливі, тому і сумісні події $(\xi < -\infty, \eta < y)$, $(\xi < x, \eta < -\infty)$ є подіями неможливими , і тому їх ймовірності дорівнюють нулю;

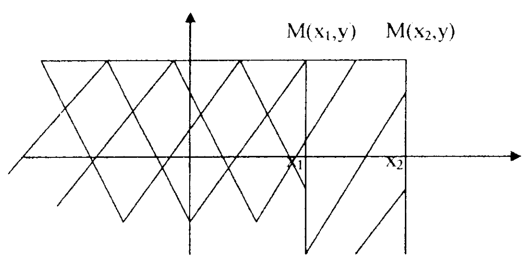
3) $F_{\xi\eta}(+\infty, +\infty) = 1$.

Події $\xi < +\infty$, $\eta < +\infty$ - вірогідні, а тому і подія $(\xi < +\infty, \eta < +\infty)$ вірогідна.

4) $F_{\xi\eta}(x, y)$ - функція неспадна за кожним з аргументів, а саме:

а) $\forall x_1, x_2 : x_1 < x_2 \Rightarrow F(x_1, y) \le F(x_2, y)$;
в) $\forall y_1, y_2 : y_1 < y_2 \Rightarrow F(x, y_1) \le F(x, y_2)$;

доведемо випадок а). Розглянемо на площині дві точки M(x_1,y) і M(x_2,y) і три події $(\xi < x_1, \eta < y)$, $(\xi < x_2, \eta < y)$, $(x_1 \le \xi < x_2, \eta < y)$



Очевидно, справедлива рівність

$$(\xi < x_2;\, \eta < y) = (\xi < x_1;\, \eta < y) + (x_1 \le \xi < x_2;\, \eta < y)\,,$$

чому доданки справа несумісні. Але тоді

$$P(\xi < x_2;\, \eta < y) = P(\xi < x_1;\, \eta < y) + P(x_1 \le \xi < x_2;\, \eta < y)\,. \qquad (*)$$

Маючи на увазі співвідношення

$$P(\xi < x_2;\, \eta < y) = F(x_2;\, y); \quad P(\xi < x_1;\, \eta < y) = F(x_1;\, y); \quad P(x_1 \le \xi < x_2;\, \eta < y) \ge 0,$$

отримуємо потрібну нерівність

$$F(x_1\,;\, y) \le F(x_2\,;\, y)\,.$$

Випадок в) доводиться аналогічно.

Зауваження. Співвідношення (*) допомагає сформулювати ще одне правило.

Ймовірність попадання випадкової точки $(\xi\,,\eta)$ у смугу, яка паралельна осі OY, дорівнює різниці значень сумісної функції розподілу на кінцях смуги

$$\boxed{P(x_1 \le \xi < x_2\,;\, \eta < y) = F(x_2\,,\, y) - F(x_1\,,\, y).}$$

Аналогічно можна знайти ймовірність попадання випадкової точки $(\xi\,,\eta)$ у смугу паралельну осі OX:

$$\boxed{P(\xi < x\,, y_1 \le \eta < y_2) = F(x\,,\, y_2) - F(x\,,\, y_1).}$$

а також ймовірність попадання випадкової точки $(\xi\,,\eta)$ у прямокутник

$$\boxed{P(x_1 \le \xi < x_2\,;\, y_1 \le \eta < y_2) = [F(x_2, y_2) - F(x_1, y_2)] - [F(x_2, y_1) - F(x_1, y_1)];}$$

5) функції розподілу координат випадкового вектора $(\xi\,,\eta)$: $F_\xi(x)$, $F_\eta(y)$:

$$F_\xi(x) = P(\xi < x) = P(\xi < x\,;\, \eta < +\infty) = F_{\xi\eta}(x\,,\, +\infty),$$

$$F_\eta(y) = P(\eta < y) = P(\xi < +\infty\,;\, \eta < y) = F_{\xi\eta}(+\infty\,,\, y)\,;$$

6) якщо випадкові величини ξ, η незалежні, тоді незалежними будуть і події $\xi < x$, $\eta < y$, а тоді ймовірність їх сумісної появи дорівнює добутку їх ймовірностей

$$P(\xi < x\,, \eta < y) = P(\xi < x) \cdot P(\eta < y),$$

тобто

$$\boxed{F_{\xi\eta}(x\,,\, y) = F_\xi(x) \cdot F_\eta(y).}$$

2.4 Дискретні випадкові вектори (ДВВ)

Означення. ДВВ називають випадковий вектор, координати якого виявляються дискретними випадковими величинами

$$(\xi\,,\eta) \;\Rightarrow \begin{cases} \xi : x_1, x_2, \cdots, x_m \\ \eta : y_1, y_2, \cdots, y_n \end{cases}$$

Зауваження. Із цього означення випливає, що множина значень, які приймає випадковий вектор (або випадкова точка), є множиною ізольованих точок площини.

Закон розподілу дискретного випадкового вектора має вигляд таблиці з двома входами, де введені такі позначення: $p_{ij} = P(\xi = x_i\,;\, \eta = y_j)$

η / ξ	y_1	y_2	…	y_n
x_1	p_{11}	p_{12}	…	p_{1n}
x_2	p_{21}	p_{22}	…	p_{2n}
…	…	…	…	…
x_m	p_{m1}	p_{m2}	…	p_{mn}

Очевидно,

$$\sum_{ij} p_{ij} = 1\,, \quad \begin{pmatrix} i = 1, 2, \cdots, m \\ j = 1, 2, \cdots, n \end{pmatrix}.$$

Приклад. Підкидаємо 2 монети. Будемо вважати, якщо випав герб – це 1, якщо – цифра – це 0. Розглянемо випадковий вектор $(\xi\,,\eta)$, де

ξ - число очок на першій монеті,

η - сума числа очок на обох монетах.

Побудуємо закон розподілу $(\xi\,,\eta)$

η / ξ	0	1	2
0	1/4	1/4	0
1	0	1/4	1/4

Усі $p_{ij} \ge 0$ і їх сума дорівнює одиниці $\sum_{i,j} p_{ij} = 1$.

2.5 Закони розподілу координат дискретного випадкового вектора (Моргинальні розподіли)

Нехай (ξ,η) - ДВВ. Визначимо $p_i = P(\xi = x_i)$, $p_j = P(\eta = y_j)$

Розглянемо подію $\xi = x_i$. Очевидно, ця подія відбувається одночасно з однією з подій $\eta = y_j$, тобто

$$(\xi = x_i) = (\xi = x_i)\cdot(\eta = y_1) + (\xi = x_i)\cdot(\eta = y_2) + \cdots + (\xi = x_i)\cdot(\eta = y_n).$$

За теоремою додавання ймовірностей несумісних подій, отримуємо

$$p_i = p_{i1} + p_{i2} + \cdots + p_{in} = \sum_{j=1}^{n} p_{ij}.$$

Аналогічно,

$$p_j = P(\eta = y_j) = P(\xi = x_1; \eta = y_j) + P(\xi = x_2; \eta = y_j) + \cdots + P(\xi = x_m; \eta = y_j) = \sum_{i=1}^{m} p_{ij}.$$

Таким чином,

$$(\xi)\quad p_i = \sum_{j=1}^{n} p_{ij}; \qquad (2.1)$$

$$(\eta)\quad p_j = \sum_{u=1}^{m} p_{ij}. \qquad (2.2)$$

η/ξ	y_1	y_2	…	y_n	ξ
x_1	p_{11}	p_{12}	…	p_{1n}	$\sum_j p_{1j}$
x_2	p_{21}	p_{22}	…	p_{2n}	$\sum_j p_{2j}$
…	…	…	…	…	…
x_m	p_{m1}	p_{m2}	…	p_{mn}	$\sum_j p_{jm}$
η	$\sum_i p_{i1}$	$\sum_i p_{i2}$	…	$\sum_i p_{in}$	

Приклад. Знайти закони розподілу координат ДВВ, якщо його закон розподілу має вигляд

η/ξ	-1	0	ξ
1	0,2	0,1	0,3
2	0,3	0,1	0,4
3	0,1	0,2	0,3
η	0,6	0,4	

Координати ДВВ такі:

ξ	1	2	3
P	0,3	0,4	0,3

η	-1	0
P	0,6	0,4

Зауваження. Якщо координати ДВВ незалежні, тоді

$$p_{ij} = P(\xi = x_i; \eta = y_j) = P(\xi = x_i)\cdot P(\eta = y_j) = p_i \cdot p_j.$$

Формули (2.1), (2.2) мають вигляд

$$p_i = \sum_{j=1}^{n} p_{ij} = \sum_{j=1}^{n} p_i \cdot p_j = p_i \sum_{j=1}^{n} p_j = p_i \cdot 1 = p_i,$$

$$p_j = \sum_{i=1}^{m} p_{ij} = \sum_{i=1}^{m} p_i \cdot p_j = p_j \sum_{i=1}^{m} p_i = p_j \cdot 1 = p_j.$$

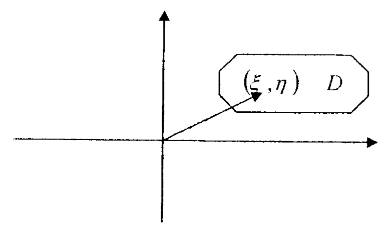
2.6 Неперервні випадкові вектори (НВВ)

Означення 1. Вектор (ξ,η) називається ***неперервним***, якщо його координати виявляються неперервними випадковими величинами.

Між неперервних векторів будемо розглядати лише ***абсолютно неперервні***, тобто такі, для яких ***існує щільність*** розподілу ймовірностей.

Означення 2. Нехай у площині (яка розглядається як двовимірний простір R^2) задана деяка область D. Невід'ємна функція $p_{\xi\eta}(x,y) \geq 0$ називається ***щільністю розподілу ймовірності*** випадкового вектора (ξ,η), якщо ймовірність події $(\xi,\eta)\in D$ $(D \in R^2)$ визначається рівністю

$$P((\xi,\eta)\in D)=\iint\limits_D p_{\xi\eta}(x,y)dxdy.$$



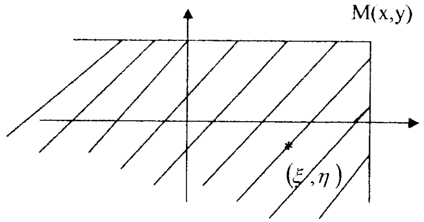
2.7 Властивості щільності розподілу

1) p(x, y) ≥ 0 ;
2) $\iint\limits_{R^2} p(x,y)dxdy=1$

(інтеграл по всій площині R^2 треба розглядати як невласний)

$$\iint\limits_{R^2} p(x,y)dxdy=\int\limits_{-\infty}^{+\infty}\left\{\int\limits_{-\infty}^{+\infty}p(x,y)dy\right\}dx;$$

3) Зв'язок між F(x,y) і p(x,y) випливає із геометричного змісту F(x,y)



$$F(x,y)=\int\limits_{-\infty}^{x}\left\{\int\limits_{-\infty}^{y}p(t,s)ds\right\}dt\ ;$$

4) $$p(x,y)=\frac{\partial^2 F}{\partial x\partial y}.$$

Дійсно,

$$\frac{\partial F(x,y)}{\partial x}=\int\limits_{-\infty}^{y}p(x,s)ds;\quad \frac{\partial^2 F(x,y)}{\partial x\partial y}=\frac{\partial}{\partial y}\left(\frac{\partial F}{\partial x}\right)=\frac{\partial}{\partial y}\left(\int\limits_{-\infty}^{y}p(x,s)ds\right)=p(x,y).$$

5) відшукання щільності розподілу координат випадкового вектора.

Нехай (ξ,η) - випадковий вектор, $p_{\xi\eta}(x,y)$ - його щільність розподілу ймовірностей. Якими будуть $p_\xi(x)$? $p_\eta(y)$?

$$p_\xi(x)=F_\xi'(x)=\{F_{\xi\eta}(x,y=+\infty)\}_x'=\left(\int\limits_{-\infty}^{x}dt\int\limits_{-\infty}^{+\infty}p_{\xi\eta}(t,s)ds\right)_x'=\int\limits_{-\infty}^{+\infty}p_{\xi\eta}(x,s)ds,$$

тобто

$$p_\xi(x)=\int\limits_{-\infty}^{+\infty}p_{\xi\eta}(x,s)ds;$$

аналогічно,

$$p_\eta(y)=\int\limits_{-\infty}^{+\infty}p_{\xi\eta}(t,y)dt.$$

Щільності розподілу ймовірностей координат випадкового вектора

$$p_\xi(x),\quad p_\eta(y)$$

називаються **моргинальними щільностями;**

6) розподіл випадкового вектора із незалежними координатами.

Функція розподілу зараз має вигляд

$$F_{\xi\eta}(x,y)=P(\xi<x;\eta<y)=P(\xi<x)\cdot P(\eta<y)=F_\xi(x)\cdot F_\eta(y).$$

Але тоді

$$p_{\xi\eta}(x,y)=\frac{\partial^2 F}{\partial x\partial y}=\frac{\partial}{\partial y}\left\{\frac{\partial}{\partial x}(F_\xi(x)\cdot F_\eta(y))\right\}=\frac{\partial}{\partial y}\left\{F_\eta(y)\cdot\frac{\partial}{\partial x}F_\xi(x)\right\}=\frac{\partial F_\eta(y)}{\partial y}\cdot\frac{\partial F_\xi(x)}{\partial x}=p_\xi(x)\cdot p_\eta(y).$$

Таким чином,

$$p_{\xi\eta}(x,y)=p_\xi(x)\cdot p_\eta(y).$$

2.8 Числові характеристики випадкового вектора

Нехай (ξ,η) - випадковий вектор.

Теорема. Нехай випадкова величина $z=f(\xi,\eta)$. Тоді математичне сподівання функції двох змінних визначається за правилом

$$M(z)=\sum_{i,j} f(x_i,y_j)p_{ij}$$

для дискретної випадкової величини (ДВВ) і

$$M(z)=\iint_{R^2} f(x,y)p_{\xi\eta}(x,y)dxdy$$

для неперервної випадкової величини (НВВ).

Приклад. Нехай $z=\xi$. Обчислимо математичне сподівання

$$M(z)=M(\xi)=\iint_{R^2} xp_{\xi\eta}(x,y)dxdy=\int_{-\infty}^{+\infty}\left\{\int_{-\infty}^{+\infty} xp_{\xi\eta}(x,y)dy\right\}dx=\int_{-\infty}^{+\infty} xp_{\xi}(x)dx.$$

Означення. Математичним сподіванням випадкового вектора (ξ,η) (або центром його розсіювання), називається невипадковий вектор $(M(\xi),M(\eta))$, де

$$M(\xi)=\int_{-\infty}^{+\infty} x\left\{\int_{-\infty}^{+\infty} p_{\xi\eta}(x,y)dy\right\}dx,\quad M(\eta)=\int_{-\infty}^{+\infty} y\left\{\int_{-\infty}^{+\infty} p_{\xi\eta}(x,y)dx\right\}dy.$$

2.9 Моменти випадкового вектора

Означення 1. Початковим моментом порядку „j+k" двовимірного вектора (ξ,η) називається математичне сподівання добутку j-го степеня координати ξ на k–ту степінь координати η, тобто

$$\nu_{jk}=M(\xi^j\cdot\eta^k).$$

(іншими словами, якщо $z=\xi^j\cdot\eta^k$, тоді початковий момент дорівнює $\nu_{jk}=M(z)$).

Означення 2. Центральним моментом порядку „j+k" випадкового вектора (ξ,η) називається математичне сподівання добутку j- го степеня відхилення координати ξ від її математичного сподівання на к –ту степінь відхилення координати η від її математичного сподівання

$$\mu_{jk}=M\left[(\xi-M(\xi))^j\cdot(\eta-M(\eta))^k\right].$$

Початкових моментів першого порядку всього два:

$$\nu_{01}=M(\xi^0\cdot\eta^1)=M(\eta),\quad \nu_{10}=M(\xi^1\cdot\eta^0)=M(\eta),$$

вони збігаються з математичними сподіваннями координат випадкового вектора.

Центральні моменти першого порядку дорівнюють нулю. Дійсно,

$$\mu_{01}=M\left[(\xi-M(\xi))^0\cdot(\eta-M(\eta))^1\right]=M(\eta-M(\eta))=M(\eta)-M(\eta)=0,$$

$$\mu_{10}=M\left[(\xi-M(\xi))^1\cdot(\eta-M(\eta))^0\right]=M(\xi-M(\xi))=M(\xi)-M(\xi)=0.$$

2.10 Центральні моменти другого порядку

$$\mu_{02}=M\left[(\xi-M(\xi))^0\cdot(\eta-M(\eta))^2\right]=M\left[(\eta-M(\eta))^2\right]=D(\eta),$$

$$\mu_{20}=M\left[(\xi-M(\xi))^2\cdot(\eta-M(\eta))^0\right]=M\left[(\xi-M(\xi))^2\right]=D(\xi),$$

$$\mu_{11}=M\left[(\xi-M(\xi))^1\cdot(\eta-M(\eta))^1\right].$$

З цих формул випливає, що μ_{02}, μ_{20} визначають характеристики окремих координат випадкового вектора.

Але числові характеристики окремих координат вектора ще не визначають поведінки випадкового вектора в цілому.

Для характеристики випадкового вектора як єдиної системи необхідно враховувати і зв'язки, які існують між його координатами.

Цей зв'язок визначається третім центральним моментом другого порядку, а саме:

$$\mu_{11}=M[(\xi-M(\xi))\cdot(\eta-M(\eta))].$$

Означення 3. Другий центральний момент другого порядку μ_{11} випадкового вектора (ξ,η) називається ***коваріацією*** випадкового вектора, або ***кореляційним моментом*** випадкового вектора

$$\mu_{11} = M[(\xi - M(\xi))\cdot(\eta - M(\eta))] = Cov(\xi,\eta) = K_{\xi\eta}.$$

Зауваження. Другі центральні моменти випадкового вектора утворюють матрицю другого порядку

$$K(\xi,\eta) = \begin{pmatrix} \mu_{20} & \mu_{11} \\ \mu_{11} & \mu_{02} \end{pmatrix} = \begin{pmatrix} D(\xi) & Cov(\xi,\eta) \\ Cov(\xi,\eta) & D(\eta) \end{pmatrix} = \begin{pmatrix} K_{\xi\xi} & K_{\xi\eta} \\ K_{\eta\xi} & K_{\eta\eta} \end{pmatrix}.$$

Означення 4. Правило обчислення коваріації випливає із її означення

$$K_{\xi\eta} = Cov(\xi,\eta) = M[(\xi - M(\xi))\cdot(\eta - M(\eta))] = M[\xi\cdot\eta - \xi\cdot M(\eta) - M(\xi)\cdot\eta + M(\xi)\cdot M(\eta)] =$$

$$= M(\xi\cdot\eta) - M(\xi)M(\eta)$$

Таким чином,

$$\underline{K_{\xi\eta} = Cov(\xi,\eta) = M(\xi\cdot\eta) - M(\xi)\cdot M(\eta).}$$

Зауваження. У випадку багатовимірного вектора $\xi(\xi_1,\xi_2,\cdots,\xi_n)$ ***кореляційною*** (коваріаційною***) матрицею*** називається матриця, елементи якої виявляються кореляційними моментами кожної пари координат

$$K = \begin{pmatrix} K_{11} & K_{12} & \cdots & K_{1n} \\ K_{21} & K_{22} & \cdots & K_{2n} \\ \approx & \approx & \approx & \approx \\ K_{n1} & K_{n2} & \cdots & K_{nn} \end{pmatrix}.$$

ТЕМА 3. ЕЛЕМЕНТИ МАТЕМАТИЧНОЇ СТАТИСТИКИ

3.1 Про задачі математичної статистики

В теорії ймовірностей вивчаються різні поняття, пов'язані з випадковими подіями і випадковими величинами; найважливіші з них – це поняття ймовірності, функції розподілу, математичного сподівання тощо.

Але у більшості випадків, що зустрічаються на практиці, точне значення ймовірності, або точний вираз функції розподілу, нам невідомі. Тому постає питання про їх експериментальне визначення.

Математична статистика вивчає методи, які дають змогу за результатами випробувань робити певні ймовірнісні висновки .

Подамо деякі найпростіші задачі, що вивчаються в математичній статистиці.

1 Нехай подія А має ймовірність, але її значення p=P(A) нам невідоме; необхідно оцінити „p" за допомогою випробувань.

2 Є цілий ряд важливих задач, пов'язаних з невідомими функціями розподілу.

Так, може статися, що нам взагалі нічого не відомо про функцію розподілу F(x) випадкової величини X; тоді може бути поставлене питання про наближене подання функції F(x) за допомогою даних, отриманих в результаті випробувань.

Проте, може бути і так, що нам відомо аналітичний вигляд функції розподілу F(x, a_1, a_2, ... ,a_k), але невідомі значення параметрів a_1, a_2, ... ,a_k, від яких залежить ця функція; треба знайти оцінки цих параметрів.

Наприклад, якщо відомо, що випадкова величина розподілена за нормальним законом, тоді її диференціальна функція розподілу (щільність ймовірності) залежить лише від двох параметрів a,σ і саме їх потрібно оцінити за результатами випробувань.

3.2 Генеральна та вибіркова сукупності

Нехай потрібно вивчити сукупність однотипних об'єктів відносно деякої ***якісної***, або ***кількісної ознаки*** , які характеризують ці об'єкти (наприклад, довжина).

На практиці, суцільне обстеження кожного з об'єктів сукупності не провадиться, або тому, що сукупність об'єктів дуже велика, або тому, що у деяких випадках обстеження пов'язане з руйнуванням об'єктів.

На практиці випадково відбирають із цієї сукупності ***обмежене*** число об'єктів і обстежують їх.

Означення 1. ***Вибірковою сукупністю*** (вибіркою) називають сукупність випадково відібраних об'єктів.

Означення 2. ***Генеральною сукупністю*** називають сукупність об'єктів, із якої добувається вибірка.

Оскільки умови експерименту вважаємо незмінними, а результати експериментів – незалежними один від іншого, то результати „n" експериментів $x_1, x_2, \ldots, x_n$ є незалежними у сукупності випадковими величинами, що мають ту саму функцію розподілу F(x)/

Зауваження. На практиці вибір з генеральної сукупності здійснюється так, що після кожного експерименту досліджуваний елемент не повертається до генеральної сукупності перед наступним відбором (так, наприклад, здійснюють вибірковий контроль якості продукції). При цьому не можна вважати, що всі експерименти проводяться в однакових умовах і їх результати незалежні один від іншого, оскільки після кожного експерименту склад генеральної сукупності змінюється.

Такий вибір називають ***вибором без повернення*** на відміну від ***простого*** випадкового вибору, або ***вибору з поверненням*** .

Якщо генеральна сукупність достатньо велика, а вибірка по своєму об'єму становить лише незначну її частину, то відміна між цими двома типами вибору незначна, і нею можна нехтувати.

3.3 Статистичний розподіл вибірки

Нехай з генеральної сукупності ми добули вибірку, причому значення „x_1" спостерігалося „n_1" разів, "x_2" – n_2 разів, ... , "x_k" – n_k разів і $\sum_{i=1}^{k} n_i = n$ - об'єм вибірки.

Спостережені значення "x_i" називають ***варіантами,*** а послідовність варіант, записана у зростаючому порядку, - ***варіаційним рядом.***

Числа спостережень „n_i" називають ***частотами***, а їх відношення до об'єму вибірки $\frac{n_i}{n} = w_i$ - ***відносними частотами***.

Означення. ***Статистичним розподілом вибірки*** називають перелік варіант і відповідних частот (або відносних частот).

Приклад.

x_i	2	6	12
n_i	3	10	7

або

x_i	2	6	12
w_i	3/20	10/20	7/20

Тут $n = 20, \quad w_i = \frac{n_i}{n}, \quad \sum w_i = 1.$

3.4 Емпірична функція розподілу

Нехай відомий статистичний розподіл частот кількісної ознаки X. Введемо позначення: n_x – число спостережень, при яких спостерігалося значення ознаки X<x; n – загальне число спостережень (об'єм вибірки).

Відносна частота події X<x дорівнює n_x/n.

Якщо змінюється „x", тоді змінюється і відносна частота, тобто відношення n_x/n є функцією аргументу "x".

Означення. ***Емпіричною*** функцією розподілу (функцією розподілу вибірки) називають функцію $F^*(x)$, яка визначає для кожного значення "x" відносну частоту події X<x

$$\underline{F^*(x) = n_x/n} .$$

На відзнаку від $F^*(x)$, функцію розподілу генеральної сукупності (ГС) F(x) = P(X<x) називають ***теоретичною*** функцією розподілу.

Зауважимо, що при великих об'ємах вибірки (n>>1) $F^*(x)$ наближається до F(x).

Властивості $F^*(x)$:

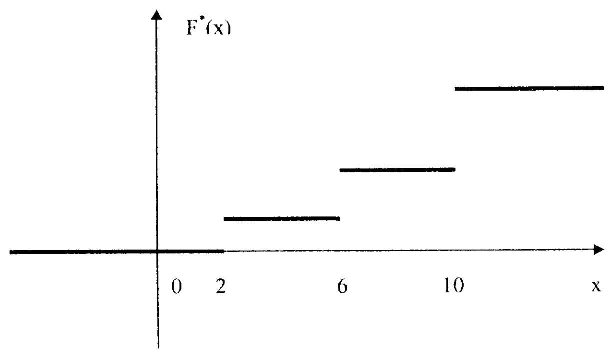
1) значення $F^*(x)$ належать відрізку [0,1] ;
2) $F^*(x)$ – монотонно зростаюча функція;
3) якщо x_1 – найменша варіанта, тоді $F^*(x) = 0$ при $x \le x_1$; якщо $x_к$ – найбільша варіанта, тоді $F^*(x) = 1$ при $x > x_k$.

Таким чином, емпірична функція розподілу може застосовуватися для оцінки теоретичної функції розподілу ГС.

Приклад. Задано вибірку об'єму n =60. Визначити емпіричну функцію розподілу $F^*(x)$ і накреслити її графік:

x_i	**2**	**6**	**10**
n_i	**12**	**18**	**30**

1) $x \le 2$, $F^*(x)=0$; 2) $2<x\le 6$, $F^*(x) = n_1/n = 12/60 = 0,2$;
3) $6<x\le 10$, $F^*(x) = (n_1+n_2)/n = 30/60 = 0,5$;
4) $x>10$, $F^*(x) = (n_1+n_2+n_3)/n = 60/60 = 1$.



3.5 Полігон і гістограма

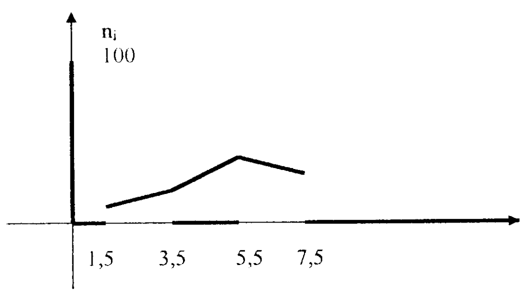
Для наочності будують різні графіки статистичного розподілу, зокрема, ***полігон*** і ***гістограму***.

Означення. ***Полігоном частот*** називають ламану, відрізки якої з'єднують точки з координатами (x_1,n_1), (x_2,n_2), … , (x_k,n_k)

Полігоном відносних частот називають ламану, відрізки якої з'єднують точки з координатами (x_1,w_1), (x_2,w_2), … , (x_k,w_k).

Приклад. Побудуємо полігон частот для вибірки об'єму $n=100$:

x_i	1,5	3,5	5,5	7,5
n_i	10	20	40	30
w_i	0,1	0,2	0,4	0,3



У випадку **неперервної ознаки** доцільно будувати ***гістограму***, для чого інтервал, в якому знаходяться усі спостережені значення ознаки (якщо вигідно, то інтервал можна збільшити), ділять на декілька часткових інтервалів довжиною „h" і обчислюють для кожного часткового інтервалу **суму частот варіант** – n_i , **які потрапили в і-ий інтервал**.

Гістограмою частот називають ступінчасту фігуру, яка будується із прямокутників, основою яких є часткові інтервали довжиною „h", а висоти дорівнюють відношенням n_i/h (щільність частоти).

Площа і-го часткового прямокутника дорівнює $h \cdot n_i / h = n_i$ – сумі частот і-го інтервалу; таким чином, **площа гістограми частот дорівнює** сумі всіх частот, тобто дорівнює **об'єму вибірки.**

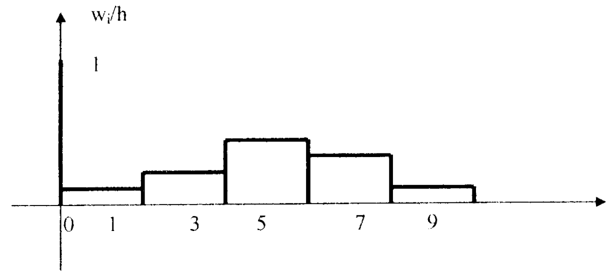
Гістограмою відносних частот називають ступінчасту фігуру, яка будується із прямокутників, основою яких є часткові інтервали довжиною „h", а висоти дорівнюють w_i/h (щільність відносної частоти).

Площа і-го прямокутника дорівнює $h \cdot w_i / h = w_i$ - відносній частоті варіант, які містяться у і-му інтервалі.

Таким чином, площа гістограми відносних частот дорівнює сумі усіх відносних частот, тобто дорівнює одиниці.

Приклад. Побудувати гістограму відносних частот вибірки об'єму $n=100$, припускаючи, що всі варіанти розташовані в інтервалі (0, 10), і $h=2$.

x_i	1	3	5	7	9
n_i	10	15	40	23	12
w_i	0,1	0,15	0,4	0,23	0,12



3.6 Статистичні оцінки параметрів розподілу

3.6.1 *Статистичні оцінки параметрів розподілу*

Нехай потрібно вивчити кількісну ознаку X генеральної сукупності (ГС). Припустимо, що з теоретичних міркувань встановлено який саме розподіл має ця ознака X.

Тоді потрібно оцінити параметри, які визначають цей розподіл.

Наприклад, якщо відомо, що ознака X, яка вивчається, в ГС розподілена за нормальним законом, тоді необхідно оцінити (знайти наближено) математичне сподівання a і середнє квадратичне відхилення σ; якщо ж можна вважати, що ознака X розподілена за законом Пуассона, тоді необхідно оцінити лише один параметр λ .

Як правило, у роспорядженні дослідника є лише елементи вибірки, наприклад, значення кількісної ознаки $x_1, x_2, \dots, x_n$, отримані в результаті спостережень (тут і далі спостереження вважаються незалежними).

Через елементи вибірки і треба виразити параметр, який оцінюється.

Крім того, ми маємо право розглядати будь-яке число вибірок об'єму „n" і тоді природно $x_1, x_2, \dots, x_n$ розглядати як незалежні випадкові величини $X_1, X_2, \dots, X_n$.

Більш того, можна сказати, що „ знайти статистичну оцінку невідомого параметра теоретичного розподілу" – це означає „ знайти функцію від випадкових величин, які спостерігаються, яка і визначає наближене значення параметра, який оцінюється":

$$\Theta^* = f(X_1, X_2, \cdots, X_n).$$

3.6.2 *Незсунені, консистентні оцінки*

Нехай Θ - деякий параметр теоретичного розподілу (наприклад, математичне сподівання, дисперсія, медіана тощо).

Оцінкою цього параметра називають таку функцію $\Theta^* = f(X_1, X_2, \dots, X_n)$ від значень, що спостерігалися, яка у певному розумінні мало відрізняється від Θ.

У математичній статистиці розглядають оцінки, які задовольняють певним умовам.

Означення 1. Оцінка $\Theta^* = f(X_1, X_2, \dots, X_n)$ параметра Θ називається ***консистентною*** оцінкою цього параметра, якщо при $n \to +\infty$ $\Theta^* = f(X_1, X_2, \dots, X_n) \to \Theta$ за ймовірністю

(тобто $\lim\limits_{n \to +\infty} P(|\Theta^* - \Theta| > \varepsilon) = 0$).

Означення 2. Оцінку $\Theta^* = f(X_1, X_2, \dots, X_n)$ називають ***незсуненою*** оцінкою параметра Θ, якщо при будь-якому натуральному „n" $M(\Theta^*) = \Theta$.

Умова незсуненості означає, що похибка від заміни Θ на Θ^* не має систематичного характеру.

3.6.3 *Генеральна середня*

Нехай вивчається дискретна генеральна сукупність відносно кількісної ознаки X.

Генеральною середньою $\overline{X_Г}$ називають середнє арифметичне значень ознаки X генеральної сукупності

$$\overline{X_Г} = \frac{1}{N}(x_1N_1 + x_2N_2 + \dots + x_kN_k), \quad \sum_{i=1}^{k} N_i = N.$$

Очевидно, що

$$P(X = x_i) = \frac{N_i}{N},$$

тому

$$\overline{X_Г} = M(X).$$

3.6.3 *Вибіркова середня*

Нехай для вивчення ГС відносно кількісної ознаки X відібрана вибірка об'єму „n".

Вибірковою середньою $\overline{x_B}$ називають середнє арифметичне значень ознаки X вибіркової сукупності

$$\overline{x_B} \frac{1}{n}(x_1n_1 + x_2n_2 + \dots + x_kn_k), \quad \sum_{i=1}^{k} n_i = n,$$

або

$$\overline{x_B} = \frac{1}{n}\sum_{i=1}^{k} x_i n_i .$$

Зауваження. У теоретичних міркуваннях вибіркові значення $x_1, x_2, \ldots, x_n$ ознаки X, отримані в результаті незалежних спостережень, розглядають як випадкові величини $X_1, X_2, \ldots, X_n$, які мають однаковий з ознакою X розподіл і такі самі числові характеристики.

3.6.5 Оцінка генеральної середньої через вибіркову середню. Стійкість вибіркових середніх

Нехай з ГС добута вибірка об'єму „n"

x_i	x_1	x_2	x_3	…	x_k
n_i	n_1	n_2	n_3	…	n_k

$$\sum_{i=1}^{k} n_i = n.$$

Нехай генеральна середня $\overline{X_Г}$ невідома і потрібно оцінити її за даними вибірки.

В якості оцінки (наближеного значення) генеральної середньої приймають вибіркову середню

$$\overline{x_B} = \frac{1}{n}\sum_{i=1}^{k} x_i n_i .$$

Доведемо, що $\overline{x_B}$ - незсунена оцінка $\overline{X_Г}$.

Маючи на увазі попереднє зауваження, розглянемо середню вибіркову як випадкову величину, $x_1, x_2, \ldots, x_к$ як незалежні випадкові величини $X_1, X_2, \ldots, X_к$ з однаковими розподілами, а тоді і з однаковими математичними сподіваннями

$$M(X)=a; \quad M(X_1)=M(X_2)=\ldots=M(X_к)=a.$$

Але тоді

$$M\left(\overline{X_B}\right) = M\left(\frac{1}{n}\sum_{i=1}^{k} X_i n_i\right) = \frac{1}{n}\sum_{i=1}^{k} M(X_i) n_i = \frac{1}{n}\sum_{i=1}^{k} a n_i = a = M(X),$$

тобто

$$M\left(\overline{X_B}\right) = \overline{X_Г}.$$

Крім того, можна довести, що вибіркова середня буде консистентною оцінкою генеральної середньої

$\overline{X_B} \to \overline{X_Г}$ за ймовірністю при умові, що $n \to +\infty$.

З цього випливає, якщо з деякої генеральної сукупності добуті декілька вибірок достатньо великого об'єму, то вибіркові середні будуть наближено рівні між собою.

В цьому полягає **властивість *стійкості* вибіркових середніх.**

3.6.6 Генеральна дисперсія

Для характеристики величини розсіювання значень кількісної ознаки X генеральної сукупності навколо свого середнього значення розглядаємо генеральну дисперсію.

Якщо значення ознаки X: $x_1, x_2, \ldots, x_к$ мають відповідно частоти $N_1, N_2, \ldots, N_k$, причому $\sum_{i=1}^{k} N_i = N$, то генеральна дисперсія дорівнює

$$D_Г = \frac{1}{N}\sum_{i=1}^{k}\left(x_i - \overline{X_Г}\right)^2 N_i .$$

Генеральним середнім квадратичним відхиленням (стандартом) називають квадратний корінь із дисперсії

$$\sigma_Г = \sqrt{D_Г}.$$

3.6.7 Вибіркова дисперсія

Для характеристики величини розсіювання значень кількісної ознаки X вибіркової сукупності навколо свого середнього значення $\overline{x_B}$, розглядаємо – вибіркову дисперсію.

Вибірковою дисперсією D_B називають ***середнє арифметичне квадратів відхилень спостережених значень ознаки X від їх середнього значення***

$$D_B = \frac{1}{n}\sum_{i=1}^{k}\left(x_i - \overline{x_B}\right)^2 n_i .$$

3.6.8 Формула для обчислення дисперсії

Теорема. Дисперсія дорівнює середньому квадратів значень ознаки X мінус квадрат середньої.

Доведення:

$$D=\frac{1}{n}\sum_{i=1}^{k}n_i(x_i-\bar{x})^2=\frac{1}{n}\sum_{i=1}^{k}n_i\left(x_i^2-2x_i\bar{x}+\bar{x}^2\right)=\frac{1}{n}\sum_{i=1}^{k}n_ix_i^2-2\frac{\bar{x}}{n}\sum_{i=1}^{k}x_in_i+\frac{[\bar{x}]^2}{n}\sum_{i=1}^{k}n_i=$$

$$=\frac{1}{n}\sum_{i=1}^{k}x_in_s-2[\bar{x}]^2+[\bar{x}]^2$$

тобто

$$\underline{D=\overline{x^2}-[\bar{x}]^2}\,. \qquad (*)$$

3.6.9 Оцінка генеральної дисперсії

Нехай із генеральної сукупності витягнута вибірка з поверненням об'єму $n=\sum_{i=1}^{k}n_i$:

$$x_1,\ x_2,\ \ldots,\ x_k$$
$$n_1,\ n_2,\ \ldots,\ n_k.$$

Потрібно за даними вибірки оцінити (наближено знайти) невідому генеральну дисперсію.

Припустимо спочатку, що в якості такої оцінки ми візьмемо вибіркову дисперсію $D_В$.

Але виявляється, що вибіркова дисперсія $D_В$ не є незсуненою оцінкою для теоретичної, тобто генеральної, дисперсії $D_Г$.

Перевіримо, що це саме так.

Введемо позначення: $y_i=x_i-a$ $\quad i=1,2,\ldots,N$ (де через "a" позначено генеральну середню $\overline{X_Г}=a$).

Тоді

$$M(y_i)=M(x_i-a)=M(x_i)-a=a-a=0.$$

(Тут ми врахували зауваження, а саме $M(X_1)=M(X_2)=\ldots=M(X_к)=a$).

Далі обчислюємо

$$M(y_i^2)=M[(x_i-a)^2]=D(x_i)=D_Г=\sigma^2,$$

$$D(y_i)=D(x_i-a)=D(x_i)-D(a)=D(x_i)=D_Г=\sigma^2.$$

За формулою (*)

$$M(D_B)=\frac{1}{n}M\left[\sum_{i=1}^{n}(x_i-\overline{x_B})^2\right]=\frac{1}{n}M\left[\sum_{i=1}^{n}(y_i-\bar{y})^2\right]=\frac{1}{n}M\left[\sum_{i=1}^{n}y_i^2\right]-\frac{2}{n}M\left[\sum_{i=1}^{n}y_i\cdot\bar{y}\right]+\frac{1}{n}M\left[n(\bar{y})^2\right]=$$

$$=\frac{1}{n}\sum_{i=1}^{k}M(y_i^2)-2M(\bar{y}^2)+M(\bar{y}^2)=\frac{1}{n}\cdot n\cdot\sigma^2-M(\bar{y}^2) \qquad (3.1)$$

Оскільки y_i і y_j незалежні при $i\neq j$, то

$$M(y_i\cdot y_j)=M(y_i)\cdot M(y_j)=0,$$

тому

$$M\left[(\bar{y})^2\right]=M\left(\frac{y_1+y_2+\ldots+y_n}{n}\right)^2=\frac{1}{n^2}M\left(\sum_{i=1}^{n}y_i^2+2\sum_{i\neq j}y_i\cdot y_j\right)=\frac{1}{n^2}\sum_{i=1}^{n}M(y_i^2)=\frac{1}{n^2}\cdot n\cdot\sigma^2=\frac{\sigma^2}{n}.$$

Повертаючись до попереднього виразу (3.1), дістаємо

$$M(D_B)=\sigma^2-\frac{\sigma^2}{n}=\frac{n-1}{n}\sigma^2\,. \qquad (3.2)$$

Отже, $M(D_B)\neq\sigma^2$, тобто $D_В$ не є незсуненою оцінкою для генеральної дисперсії.

За формулою (3.2) можна дістати незсунену оцінку для генеральної дисперсії. Справді, розглянемо „виправлену" дисперсію

$$S^2=\frac{n}{n-1}D_B=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{x_B})^2\,. \qquad (3.3)$$

Очевидно,

$$M(S^2)=M\left(\frac{n}{n-1}D_B\right)=\frac{n}{n-1}M(D_B)=\frac{n}{n-1}\cdot\frac{n-1}{n}\sigma^2=\sigma^2$$

і тому S^2 є незсуненою оцінкою для генеральної (теоретичної) дисперсії σ^2.

При $n\to+\infty$ обидві величини $D_В$ і S^2 прямують до одної границі і тому ця оцінка є також консистентною.

Для оцінки середнього квадратичного відхилення використовують „виправлене" середнє квадратичне відхилення

$$S=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{x_B})^2}\,.$$

Підкреслимо, що S не є незсуненою оцінкою і тому ми будемо говорити так: **„виправлене" середнє квадратичне відхилення.**

Зауваження. На практиці виправленою дисперсією користуються лише, якщо n<30.

3.7 Точкові оцінки, надійна ймовірність (надійність), надійний інтервал

Означення 1. ***Точковою*** називають оцінку, яка визначається ***одним числом***.

Для вибірки малого об'єму точкова оцінка може значно відрізнятися від параметра, який оцінюється, тобто приводити до грубих помилок. Тому, при невеликому об'ємі вибірки треба користуватися інтервальними оцінками.

Означення 2. ***Інтервальною*** називають оцінку, яка визначається ***двома числами*** – кінцями інтервалу.

Якщо Θ^*, яка знайдена за елементами вибірки, є оцінкою (наближеним значенням) для невідомого параметра Θ (який вважається величиною сталою), тоді Θ^* тим краще визначає Θ, чим меншим буде $|\Theta^* - \Theta|$, тобто, якщо $\delta > 0$ і $|\Theta - \Theta^*| < \delta$, то чим менше δ, тим точніше буде оцінка Θ^*.

Таким чином, $\delta > 0$ характеризує **точність оцінки.**

Але статистичні методи не дозволяють категорично стверджувати, що оцінка Θ^* задовольняє нерівності $|\Theta - \Theta^*| < \delta$; можна лише говорити про ймовірність γ, з якою виконується ця нерівність.

Означення 3. ***Надійністю (надійною ймовірністю)*** оцінки Θ^* для Θ називають ймовірність γ, з якою виконується нерівність $|\Theta - \Theta^*| < \delta$:

$$P(|\Theta - \Theta^*| < \delta) = \gamma \ . \qquad (3.4)$$

Як правило, надійність задається наперед, причому в якості γ визначають число близьке до 1, наприклад, 0,95; 0,99; 0,999.

Рівність (3.4) можна переписати так:

$$P(\Theta^* - \delta < \Theta < \Theta^* + \delta) = \gamma \ . \qquad (3.5)$$

Останнє співвідношення треба розуміти так:

Ймовірність того, що випадковий інтервал $(\Theta^* - \delta, \Theta^* + \delta)$ **містить у собі (накриває) невідомий параметр** Θ, **дорівнює** γ .

Надійним називають інтервал $(\Theta^* - \delta, \Theta^* + \delta)$, який накриває невідомий параметр Θ із заданою надійністю (Нейман-Фішер).

Зауваження. Θ^* - випадкова величина і тому інтервал $(\Theta^* - \delta, \Theta^* + \delta)$ має випадкові межі.

Тому, що випадковою величиною є не параметр (стала), який оцінюється, а надійний інтервал, то правильніше говорити не про ймовірність попадання Θ в надійний інтервал, **а про ймовірність того, що надійний інтервал накриває** Θ.

Як приклад, розглянемо надійний інтервал для оцінки математичного сподівання нормального розподілу, якщо середнє квадратичне відхилення $\sigma = \sqrt{D_\Gamma}$ є відомою величиною.

Нехай кількісна ознака X генеральної сукупності розподілена за нормальним законом, причому середнє квадратичне відхилення σ цього розподілу відоме.

При цій умові треба оцінити невідоме математичне сподівання „a" за допомогою вибіркової середньої $\overline{x_B}$.

Сформулюємо задачу так: **знайти надійний інтервал, який накриває „a" з надійністю** γ .

Будемо розглядати вибіркову середню $\overline{x_B}$ як випадкову величину $\overline{X_B}$, а вибіркові значення $x_1, x_2, \cdots, x_n$ - як однаково розподілені випадкові величини X_1, X_2, ... , X_n з однаковими математичними сподіваннями і дисперсіями

$$M(X_i) = a, \quad \sqrt{D(X_i)} = \sigma .$$

Припустимо, без доведення, що, якщо X має нормальний розподіл, тоді і $\overline{X_B}$, знайдена за допомогою незалежних випробувань, теж розподілена нормально, причому має такі параметри:

$$M(\overline{X_B}) = M\left(\frac{X_1 + X_2 + \ldots + X_n}{n}\right) = \frac{1}{n}\sum_{i=1}^{n} M(X_i) = \frac{1}{n}(na) = a ,$$

$$D(\overline{X_B}) = D\left(\frac{X_1 + X_2 + \ldots + X_n}{n}\right) = \frac{1}{n^2}\sum_{i=1}^{n} D(X_i) = \frac{1}{n^2} \cdot n\sigma^2 = \frac{\sigma^2}{n} ,$$

$$\sigma(\overline{X_B}) = \sqrt{D(\overline{X_B})} = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}} .$$

Нехай тепер виконується співвідношення

$$P\left(\left|\overline{X_B}-a\right|<\delta\right)=\gamma, \qquad (3.6)$$

де γ - задана надійність.

Використаємо формулу

$$P\left(|X-a|<\delta\right)=2\Phi\left(\frac{\delta}{\sigma}\right).$$

(Дійсно,

$$P\left(|X-a|<\delta\right)=P(a-\delta<X<a+\delta)=\Phi\left(\frac{a+\delta-a}{\sigma}\right)-\Phi\left(\frac{a-\delta-a}{\sigma}\right)=\Phi\left(\frac{\delta}{\sigma}\right)-\Phi\left(\frac{-\delta}{\sigma}\right),$$

але функція Лапласа $\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}dt$ - непарна $\Phi(-x)=-\Phi(x)$, і тому

$$P\left(|X-a|<\delta\right)=2\Phi\left(\frac{\delta}{\sigma}\right)), \qquad (3.7)$$

в якій замінемо X на $\overline{X_B}$, $\sigma - \frac{\sigma}{\sqrt{n}}$, тобто

$$P\left(\left|\overline{X_B}-a\right|<\delta\right)=2\Phi\left(\frac{\delta\sqrt{n}}{\sigma}\right),$$

або

$$P\left(\left|\overline{X_B}-a\right|<\delta\right)=2\Phi(t), \quad t=\frac{\delta\sqrt{n}}{\sigma}. \qquad (3.8)$$

Із рівностей (3.6) і (3.8) випливає, що

$$2\Phi(t)=\frac{\gamma}{2},$$

звідки, при відомій надійності γ можна знайти "t", якщо використати таблиці значень інтеграла Лапласа $\Phi(x)$:

$$\Phi(t)=\frac{t}{2} \Rightarrow t \Rightarrow \delta=\frac{t\sigma}{\sqrt{n}}.$$

Таким чином, можна записати таку робочу формулу

$$P\left(\overline{X_B}-t\frac{\sigma}{\sqrt{n}}<a<\overline{X_B}+t\frac{\sigma}{\sqrt{n}}\right)=\gamma, \quad \gamma=2\Phi(t).$$

Зміст отриманого співвідношення такий:

з надійністю γ можна стверджувати, що надійний інтервал

$$\left(\overline{X_B}-t\frac{\sigma}{\sqrt{n}}, \overline{X_B}+t\frac{\sigma}{\sqrt{n}}\right) \qquad (3.9)$$

накриває невідоме математичне сподівання „a"; точність оцінки дорівнює

$$\delta=\frac{t\sigma}{\sqrt{n}}.$$

Зауваження 1. Із формули $\delta=\frac{t\sigma}{\sqrt{n}}$ можна зробити такі висновки:

1) при зростанні об'єму вибірки „n" число δ спадає, отже точність оцінки збільшується;

2) зростання надійності $\gamma=2\Phi(t)$ призводить до зростання „t", тому що $\Phi(t)$ зростаюча функція, отже і до зростання δ ;

Приклад. Випадкова величина **X** має нормальний розподіл з відомим $\sigma=3$.

Знайти надійний інтервал для оцінки невідомого математичного сподівання „a" за допомогою вибіркового середнього $\overline{X_B}$, якщо об'єм вибірки n=36 і задана надійність оцінки $\gamma=0{,}95$.

Розв'язання: знайдемо „t":

$$\gamma=2\Phi(t) \Rightarrow \Phi(t)=\frac{\gamma}{2}=0{,}475.$$

За допомогою таблиці знаходимо

$$t = 1{,}96.$$

Далі відшукуємо точність оцінки

$$\delta = \frac{t\sigma}{\sqrt{n}} = \frac{1{,}96\cdot 3}{\sqrt{36}} = 0{,}98.$$

Надійний інтервал такий

$$\left(\overline{X_B} - 0{,}98\,,\ \overline{X_B} + 0{,}98\right).$$

Наприклад, якщо $\overline{X_B} = 4{,}1$, тоді надійний інтервал має надійні межі (4,1-0,98) і (4,1+0,98), тобто

$$(3{,}18 < a < 5{,}08).$$

Зауваження 2. Задача побудови надійного інтервалу розв'язана для випадку, коли генеральна дисперсія $D_Г = \sigma^2$ відома.

Але, як правило, величина σ^2 нам невідома. Тому на практиці її замінюють відповідною оцінкою, тобто величиною S^2. При цьому для надійності $\gamma = 2\Phi(t)$ для математичного сподівання розглядають надійний інтервал

$$\left(\overline{X_B} - \frac{S}{\sqrt{n}}t\,;\ \overline{X_B} + \frac{S}{\sqrt{n}}t\right).$$

Якщо об'єм вибірки недостатньо великий (n<30), то знайдений надійний інтервал не заслуговує довіри тому, що ми замінили невідому дисперсію її оцінкою S^2.

Зауваження 3. Співвідношенням (3.9) для надійного інтервалу для математичного сподівання можна користуватися при будь-якому розподілі випадкової величини. тому що рівність (3.7) і в цьому випадку має зміст, але в ньому, на основі центральної граничної теореми, ліва частина лише наближено дорівнює правій при великих n>>1.

3.8 Кореляційний зв'язок між випадковими величинами. Регресія

3.8.1 Вивчення зв'язку за допомогою прямих регресії

В частині першій конспекту лекцій було введено поняття коефіцієнта кореляції $\left(\rho_{XY} = \dfrac{K_{XY}}{\sigma_X \cdot \sigma_Y}\right)$ між двома випадковими величинами X і Y.

При цьому було показано, що коли X і Y незалежні, то $\rho_{XY} = 0$, а коли X і Y зв'язані лінійною залежністю (і тільки в цьому випадку) $|\rho_{XY}| = 1$.

Якщо $0 < \rho_{XY} < 1$, то X і Y залежні випадкові величини, і, природно вважати, що, чим ближче $|\rho_{XY}|$ до 1, тим „ближче" залежність між X і Y до функціональної залежності виду Y=aX+b.

Тому цілком природним є питання про наближене зображення зв'язку між X і Y у вигляді лінійної функції (для цієї мети використовують також поліноми вищих степенів).

Означення. Пряма $y = \alpha x + \beta$ називається ***прямою регресії*** (або прямою середньої квадратичної регресії) Y на X, якщо коефіцієнти α, β обрано так, щоб середнє квадратичне відхилення $\alpha X + \beta$ від Y було мінімальним

$$M[Y - (\alpha X + \beta)]^2 = \min. \qquad (3.10)$$

Аналогічно визначається пряма регресії X на Y:

$$M[X - (\alpha_1 Y + \beta_1)]^2 = \min.$$

З умови (3.10) легко визначити α і β. Виконавши елементарні тотожні перетворення, дістанемо:

$$M[Y - (\alpha X + \beta)]^2 = M[\{(Y - M(Y)) - \alpha(X - M(X))\} + (M(Y) - \alpha M(X) - \beta)]^2 = M[Y - M(Y)]^2 -$$

$$-2\alpha M[(X - M(X))\cdot(Y - M(Y))] + \alpha^2 M[X - M(X)]^2 +$$

$$2M\{[(Y - M(Y)) - \alpha(X - M(X))][M(Y) - \alpha M(X) - \beta]\} + M[M(Y) - \alpha M(X) - \beta]^2 =$$

$$= M[Y - M(Y)]^2 - 2\alpha M[(X - M(X))(Y - M(Y))] + \alpha^2 M[X - M(X)]^2 + M[(M(Y) - \alpha M(X) - \beta)^2] =$$

$$= \sigma_Y^2 - 2\alpha K_{XY} + \alpha^2 \sigma_X^2 + (M(Y) - \alpha M(X) - \beta)^2 =$$

$$= \{\sigma_Y^2 - \rho_{XY}^2 \sigma_Y^2\} + \{\rho_{XY}^2 \sigma_Y^2 - 2\alpha \rho_{XY} \sigma_X \sigma_Y + \alpha^2 \sigma_X^2\} + (M(Y) - \alpha M(X) - \beta)^2 =$$

$$= \sigma_Y^2 (1 - \rho_{XY}^2) + (\rho_{XY} \sigma_Y - \alpha \sigma_X)^2 + (M(Y) - \alpha M(X) - \beta)^2,$$

тобто маємо рівність

$$M[Y - (\alpha X + \beta)]^2 = \sigma_Y^2 (1 - \rho_{XY}^2) + (\rho_{XY} \sigma_Y - \alpha \sigma_X)^2 + (M(Y) - \alpha M(X) - \beta)^2, \quad (3.11)$$

із якої випливає, що умову (3.10) буде виконано, якщо другий і третій доданки правої частини дорівнюють нулю

$$(\rho_{XY} \sigma_Y - \alpha \sigma_X) = 0, \quad (M(Y) - \alpha M(X) - \beta) = 0,$$

звідки

$$\alpha = \rho_{XY} \cdot \frac{\sigma_Y}{\sigma_X}, \quad \beta = M(Y) - \alpha M(X) \quad . \quad (3.12)$$

Коефіцієнт α називається ***коефіцієнтом регресії Y на X***.
Рівняння прямої регресії Y на X має вигляд

$$y = \rho_{XY} \frac{\sigma_Y}{\sigma_X} x + (M(Y) - \alpha M(X)),$$

або

$$y - M(Y) = \rho_{XY} \frac{\sigma_Y}{\sigma_X} x + \left(M(Y) - \rho_{XY} \frac{\sigma_Y}{\sigma_X} M(X) \right),$$

або

$$y - M(Y) = \rho_{XY} \frac{\sigma_Y}{\sigma_X} (x - M(X)). \quad (3.13)$$

Аналогічно доводиться, що пряма регресії X на Y має рівняння

$$x - M(X) = \rho_{XY} \frac{\sigma_X}{\sigma_Y} (y - M(Y)). \quad (3.14)$$

Теорема. Пряма регресії Y на X збігається з прямою регресії X на Y тоді і лише тоді, коли $|\rho_{XY}| = 1$.

Доведення. Порівнюючи рівняння (3.13) і (3.14), неважко перевірити, що вони визначають одну пряму тоді і лише тоді, коли $\rho_{XY} = \frac{1}{\rho_{XY}}$, тобто $\rho_{XY}^2 = 1$.

Зауваження 1. З рівності (3.11) випливає, що для прямої регресії Y на X має місце співвідношення

$$M[(Y - (\alpha X + \beta))^2] = (1 - \rho_{XY}^2)\sigma_Y^2 \quad (3.15)$$

(цю величину іноді називають ***залишковою дисперсією***) ; аналогічне співвідношення має місце і для прямої регресії X на Y:

$$M[(X - (\alpha_1 Y + \beta_1))^2] = (1 - \rho_{XY}^2)\sigma_X^2. \quad (3.16)$$

Звідси випливає, що чим менше відрізняється $|\rho_{XY}|$ від 1, тим кращим можна вважати наближення Y лінійним виразом $\alpha X + \beta$.

Зауваження 2. Із рівнянь (3.13), (3.14) випливає, що обидві прямі регресії проходять через точку $(M(X), M(Y))$, яку називають ***центром спільного розподілу випадкових величин X і Y.***

3.8.2 Вибіркові характеристики зв'язку

Нехай X і Y – випадкові величини, зв'язок між якими треба вивчити; в результаті „n" випробувань дістали "n" пар значень цих величин (x_1,y_1), (x_2,y_2), ... , (x_n,y_n).

Вибіркові середні і дисперсії позначимо відповідно $\bar{x}, \bar{y}, \overline{\sigma_X^2}, \overline{\sigma_Y^2}$.

За оцінку коефіцієнта кореляції $\rho_{XY} = \rho$ можна взяти вибірковий коефіцієнт кореляції "r" – коефіцієнт кореляції між розподілами вибірок $x_1, x_2, \dots, x_n$ і $y_1, y_2, \dots, y_n$; при цьому кожній парі (x_i, y_i), як і при побудові одномірної функції розподілу вибірки, ми приписуємо ймовірність 1/n.

Згідно з означенням коефіцієнта кореляції

$$r = \frac{\overline{x \cdot y} - \bar{x} \cdot \bar{y}}{\overline{\sigma_X} \cdot \overline{\sigma_Y}} = \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x} \cdot \bar{y}}{n \overline{\sigma_X} \cdot \overline{\sigma_Y}}. \quad (3.17)$$

Аналогічно визначаються рівняння вибіркових прямих регресії : прямої регресії Y на X

$$y - \overline{y} = r\frac{\overline{\sigma_Y}}{\sigma_X}\left(x - \overline{x}\right) \qquad (3.18)$$

і прямої регресії X на Y

$$x - \overline{x} = r\frac{\overline{\sigma_X}}{\sigma_Y}\left(y - \overline{y}\right). \qquad (3.19)$$

Можна довести, що вибірковий коефіцієнт кореляції „r", а отже і вибіркові коефіцієнти регресії $r\frac{\overline{\sigma_Y}}{\sigma_X}, r\frac{\overline{\sigma_X}}{\sigma_Y}$ є консистентними оцінками відповідних теоретичних параметрів.

Тому, при достатньо великих „n" можна сподіватися, що вибірковий коефіцієнт кореляції і вибіркові прямі регресії будуть мало відрізнятися від теоретичних.

З'ясуємо геометричний зміст вибіркових прямих регресії. Якщо $y = \alpha x + \beta$ - рівняння вибіркової прямої регресії Y на X, то, згідно з означенням прямої регресії за 3.8.1, повинно бути

$$\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \alpha x_i - \beta\right) = \min.$$

Інакше кажучи, пряма регресії – це пряма, яку наближено проведено через точки (x_i, y_i) (i=1,2, ... , n) за методом найменших квадратів.

Отже, рівняння прямих регресії доцільно знаходити лише у тому випадку, коли точки (x_i,y_i) (i=1,2,...,n) групуються поблизу деякої прямої (це можна перевірити, наприклад, побудувавши ці точки в Декартові системі координат). Якщо для точок (x_i, y_i) не спостерігається розміщення поблизу від чогось прямолінійного, то замість прямих регресії можна розглядати лінії вищих порядків. Так, наприклад, можна розглядати параболічну регресію

$$y = \alpha x^2 + \beta x + \gamma ,$$

замінюючи умову (3.10) на умову

$$M\left[\left(Y - \left(\alpha x^2 + \beta x + \lambda\right)\right)^2\right] = \min .$$

3.9 Статистична перевірка статистичних гіпотез

3.9.1 Статистична гіпотеза. Нульова і конкуруюча, проста і складена гіпотези

Означення. ***Статистичною*** називають ***гіпотезу*** про вид невідомого розподілу, або про параметри відомих розподілів.

Наприклад, статистичними є гіпотези:

1) генеральна сукупність розподілена за законом Пуассона;
2) дисперсії двох нормальних сукупностей рівні між собою.

Поряд з висунутою гіпотезою розглядають і гіпотезу, яка їй суперечить. Якщо висунута гіпотеза відкидається, тоді має місце гіпотеза, яка їй суперечить.

Першу гіпотезу називають ***нульовою*** (основною) $\mathbf{H_0}$, другу – ***конкуруючою*** (альтернативною) $\mathbf{H_1}$.

Наприклад: H_0: a=10 ; H_1: a≠ 10 .

Простою називають гіпотезу, яка містить у собі лише ***одне*** припущення, ***складеною – два або декілька***.

3.9.2 Помилки першого та другого роду

Висунута гіпотеза може бути правильною або неправильною (хибною).

Оскільки перевірку здійснюють статистичними методами, гіпотезу називають ***статистичною гіпотезою.***

В результаті статистичної перевірки гіпотези в ***двох*** випадках може бути прийнято невірне рішення, тобто можуть бути допущені ***помилки двох родів***.

Помилка першого роду полягає в тому, що буде ***відкинута правильна гіпотеза.***

Помилка другого роду полягає в тому, що буде ***прийнята неправильна гіпотеза***.

Зауваження 1. Правильне рішення може бути прийнято теж у двох випадках: 1) приймається вірна гіпотеза; 2) відкидається невірна гіпотеза.

Зауваження 2. Ймовірність здійснити помилку першого роду позначають α і називають рівнем ***значущості*** .

Для α часто використовують значення 0,05 або 0,01.

Наприклад, якщо рівень значущості α =0,05, то це означає, що лише у 5 випадках із 100 буде відкинута правильна гіпотеза.

3.9.3 Перевірка гіпотези про нормальний розподіл генеральної сукупності за допомогою ексцесу і асиметрії

Введемо такі означення.

Нехай деяка випадкова величина задана своєю щільністю ймовірності f(x).

1 Центральним моментом к-го порядку називають вираз

$$M_k = \int_{-\infty}^{+\infty} (x - M(X))^k f(x)dx, \quad k = 1, 2, 3, \cdots$$

Очевидно, при к=1

$$M_1 = \int_{-\infty}^{+\infty} (x - M(X)) f(x)dx = \int_{-\infty}^{+\infty} x f(x)dx - M(X)\int_{-\infty}^{+\infty} f(x)dx = M(X) - M(X) = 0$$

тобто центральний момент першого порядку завжди дорівнює нулю.

При к=2

$$M_2 = \int_{-\infty}^{+\infty} (x - M(X))^2 f(x)dx = D(X).$$

2 Асиметрією випадкової величини X називають відношення

$$A = \frac{M_3}{\sigma^3} = \frac{M_3}{(M_2)^{3/2}}.$$

3 Ексцесом випадкової величини X називають різницю

$$E = \frac{M_4}{\sigma^4} - 3$$

або

$$E = \frac{M_4}{(M_2)^2} - 3.$$

Зауваження. Якщо розподіл випадкової величини X симетричний відносно прямої $x=m_X$ ($M(X)=m_X$), тобто f(x) є парною функцією відносно аргументу $x-m_X$, тоді всі непарні центральні моменти дорівнюють нулю (як інтеграли в симетричних границях від непарних функцій), тобто і асиметрія A для таких випадкових величин дорівнює нулю.

Розглянемо тепер випадкову величину X розподілену за нормальним законом

$$f(x) = \frac{1}{\sqrt{2\pi}\cdot\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}, \quad a = m_X$$

і обчислимо для неї ***асиметрію*** і ***ексцес*** :

$$A = \frac{M_3}{\sigma^3} = \frac{1}{\sigma^3}\cdot\frac{1}{\sqrt{2\pi}\cdot\sigma}\int_{-\infty}^{+\infty}(x-m_X)^3 e^{-\frac{1}{2}\left(\frac{x-m_X}{\sigma}\right)^2}dx = 0 ;$$

$$E = \frac{M_4}{\sigma^4} - 3 = \frac{1}{\sqrt{2\pi}\cdot\sigma^5}\int_{-\infty}^{+\infty}(x-m_X)^4 e^{-\frac{1}{2}\left(\frac{x-m_X}{\sigma}\right)^2}dx - 3 = \left| \begin{array}{ll} \frac{x-m_X}{\sigma} = z & \frac{dx}{\sigma} = dz \\ \int_{-\infty}^{+\infty} \approx \int_{-\infty}^{+\infty} & \end{array} \right| =$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} z^4 e^{-\frac{z^2}{2}}dz - 3 = \left| \begin{array}{ll} u = z^3 & du = 3z^2dz \\ dv = ze^{-\frac{z^2}{2}}dz & v = -e^{-\frac{z^2}{2}} \end{array} \right| = \frac{1}{\sqrt{2\pi}}\Big|_{-\infty}^{+\infty}\left(-z^3 e^{-\frac{z^2}{2}}\right) + 3\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} z^2 e^{-\frac{z^2}{2}}dz - 3 =$$

$$= \left| \begin{array}{ll} u = z & du = dz \\ dv = ze^{-\frac{z^2}{2}}dz & v = -e^{-\frac{z^2}{2}} \end{array} \right| = \frac{3}{\sqrt{2\pi}}\Big|_{-\infty}^{+\infty}\left(-ze^{-\frac{z^2}{2}}\right) + \frac{3}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{z^2}{2}}dz - 3 = 3 - 3 = 0.$$

Отже, для теоретичного нормального розподілу випадкової величини X

A=0 і E=0.

Розглянемо тепер деяку генеральну сукупність і будемо вивчати її відносно кількісної ознаки X.

Відомо, що таке дослідження має своєю метою відшукання закону розподілу ознаки X та його параметрів.

За елементами вибірки ми можемо визначити наближені значення його параметрів

$$\overline{x_B}, \quad \sigma_B = \sqrt{D_B}$$

і побудувати гістограму відносних частот. Вид гістограми може нам допомогти висловити припущення (гіпотезу) про закон розподілу X .

Наприклад, ми можемо висунути гіпотезу про нормальний розподіл ознаки X генеральної сукупності.

Найбільш короткий шлях перевірки такої гіпотези полягає в обчисленні асиметрії А і ексцесу Е за допомогою експериментальних даних

$$A_B = \frac{M_3}{(M^2)^{3/2}}, \quad E_B = \frac{M_4}{(M^2)^2} - 3,$$

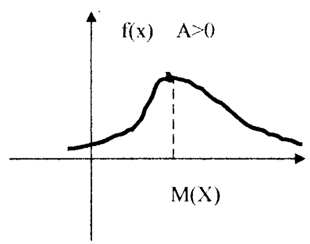
де центральний момент будь-якого к – го порядку (M_k) обчислюється за формулою

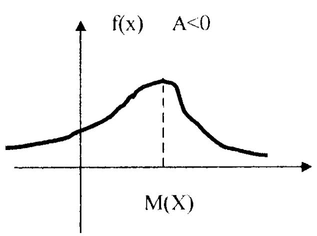
$$M_k = \frac{1}{n}\sum_{i=1}^{m}(x_i - \overline{x_B})^k n_i, \quad k = 2, 3, 4.$$

Якщо виявиться, що A_B і E_B мають невеликі значення, тоді можна вважати, що гіпотеза про нормальний розподіл ознаки X генеральної сукупності узгоджується з результатами експерименту.

Великі значення асиметрії A_B і ексцесу E_B вказують на значне відхилення емпіричного розподілу від нормального.

Зауваження. Для різних знаків асиметрії А мають місце такі види графіків щільності f(x):

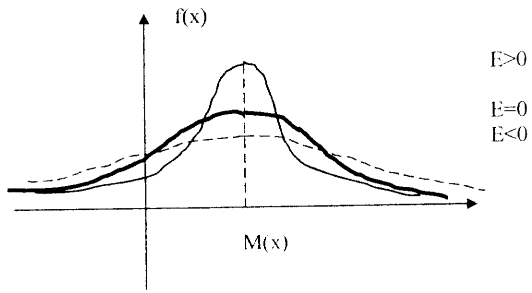




Тобто асиметрія ***додатня***, якщо „довга" частина кривої розподілу розташована з ***правого*** боку від математичного сподівання; асиметрія ***від'ємна***, якщо „довга" частина кривої розподілу розташована з ***лівого*** боку від математичного сподівання.

На практиці знак асиметрії визначають по розташуванню кривої розподілу відносно ***моди*** (точки максимуму f(x)).

Для оцінки „крутості", тобто більшого або меншого підйому кривої розподілу в порівнянні з нормальною кривою, використовують ***ексцес***:



3.9.4 *Статистичні критерії перевірки нульової гіпотези*

Для перевірки нульової гіпотези використовують ***спеціально підібрану випадкову величину***, точний, або наближений, розподіл якої виявляється відомим.

Цю величину називають ***статистичним критерієм.***

Після вибору критерію, множину усіх його можливих значень поділяють на дві множини, які не перетинаються :

Одна з них містить у собі значення критерію, при яких нульова гіпотеза відхиляється – ***критична область*** , друга – містить у собі значення критерію, при яких нульова гіпотеза приймається - ***область прийняття гіпотези.***

Основний принцип перевірки статистичної гіпотези можна сформулювати так: за експериментальними даними обчислюють ***спостережене значення критерію***.

Якщо спостережене значення критерію належить області прийняття гіпотези – гіпотезу приймають;

якщо спостережене значення критерію належить критичній області – гіпотезу відхиляють.

Критерій перевірки гіпотези ***про можливий закон розподілу*** випадкової величини X називають ***критерієм згоди.***

3.9.5 Критерій згоди Пірсона

Нехай за вибіркою об'єму „n" отримано статистичний розподіл

x_i	x_1	x_2	...	x_k
n_i	n_1	n_2	...	n_k

Припустимо, що ознака X генеральної сукупності розподілена нормально і, при цьому припущенні, обчислені теоретичні частоти n_i'.

При рівні значущості α треба перевірити нульову гіпотезу : ***ознака X генеральної сукупності розподілена нормально.***

В якості критерію перевірки нульової гіпотези візьмемо випадкову величину

$$\chi^2_{спост}=\sum_{i=1}^{k}\left(n_i-n_i'\right)^2 \Big/ n_i' \qquad (*)$$

Ця величина випадкова, тому що в різних випробуваннях вона отримує різні значення, які наперед не можна передбачити.

Але ясно, що чим менше відрізняються емпіричні і теоретичні частоти, тим меншою буде величина критерію (*), отже, він в певній степені характеризує близькість емпіричного і теоретичного розподілів.

Доведено, що при $n\to+\infty$ закон розподілу випадкової величини (*), незалежно від того, якому закону розподілу підпорядкована генеральна сукупність, прямує до закону розподілу χ^2 з „к" степенями волі.

Тому випадкова величина (*) позначена через χ^2, а сам критерій називають **критерієм згоди „хі квадрат".**

Число степенів волі знаходять за рівністю

$$k=r-l-1,$$

де r – число груп (часткових інтервалів) вибірки;
l – число параметрів припущеного розподілу, які оцінені за даними вибірки.

Зокрема, якщо припущений розподіл – нормальний, тоді оцінюють два параметри (математичне сподівання „a" і середнє квадратичне відхилення σ), тому l=2 і число степенів волі

$$k=r-2-1=r-3.$$

Якщо, наприклад, припускають, що генеральна сукупність розподілена за законом Пуассона, тоді оцінюють лише один параметр λ, тому

$$k=r-1-1=r-2.$$

Правило. Для того, щоб при даному рівні значущості перевірити нульову гіпотезу H_0 : ***ознака X генеральної сукупності розподілена за нормальним законом.***

Треба спочатку обчислити теоретичні частоти, а потім спостережене значення критерію

$$\chi^2_{спост}=\sum_{i=1}^{k}\left(n_i-n_i'\right)^2 \Big/ n_i'$$

і в таблиці критичних точок розподілу χ^2, за даним рівнем значущості α і числом степенів волі k=r-3, знайти критичну точку $\chi^2_{kp}(\alpha, k)$.

Якщо

$\chi^2_{спост}<\chi^2_{kp}$ - немає підстав відхиляти нульову гіпотезу.

Якщо

$\chi^2_{спост}>\chi^2_{kp}$ - нульову гіпотезу відхиляють.

СПИСОК ЛІТЕРАТУРИ

1 Гмурман В.Е. Теория вероятностей и математическая статистика. - М.: Высшая школа, 1977.

2 Гмурман В.Е. Сборник задач по теории вероятностей и математической статистике. - М.: Высшая школа, 1990.