Probability Theory Notes

November 30, 2021

Contents

Chapter 1

Числові характеристики випадкових величин

1.1 Попередні зауваження

Розглянемо дискретну випадкову величину ξ

$$\xi(\omega) = \sum_{i=1}^{n} x_i \mathbb{1}_{A_i}(\omega)$$

$$\{A_1, \ldots, A_n\} \text{ - повна група подій}$$

Проведено n незалежних випробувань в кожному з яких спостерігається

$$\xi_n(\omega) = \sum_{i=1}^{m} x_i \cdot \mathbb{1}^n_{A_i}(\omega)$$

Розглянемо

$$\hat{\xi} = \frac{\xi_1 + \cdots + \xi_n(\omega)}{n} = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{m} x_j \cdot \mathbb{1}^i_{A_j}(\omega) =$$

$$= \sum_{i=1}^{m} x_j \cdot \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}^i_{A_j}(\omega)$$

$$\frac{1}{n} \sum_{i=1}^{n} \mathbb{1}^i_{A_j}(\omega) \text{ - частота появи } A_j \text{ в } n \text{ випробуваннях } \rightarrow_{n\to\infty} P(A_j).$$

$$\hat{\xi} = \frac{\xi_1 + \cdots + \xi_n}{n} \rightarrow_{n\to\infty} \sum_{j=1}^{m} x_j \cdot P(A_j).$$

Припустимо $\Omega[0,1];\ \mathcal{F} = \mathcal{B}([0,1])$, P міра Лебега, $P((a,b]) = b - a$
для дискретної ймовірності:

$$S_1 = x_1 \cdot P(A_1) = x_1 \cdot |A_1|$$
$$S_2 = x_2 \cdot P(A_2) = x_2 \cdot |A_2|$$
$$S \sim \sum_{j=1}^{m} x_j P(A_j) \text{ - площа}$$

для неперевного випадку:

$$\hat{\xi} \sim \int\limits_{\Omega} \xi(\omega) P(d\omega).$$

1.2 Definition and examples of expected value

Нехай $(\omega, \mathcal{F}, P)$ - ймовірністний простір. ξ - випадкова величина на цьому просторі.

Definition 1. *Математичним сподіванням випадкової величини ξ називається число*

$$M\xi = \int\limits_{\Omega} \xi(\omega) P(d\omega).$$

$$\textit{(expectation)} \quad E\xi = \int\limits_{\Omega} \xi(\omega) P(d\omega).$$

ξ індукує міру P_ξ на $\mathbb{R}$:

$$P_\xi((a,b]) = F_\xi(b) - F_\xi(a).$$

Заміна $\xi(\omega) = x$ приводить до інтеграла Лебега-Стілтьєса:

$$M\xi = \int\limits_{\mathbb{R}} x P_\xi(dx).$$

Звідси маємо інтеграл Стілтьєса:

$$M\xi = \int\limits_{\mathbb{R}} x dF_\xi(x).$$

Для дискретної випадкової величини ξ:

$$E\xi = \sum_{i=1}^{\infty} x_i \cdot P(\xi = x_i) \tag{1.1}$$

Якщо ξ має щільність $f_\xi(x)$:

$$E\xi = \int\limits_{\mathbb{R}} x f_x(x) dx \tag{1.2}$$

Remark. *It is considered that expectation exists if series* (1.1) *or integral* (1.2) *is absolutely convergent.*

Example. *If $A \in \mathcal{F}$ then $\xi(\omega) = \mathbb{1}_A(\omega)$*

$$E\xi = 0 \cdot P(\xi = 0) + 1 \cdot P(\xi = 1) = P(A).$$

Example.

$$P(\xi = i) = \frac{1}{i(i+1)}, i = 1, 2, \ldots.$$

$$\sum_{i=1}^{\infty} i \cdot P(\xi = i) = \sum_{i=1}^{\infty} \frac{1}{i+1} = +\infty \implies E\xi \textit{ does not exist}.$$

Example.

$$\xi \sim U(a,b);\ f_\xi(x) = \frac{1}{b-a}\mathbb{1}(x \in (a,b]).$$

$$E\xi = \frac{1}{b-a}\int\limits_a^b x dx = \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2}.$$

For uniform distribution the expectation is the middle of the segment.

Example.

$$\xi \sim C(0,1);\ f_\xi(x) = \frac{1}{\pi(a+x^2)}.$$

Whereas $\int\limits_{-\infty}^{\infty} \frac{xdx}{\pi(1+x^2)}$ *- divergent then* $E\xi$ *does not exist.*

Let g - Borel function. Then $g(\xi)$ - stochastic variable. For $Mg(\xi)$ have:

$$Eg(\xi) = \int\limits_\Omega g(\xi(\omega))P(d\omega) = \int\limits_\mathbb{R} g(x)P_\xi(dx) = \int\limits_\mathbb{R} g(x)dF_\xi(x).$$

For discrete stochastic variable:

$$Eg(\xi) = \sum_{i=1}^{+\infty} g(x_i) \cdot P(\xi = x_i).$$

For absolutely continuous:

$$Eg(\xi) = \int\limits_\mathbb{R} g(x)f_\xi(x)dx.$$

If $\xi = (\xi_1, \dots, \xi_n)$ with density $f_\xi(x_1, \dots, x_n),\ g : \mathbb{R}^n \to \mathbb{R}$ - Borel function.

$$Eg(\xi_1, \dots, \xi_n) = \int \cdots \int\limits_{\mathbb{R}^n} g(x_1, \dots, x_n) f_\xi(x_1, \dots, x_n) dx_1 \dots dx_n.$$

Theorem 1. *Properties of expectation*

1. $Ec = c,\ c = const$
2. $E(a\xi + b) = a \cdot E\xi + b,\ a, b = const$
3. $E(\xi_1 + \xi_2) = E\xi_1 + E\xi_2$
4. $E[\xi_1 \cdot \xi_2] = E\xi_1 \cdot E\xi_2$
   ξ_1, ξ_2 *are independent stochastic variables*
5. $\xi \geq 0 \implies M\xi \geq 0$
   $\xi \leq \eta \implies E\xi \leq E\eta$
6. $|E\xi| \leq E|\xi|$

Proof. 4. Let ξ_1, ξ_2 - absolutely continuous stochastic variables with densities

$$f_{\xi_1}(x), f_{\xi_2}(y).$$

$$E[\xi_1, \xi_2] = \iint\limits_{\mathbb{R}^2} x \cdot y \cdot f_{(\xi_1,\xi_2)}(x, y)dxdy =$$
$$= \iint\limits_{\mathbb{R}^2} x \cdot y \cdot f_{\xi_1}(x) \cdot f_{\xi_2}(y)dxdy =$$
$$= \int\limits_{\mathbb{R}} xf_{\xi_1}(x)dx \int\limits_{\mathbb{R}} yf_{\xi_2}(y)dy = E\xi_1 \cdot E\xi_2.$$

□

Remark. *For arbitrary number of stochastic variables:*

$$E(\xi_1 + \cdots + \xi_n) = \sum_{i=1}^{n} E\xi_i.$$

$$E(\xi_1 \cdot \cdots \cdot \xi_n) = \prod_{i=1}^{n} E\xi_i.$$

for $\xi_1, \ldots, \xi_n$ that are independent together.

Example. *let $\xi \sim Bin(n, p)$; $\quad E\xi - ?$*

$$P(\xi = k) = C_n^k p^k (1 - p)^{n-k}, \; k = \overline{0, n}.$$

$$M\xi = \sum_{k=0}^{n} k \cdot C_n^k p^k (1 - p)^{n-k}.$$

Using:

$$\xi = \sum_{i=1}^{n} \xi_i \text{ where } \xi_i \sim B(p) : .$$

$$P(\xi_i = 1) = p; P(\xi_i = 0) = 1 - pM\xi_i = 1 \cdot p + 0 \cdot (1 - p) = p.$$

Then:

$$M\xi = \sum_{i=1}^{n} M\xi_i = n \cdot p.$$

1.3 Dispersion

Definition 2. *Dispersion of stochastic variable is called a number*

$$\mathcal{D}\xi = M(\xi - M\xi)^2.$$

Remark.

$$\mathcal{D}\xi = M(\xi^2 - 2M\xi \cdot \xi + (M\xi)^2) = M\xi^2 - 2M\xi \cdot M\xi + (M\xi)^2 = M\xi^2 - (M\xi)^2.$$

$$\mathcal{D}\xi = M\xi^2 - (M\xi)^2 \tag{1.3}$$

Definition 3. *Number $M\xi^2$ is called second momentum os stochastic variable ξ.*

Example.

.

$$\xi \sim B(p);\ M\xi = p;$$
$$M\xi^2 = 1 \cdot P(\xi = 1) + 0 \cdot P(\xi = 0) = p\mathcal{D}\xi = p - p^2 = p(1-p).$$

Example.

$$\xi \sim U(a,b);\ M\xi = \frac{a+b}{2}$$
$$M\xi^2 = \int_a^b x^2 \frac{1}{b-a} dx = \frac{b^2 - a^2}{3(b-a)} = \frac{a^2 + ab + b^2}{3}$$
$$\mathcal{D}\xi = \frac{a^2 + ab + b^2}{3} - \frac{(a+b)^2}{4} = \frac{4(a^2 + ab + b^2) - 3(a+b)^2}{12} = \frac{(b-a)^2}{12}$$

Theorem 2 (properties of dispersion)**.** *1.*

$$\mathcal{D}\xi \geq 0$$
$$\mathcal{D}\xi = 0 \iff \xi = c = const$$

2.

$$\mathcal{D}(a\xi + b) = a^2 \cdot \mathcal{D}\xi$$

3. If ξ_1 and ξ_2 are independent, then

$$\mathcal{D}(\xi_1 + \xi_2) = \mathcal{D}\xi_1 + \mathcal{D}\xi_2$$

Proof. 1.

$$\mathcal{D}\xi = M(\xi - M\xi)^2$$
$$(\xi - M\xi)^2 \geq 0 \implies M(\xi - M\xi)^2 \geq 0$$
$$M(\xi - M\xi)^2 = 0 \iff (\xi - M\xi)^2 = 0 \iff \xi = M\xi = const$$

2.

$$\mathcal{D}(a\xi + b) = M((a\xi + b) - m(a\xi + b))^2 = M(a\xi + b - aM\xi - b)^2 = Ma^2(\xi - m\xi)^2 =$$
$$= a^2 \cdot M(\xi - M\xi)^2 = a^2 \cdot \mathcal{D}\xi.$$

3. Let ξ_1 and ξ_1 independent.

$$\mathcal{D}(\xi_1 + \xi_2) = M(\xi_1 + \xi_2 - M(\xi_1 + \xi_2))^2 = M((\xi_1 - M\xi_1) + (\xi_2 - M\xi_2))^2 =$$
$$= M((\xi_1 - M\xi_1)^2 + 2(\xi_1 - M\xi_1)(\xi_2 - M\xi_2) + (\xi_2 - M\xi_2)^2) =$$
$$= \mathcal{D}\xi_1 + 2 \cdot M[(\xi_1 - M\xi_1)(\xi_2 - M\xi_2)] + \mathcal{D}\xi_2.$$
$$\xi_1, \xi_2 \text{ independent } \implies \mathcal{D}\xi_1 + \mathcal{D}\xi_2 + 2M(\xi_1 - M\xi_1) \cdot M(\xi_2 - M\xi_2) = \mathcal{D}\xi_1 + \mathcal{D}\xi_2.$$

□

Example.

$$\xi \sim \textit{Bin}(n,p);\ M\xi = n \cdot p;\ \mathcal{D}\xi = ?$$
$$M\xi^2 = \sum_{k=0}^{n} k^2 \cdot C_n^k \cdot p^k \cdot (1-p)^{n-k}$$
$$\xi = \sum_{i=1}^{n} \xi_i,\ \xi_i \sim B(p),\ \xi_2, \dots, \xi_n \text{ - \textit{independent}}$$
$$\mathcal{D}\xi = \sum_{i=1}^{n} \mathcal{D}\xi_i = \sum_{i=1}^{n} p \cdot (1-p) = np(1-p).$$

Remark.

$$M\xi = \underset{a}{\textit{argmin}} M(\xi - a)^2.$$

$$M(\xi - a)^2 = M((\xi - M\xi) + (M\xi - a))^2 =$$
$$M(\xi - M\xi)^2 + 2(M\xi - a)M(\xi - M\xi) + (M\xi - a)^2 =$$
$$\mathcal{D}\xi + (M\xi - a)^2 \geq \mathcal{D}\xi$$
$$\textit{moreover } M(\xi - a)^2 = \mathcal{D}\xi \iff (M\xi - a)^2 = 0$$
$$\implies a = M\xi.$$

Example. *Numerical characteristics of the main probability distributions*

1. $\xi \sim B(p), M\xi = p, \mathcal{D}\xi = p(1-p)$
2. $\xi \sim \textit{Bin}(n,p), M\xi = np,\ \ \mathcal{D}\xi = np(1-p)$
3. $\xi \sim \textit{Poiss}(\lambda)$

$$M\xi = \sum_{k=0}^{\infty} k \cdot \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \cdot \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} = \lambda e^{-\lambda} \cdot \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = \lambda e^{-\lambda} \cdot e^{\lambda} = \lambda$$
$$M\xi^2 = \sum_{k=0}^{\infty} k^2 \cdot \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=1}^{\infty} k \frac{\lambda^k}{(k-1)!} =$$
$$e^{-\lambda} \sum_{k=1}^{\infty} ((k-1)+1) \cdot \frac{\lambda^k}{(k-1)!} = e^{-\lambda}(\lambda^2 \sum_{k=2}^{\infty} \frac{\lambda^{k-1}}{(k-2)!} + \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!}) =$$
$$= \lambda^2 e^{-\lambda} e^{\lambda} + \lambda \cdot e^{-\lambda} \cdot e^{\lambda} = \lambda^2 + \lambda$$
$$\mathcal{D}\xi = \lambda^2 + \lambda - \lambda^2 = \lambda$$

4. $\xi \sim \textit{Geom}(p) : p(\xi = k) = (1-p)^k,\ k = 0, 1, \dots$

$$M\xi = \sum_{k=0}^{\infty} k \cdot (1-p)^k \cdot p$$

$$\sum_{k=0}^{\infty} (1-p)^k = \frac{1}{p}$$

$$\sum_{k=0}^{\infty} k(1-p)^{k-1} = \frac{1}{p^2} \quad | \, p \cdot (1-p)$$

$$\sum_{k=0}^{\infty} k(1-p)^k \cdot p = \frac{1-p}{p}$$

$$M\xi^2 = \sum_{k=0}^{\infty} k^2 (1-p)^k \cdot p$$

$$\sum_{k=0}^{\infty} k(1-p)^k = \frac{1-p}{p^2} = \frac{1}{p^2} - \frac{1}{p}$$

$$\sum_{k=0}^{\infty} k^2(1-p)^{k-1} = \frac{2}{p^3} - \frac{1}{p^2} \quad | \, (1-p) \cdot p$$

$$\sum_{k=0}^{\infty} k^2(1-p)^k \cdot p = \frac{2(1-p)}{p^2} - \frac{1-p}{p}$$

$$\mathcal{D} = \frac{2(1-p)}{p^2} - \frac{1-p}{p} - \left(\frac{1-p}{p}\right)^2 = \frac{2(1-p)}{p^2} - \frac{1-p}{p}\left(1 + \frac{1-p}{p}\right) =$$
$$= \frac{2(1-p)}{p^2} - \frac{1-p}{p^2} = \frac{1-p}{p^2};$$
$$\mathcal{D}\xi = \frac{1-p}{p^2}; \quad M\xi = \frac{1-p}{p}$$

5. $\xi \sim U(a, b); \quad M\xi = \frac{a+b}{2}; \quad \mathcal{D}\xi = \frac{(b-a)^2}{12}$

6. $\xi \sim \textit{Exp}(\lambda) : \ f_\xi(x) = \lambda e^{-\lambda x} \cdot \mathbb{1}(x \geq 0)$

$$M\xi = \int_0^\infty x\lambda e^{-\lambda x} dx = -\int_0^\infty x d e^{-\lambda x} = -x \cdot e^{-\lambda x} \ |_0^\infty + \int_0^\infty e^{-\lambda x} dx = \frac{1}{\lambda}.$$

$$M\xi^2 = \int_0^\infty x^2 \lambda e^{-\lambda x} dx = \int_0^\infty x^2 d e^{-\lambda x} = \int_0^\infty 2x e^{-\lambda x} dx = \frac{2}{\lambda} \int_0^\infty x\lambda e^{-\lambda x} dx = \frac{2}{\lambda^2}$$

$$\mathcal{D}\xi = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2} \quad M\xi = \frac{1}{\lambda}$$

7. $\xi \sim \textit{N}(a, \sigma^2);$

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

$$M\xi = a; \quad \mathcal{D}\xi = \sigma^2$$

$$M\xi = \int\limits_{\mathbb{R}} x \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{(x-a)^2}{2\sigma^2}} dx =$$

$$\frac{x-a}{\sigma} = z$$

$$dx = \sigma dz$$

$$= \frac{1}{\sqrt{2\pi}} \int\limits_{\mathbb{R}} (\sigma z + a) e^{-z^2/2} dz == \frac{\sigma}{\sqrt{2\pi}} \cdot \int\limits_{\mathbb{R}} z e^{-z^2/2} dz + \frac{a}{\sqrt{2\pi}} \int\limits_{\mathbb{R}} e^{-z^2/2} dz = a$$

$$\mathcal{D}\xi = M(\xi - M\xi)^2 = M(\xi - a)^2 = \int\limits_{\mathbb{R}} (x-a)^2 \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}} dx =$$

$$\frac{x-a}{\sigma} = z$$

$$dx = \sigma dz$$

$$= \frac{1}{\sqrt{2\pi}} \int\limits_{\mathbb{R}} \sigma^2 z^2 \cdot e^{-z^2/2} dz = \frac{2\sigma^2}{\sqrt{2\pi}} \int\limits_0^\infty z^2 e^{-z^2/2} dz = -\frac{2\sigma^2}{\sqrt{2\pi}} \int\limits_0^\infty z d e^{-z^2/2} =$$

$$= -\frac{2\sigma^2}{\sqrt{2\pi}} z \cdot e^{-z^2/2} \Big|_0^\infty + \frac{2\sigma^2}{\sqrt{2\pi}} \int\limits_0^\infty e^{-z^2/2} dz = \sigma^2 \cdot \frac{2}{\sqrt{2\pi}} \int\limits_0^\infty e^{-z^2/2} dz =$$

$$= \sigma^2 \cdot \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{+\infty} e^{-z^2/2} dz = \sigma^2$$

Chapter 2

Covariance of random variables. Correlation coefficient.

Consider $\xi = (\xi_1, \xi_2)$ - random vector.

Definition 4. *Covariation of stochastic variables ξ_1, ξ_2 is a number:*

$$\textit{cov}(\xi_1, \xi_2) = M[(\xi_1 - M\xi_1) \cdot (\xi_2 - M - \xi_2)] \tag{2.1}$$

(assuming that $M\xi_i$ exist)

If ξ_1, ξ_2 are discrete random variables, then

$$\text{cov}(\xi_1, \xi_2) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (x_i - M\xi_1) \cdot (y_j - M\xi_2) \cdot P(\xi_1 = x_i, \xi_2 = y_j). \tag{2.2}$$

If ξ_1, ξ_2 have common distribution density $f_\xi(x, y)$, then

$$\text{cov}(\xi_1, \xi_2) = \iint\limits_{\mathbb{R}^2} (x - M\xi_1)(y - M\xi_2) f_\xi(x, y) dx dy \tag{2.3}$$

From definition implies:

$$\text{cov}(\xi_1, \xi_2) = M[\xi_1 \cdot \xi_2 - \xi_1 \cdot M\xi_2 - \xi_2 \cdot M\xi_1 + M\xi_1 \cdot M\xi_2] =$$
$$M[\xi_1 \cdot \xi_2] - M\xi_2 \cdot M\xi_1 - M\xi_1 \cdot M\xi_2 + M\xi_1 \cdot M\xi_2 = M[\xi_1 \cdot \xi_2] - M\xi_1 \cdot M\xi_2.$$

$$\text{cov}(\xi_1, \xi_2) = M[\xi_1 \cdot \xi_2] - M\xi_1 \cdot M\xi_2 \tag{2.4}$$

Proposition 3. *If ξ_1, ξ_2 are independent, then*

$$\textit{cov}(\xi_1, \xi_2) = 0.$$

It is said, that ξ_1, ξ_2 are uncorrelated.

Indeed, if ξ_1, ξ_2 are independent, then from the properties of expectation:

$$M[\xi_1 \cdot \xi_2] = M\xi_1 \cdot M\xi_2.$$

then

$$\textit{cov}(\xi_1, \xi_2) = M\xi_1 \cdot M\xi_2 - M\xi_1 \cdot M\xi_2 = 0.$$

Inverse statement is not true: from uncorrelated does not implies independency.

Chapter 3

Inequalities. The law of large numbers in the form of Chebyshev. Borel-Cantelli lemma