Probability theory

Lecture 6: Variance and covariance

Maksim Zhukovskii

MIPT

Definitions

▶ Variance of ξ is $Var\xi = E(\xi - E\xi)^2$.

Definitions

▶ Variance of
$$\xi$$
 is $Var\xi = E(\xi - E\xi)^2$.

► Covariance of
$$\xi, \eta$$
 is $cov(\xi, \eta) = E((\xi - E\xi)(\eta - E\eta)).$

Definitions

▶ Variance of
$$\xi$$
 is $Var\xi = E(\xi - E\xi)^2$.

► Covariance of ξ, η is $cov(\xi, \eta) = E((\xi - E\xi)(\eta - E\eta)).$

Clearly, $Var \xi = cov(\xi, \xi)$.

1.
$$Var\xi = E\xi^2 - (E\xi)^2$$
.

1.
$$Var\xi = E\xi^2 - (E\xi)^2$$
. Proof.

1.
$$Var \xi = E \xi^2 - (E \xi)^2$$
.
Proof. $Var \xi = E(\xi - E \xi)^2 =$

1.
$$Var\xi = E\xi^2 - (E\xi)^2$$
.

Proof.
$$Var\xi = E(\xi - E\xi)^2 = E[\xi^2 - 2\xi E\xi + (E\xi)^2] =$$

1.
$$Var \xi = E \xi^2 - (E \xi)^2$$
.

Proof.
$$Var\xi = E(\xi - E\xi)^2 = E[\xi^2 - 2\xi E\xi + (E\xi)^2] = E\xi^2 - 2(E\xi)^2 + (E\xi)^2 = E\xi^2 - 2(E\xi)^2 + (E\xi)^2 = E\xi^2 - E(\xi)^2 + (E\xi)^2 = E(\xi)^2 + (E\xi)^2 + (E\xi)^2 = E(\xi)^2 + (E\xi)^2 + (E\xi)^2 + (E\xi)^2 = E(\xi)^2 + (E\xi)^2 + (E\xi)^$$

1.
$$Var \xi = E \xi^2 - (E \xi)^2$$
.

Proof.
$$Var\xi = E(\xi - E\xi)^2 = E[\xi^2 - 2\xi E\xi + (E\xi)^2] = E\xi^2 - 2(E\xi)^2 + (E\xi)^2 = E\xi^2 - (E\xi)^2.$$

1.
$$Var\xi = E\xi^2 - (E\xi)^2$$
.
Proof. $Var\xi = E(\xi - E\xi)^2 = E(\xi^2 - \xi^2)^2$

Proof.
$$Var\xi = E(\xi - E\xi)^2 = E[\xi^2 - 2\xi E\xi + (E\xi)^2] = E\xi^2 - 2(E\xi)^2 + (E\xi)^2 = E\xi^2 - (E\xi)^2.$$

2.
$$Var \xi \ge 0$$
.

1.
$$\operatorname{Var} \xi = \operatorname{\mathsf{E}} \xi^2 - (\operatorname{\mathsf{E}} \xi)^2$$
.
Proof. $\operatorname{Var} \xi = \operatorname{\mathsf{E}} (\xi - \operatorname{\mathsf{E}} \xi)^2 = \operatorname{\mathsf{E}} [\xi^2 - 2\xi \operatorname{\mathsf{E}} \xi + (\operatorname{\mathsf{E}} \xi)^2] = \operatorname{\mathsf{E}} \xi^2 - 2(\operatorname{\mathsf{E}} \xi)^2 + (\operatorname{\mathsf{E}} \xi)^2 = \operatorname{\mathsf{E}} \xi^2 - (\operatorname{\mathsf{E}} \xi)^2.$

Since
$$(\xi - E)$$

2. $Var \varepsilon > 0$.

Since $(\xi - \mathsf{E}\xi)^2 \ge 0$.

1.
$$cov(\xi, \eta) = \mathsf{E}(\xi \eta) - \mathsf{E}\xi \mathsf{E}\eta$$
.

1.
$$cov(\xi, \eta) = E(\xi \eta) - E\xi E\eta$$
.
Proof.

1.
$$cov(\xi, \eta) = E(\xi \eta) - E\xi E\eta$$
.

Proof.
$$cov(\xi, \eta) = E(\xi - E\xi)(\eta - E\eta)$$

1.
$$cov(\xi, \eta) = E(\xi \eta) - E\xi E\eta$$
.

Proof.
$$cov(\xi, \eta) = E(\xi - E\xi)(\eta - E\eta)$$

= $E[\xi \eta - \xi E\eta - \eta E\xi + E\xi E\eta] =$

1.
$$cov(\xi, \eta) = E(\xi \eta) - E\xi E\eta$$
.

Proof.
$$cov(\xi, \eta) = E(\xi - E\xi)(\eta - E\eta)$$

= $E[\xi\eta - \xi E\eta - \eta E\xi + E\xi E\eta] =$
 $E\xi\eta - 2E\xi E\eta + E\xi E\eta =$

1.
$$\operatorname{cov}(\xi, \eta) = \mathsf{E}(\xi \eta) - \mathsf{E}\xi \mathsf{E}\eta$$
.

Proof.
$$cov(\xi, \eta) = E(\xi - E\xi)(\eta - E\eta)$$

= $E[\xi \eta - \xi E\eta - \eta E\xi + E\xi E\eta] =$
 $E\xi \eta - 2E\xi E\eta + E\xi E\eta = E\xi \eta - E\xi E\eta$.

1.
$$cov(\xi, \eta) = E(\xi \eta) - E\xi E\eta$$
.
Proof. $cov(\xi, \eta) = E(\xi - E\eta)$

Proof.
$$cov(\xi, \eta) = E(\xi - E\xi)(\eta - E\eta)$$

= $E[\xi\eta - \xi E\eta - \eta E\xi + E\xi E\eta] =$
 $E\xi\eta - 2E\xi E\eta + E\xi E\eta = E\xi\eta - E\xi E\eta$.

2. If ξ, η are independent, then $cov(\xi, \eta) = 0$.

1.
$$cov(\xi, \eta) = E(\xi\eta) - E\xi E\eta$$
.
Proof. $cov(\xi, \eta) = E(\xi - E\xi)(\eta - E\eta)$
 $= E[\xi\eta - \xi E\eta - \eta E\xi + E\xi E\eta] =$
 $E\xi\eta - 2E\xi E\eta + E\xi E\eta = E\xi\eta - E\xi E\eta$.

2. If
$$\xi, \eta$$
 are independent, then $cov(\xi, \eta) = 0$.

Since $E\xi\eta = E\xi E\eta$.

3. Covariance is **symmetric**: $cov(\xi, \eta) = cov(\eta, \xi)$.

3. Covariance is **symmetric**:
$$cov(\xi, \eta) = cov(\eta, \xi)$$
.

4. Covariance is **bilinear**:
$$cov(a_1\xi_1 + a_2\xi_2, \eta) = a_1cov(\xi_1, \eta) + a_2cov(\xi_2, \eta).$$

3. Covariance is **symmetric**:
$$cov(\xi, \eta) = cov(\eta, \xi)$$
.

4. Covariance is **bilinear**:
$$cov(a_1\xi_1 + a_2\xi_2, \eta) = a_1cov(\xi_1, \eta) + a_2cov(\xi_2, \eta)$$
. Proof.

3. Covariance is **symmetric**:
$$cov(\xi, \eta) = cov(\eta, \xi)$$
.

$$cov(a_1\xi_1 + a_2\xi_2, \eta) =
a_1cov(\xi_1, \eta) + a_2cov(\xi_2, \eta).$$

Proof.
$$\operatorname{cov}(a_1\xi_1 + a_2\xi_2, \eta) =$$

3. Covariance is symmetric:
$$cov(\xi, \eta) = cov(\eta, \xi)$$
.

$$cov(\xi, \eta) = cov(\eta, \xi).$$

 $cov(a_1\xi_1 + a_2\xi_2, \eta) =$

 $a_1 cov(\xi_1, \eta) + a_2 cov(\xi_2, \eta).$

Proof. $cov(a_1\xi_1 + a_2\xi_2, \eta) =$

 $E(a_1\xi_1\eta + a_2\xi_2\eta) - E(a_1\xi_1 + a_2\xi_2)E\eta =$

3. Covariance is **symmetric**:
$$cov(\xi, \eta) = cov(\eta, \xi)$$
.

$$cov(\xi, \eta) = cov(\eta, \xi).$$
4. Covariance is **bilinear**:

$$cov(a_1\xi_1 + a_2\xi_2, \eta) =$$
 $a_1cov(\xi_1, \eta) + a_2cov(\xi_2, \eta).$

Proof. $cov(a_1\xi_1 + a_2\xi_2, \eta) =$
 $E(a_1\xi_1\eta + a_2\xi_2\eta) - E(a_1\xi_1 + a_2\xi_2)E\eta =$
 $a_1E\xi_1\eta + a_2E\xi_2\eta - a_1E\xi_1E\eta - a_2E\xi_2E\eta$

3. Covariance is **symmetric**:
$$cov(\xi, \eta) = cov(\eta, \xi)$$
.

$$cov(\xi, \eta) = cov(\eta, \xi).$$

4. Covariance is **bilinear**:

$$cov(a_1\xi_1 + a_2\xi_2, \eta) =$$

$$a_1cov(\xi_1, \eta) + a_2cov(\xi_2, \eta).$$
Proof.
$$cov(a_1\xi_1 + a_2\xi_2, \eta) =$$

$$E(a_1\xi_1\eta + a_2\xi_2\eta) - E(a_1\xi_1 + a_2\xi_2)E\eta =$$

$$a_1E\xi_1\eta + a_2E\xi_2\eta - a_1E\xi_1E\eta - a_2E\xi_2E\eta$$

 $= a_1 cov(\xi_1, \eta) + a_2 cov(\xi_2, \eta).$

$$cov(\xi, \eta) = cov(\eta, \xi).$$
4. Covariance is **bilinear**:

Claim

$$Var(\xi_1 + \ldots + \xi_n) = Var\xi_1 + \ldots + Var\xi_n + \sum_{i \neq j} cov(\xi_i, \xi_j).$$

Claim

$$Var(\xi_1 + \ldots + \xi_n) = Var\xi_1 + \ldots + Var\xi_n + \sum_{i \neq j} cov(\xi_i, \xi_j).$$

Proof.

Claim

$$\operatorname{Var}(\xi_1 + \ldots + \xi_n) =$$

Proof. $Var(\xi_1 + \ldots + \xi_n) =$

$$\operatorname{Var}\xi_1 + \ldots + \operatorname{Var}\xi_n + \sum_{i\neq j} \operatorname{cov}(\xi_i, \xi_j).$$

Claim

$$Var(\xi_1 + \ldots + \xi_n) = Var\xi_1 + \ldots + Var\xi_n + \sum_{i \neq j} cov(\xi_i, \xi_j).$$

$$Var\xi_1 + \ldots + Var\xi_n$$

Proof. $Var(\xi_1 + \ldots + \xi_n) =$

 $cov(\xi_1 + \ldots + \xi_n, \xi_1 + \ldots + \xi_n) =$

Claim

$$Var(\xi)$$

$$\operatorname{Var}(\xi_1 + \ldots + \xi_n) =$$

$$\operatorname{Var}\xi_1 + \ldots + \operatorname{Var}\xi_n + \sum_{i \neq j} \operatorname{cov}(\xi_i, \xi_j).$$

 $\operatorname{cov}(\xi_1 + \ldots + \xi_n, \xi_1 + \ldots + \xi_n) =$

 $\sum_{i=1}^{n} \operatorname{cov}(\xi_i, \xi_i) + \sum_{i \neq i} \operatorname{cov}(\xi_i, \xi_i) =$

Proof.
$$Var(\xi_1 + \ldots + \xi_n) =$$

Claim

$$\operatorname{Var}(\xi_1 + \ldots + \xi_n) =$$

$$Var(\xi)$$

$$\operatorname{Var}\xi_1 + \ldots + \operatorname{Var}\xi_n + \sum_{i \neq j} \operatorname{cov}(\xi_i, \xi_j).$$

$$+\ldots+$$
 Va

Proof.
$$Var(\xi_1 + \ldots + \xi_n) =$$

 $\operatorname{Var}\xi_1 + ... + \operatorname{Var}\xi_n + \sum_{i \neq i} \operatorname{cov}(\xi_i, \xi_i).$

$$+\xi_n$$
) =

$$\operatorname{cov}(\xi_1 + \ldots + \xi_n, \xi_1 + \ldots + \xi_n) = \sum_{i=1}^n \operatorname{cov}(\xi_i, \xi_i) + \sum_{i \neq i} \operatorname{cov}(\xi_i, \xi_i) =$$

Claim

$$Var(\xi_1 + \ldots + \xi_n) = Var\xi_1 + \ldots + Var\xi_n + \sum_{i \neq j} cov(\xi_i, \xi_j).$$

$$\operatorname{Proof.} \operatorname{Var}(\xi_1 + \ldots + \xi_n) = \\ \operatorname{cov}(\xi_1 + \ldots + \xi_n, \xi_1 + \ldots + \xi_n) = \\$$

Proof.
$$Var(\xi_1 + \ldots + \xi_n) =$$

$$cov(\xi_1 + \ldots + \xi_n, \xi_1 + \ldots + \xi_n) =$$

$$\sum_{n=1}^{n} cov(\xi_1, \xi_2) + \sum_{n=1}^{n} cov(\xi_1, \xi_2) =$$

$$cov(\xi_1 + \ldots + \xi_n, \xi_1 + \ldots + \xi_n) =$$

$$\sum_{i=1}^n cov(\xi_i, \xi_i) + \sum_{i \neq j} cov(\xi_i, \xi_j) =$$

$$Var\xi_1 + \ldots + Var\xi_n + \sum_{i \neq j} cov(\xi_i, \xi_j).$$
If $\xi_1 = \xi_n$ are pairwise independent, then

$$\operatorname{Cov}(\xi_{1} + \ldots + \xi_{n}, \xi_{1} + \ldots + \xi_{n}) = \sum_{i=1}^{n} \operatorname{cov}(\xi_{i}, \xi_{i}) + \sum_{i \neq j} \operatorname{cov}(\xi_{i}, \xi_{j}) = \operatorname{Var}\xi_{1} + \ldots + \operatorname{Var}\xi_{n} + \sum_{i \neq j} \operatorname{cov}(\xi_{i}, \xi_{j}).$$
If ξ_{1}, \ldots, ξ_{n} are pairwise independent, then
$$\operatorname{Var}(\xi_{1} + \ldots + \xi_{n}) = \operatorname{Var}(\xi_{1}) + \ldots + \operatorname{Var}(\xi_{n}).$$

Proof.
$$\operatorname{Var}(\xi_1 + \ldots + \xi_n) = \operatorname{cov}(\xi_1 + \ldots + \xi_n, \xi_1 + \ldots + \xi_n) = \sum_{n=0}^{n} \operatorname{cov}(\xi_n, \xi_n) + \sum_{n=0}^{n} \operatorname{cov}(\xi_n, \xi_n) = \operatorname{cov}(\xi_n,$$

Variance of a binomial random variable

 $\xi \sim Bin(n, p)$.

Variance of a binomial random variable $\xi \sim Bin(n, p)$.

$$\mathsf{E}\xi^2 = \sum_{k=0}^n k^2 \binom{n}{k} p^k (1-p)^{n-k} =$$

Variance of a binomial random variable $\xi \sim Bin(n, p)$.

$$\mathsf{E}\xi^2 = \sum_{k=0}^n k^2 \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^n k(k-1) \binom{n}{k} p^k (1-p)^{n-k} + \sum_{k=0}^n k(k-1)$$

$$\sum_{k=1}^{n} k(k-1) \binom{n}{k} p^{k} (1-p)^{n-k} +$$

$$\sum_{k=1}^{n} k \binom{n}{k} p^{k} (1-p)^{n-k} =$$

$\xi \sim Bin(n, p)$.

 $\sum_{k=1}^{n} k(k-1) \binom{n}{k} p^{k} (1-p)^{n-k} + np =$

Variance of a binomial random variable

$$\mathsf{E}\xi^2 = \sum_{k=0}^n k^2 \binom{n}{k} p^k (1-p)^{n-k} =$$

$$\sum_{k=1}^{n} k(k-1) \binom{n}{k} p^{k} (1-p)^{n-k} +$$

$$\sum_{k=1}^{n} k(k-1) \binom{n}{k} p^{k} (1-p)^{n-k} -$$

$$\frac{\sum_{k=1}^{n} k(k-1)\binom{n}{k} p^{k} (1-p)^{n-k}}{\sum_{k=1}^{n} k\binom{n}{k} p^{k} (1-p)^{n-k}} =$$

$\xi \sim Bin(n, p)$. $\mathsf{E}\xi^2 = \sum_{k=0}^n k^2 \binom{n}{k} p^k (1-p)^{n-k} =$

Variance of a binomial random variable

 $n(n-1)p^2 \sum_{k=2}^{n} {n-2 \choose k-2} p^{k-2} (1-p)^{n-k} + np =$

$$\sum_{k=1}^{n} k(k-1) \binom{n}{k} p^{k} (1-p)^{n-k} + \sum_{k=1}^{n} k \binom{n}{k} p^{k} (1-p)^{n-k} =$$

$$\sum_{k=1}^{k=1} k(k-1)\binom{n}{k} p^{k} (1-p)^{n-k} =$$

$$\sum_{k=1}^{n} k \binom{n}{k} p^{k} (1-p)^{n-k} =$$

$$\sum_{k=1}^{n} k \binom{n}{k} p^{k} (1-p)^{n-k} + np =$$

$\mathcal{E} \sim Bin(n, p)$. $\mathsf{E}\xi^2 = \sum_{k=0}^n k^2 \binom{n}{k} p^k (1-p)^{n-k} =$

$$\sum_{k=1}^{n} k(k-1) \binom{n}{k} p^{k} (1-p)^{n-k} +$$

$$\sum_{k=1}^{n} k \binom{n}{k} p^{k} (1-p)^{n-k} =$$

$$\sum_{k=1}^{n} k(k-1) \binom{n}{k} p^{k} (1-p)^{n-k} +$$

 $n(n-1)p^2 + np$.

$$\sum_{k=1}^{n} k \binom{n}{k} p^{k} (1-p)^{n-k} =$$

$$\sum_{k=1}^{n} k(k-1) \binom{n}{k} p^{k} (1-p)^{n-k} + np =$$

$$n(n-1) p^{2} \sum_{k=2}^{n} \binom{n-2}{k-2} p^{k-2} (1-p)^{n-k} + np =$$

Variance of a binomial random variable

$$\sum_{k=1}^{n} k(k-1) \binom{n}{k} p^{k} (1-p)^{n-k} + np = n(n-1)p^{2} \sum_{k=2}^{n} \binom{n-2}{k-2} p^{k-2} (1-p)^{n-k} + np$$

$$\sum_{k=1}^{n} k(k-1) \binom{n}{k} p^{k} (1-p)^{n-k} + np =$$

$$n(n-1) p^{2} \sum_{k=2}^{n} \binom{n-2}{k-2} p^{k-2} (1-p)^{n-k} + np =$$

$$n(n-1) p^{2} \sum_{j=0}^{n-2} \binom{n-2}{j} p^{j} (1-p)^{n-2-j} + np =$$

 $\mathcal{E} \sim Bin(n, p)$. $\mathsf{E}\xi^2 = \sum_{k=0}^n k^2 \binom{n}{k} p^k (1-p)^{n-k} =$

Variance of a binomial random variable

$$\sum_{k=1}^{n} k(k-1) \binom{n}{k} p^{k} (1-p)^{n-k} +$$

$$\sum_{k=1}^{n} k \binom{n}{k} p^{k} (1-p)^{n-k} =$$

$$\sum_{k=1}^{n} k(k-1) \binom{n}{k} p^{k} (1-p)^{n-k} + np =$$

$$p(n-1) p^{2} \sum_{k=1}^{n} \binom{n-2}{k} p^{k-2} (1-p)^{n-k} + 1$$

 $n(n-1)p^2 \sum_{k=2}^{n} {n-2 \choose k-2} p^{k-2} (1-p)^{n-k} + np =$ $n(n-1)p^2 \sum_{i=0}^{n-2} {n-2 \choose i} p^j (1-p)^{n-2-j} + np = 0$ $n(n-1)p^2 + np$.

$$Var\xi = n^2p^2 - np^2 + np - n^2p^2 = np(1-p).$$

$$\xi \sim Bin(n,p)$$
.

$$\xi \sim Bin(n, p)$$
.

Let $\xi_1, ..., \xi_n$ are independent Bernoulli random variables.

$$\xi \sim Bin(n, p)$$
.

Let $\xi_1, ..., \xi_n$ are independent Bernoulli random variables.

Then
$$\xi \stackrel{d}{=} \xi_1 + \ldots + \xi_n$$
.

$$\xi \sim Bin(n, p)$$
.

Let $\xi_1, ..., \xi_n$ are independent Bernoulli random variables.

Then
$$\xi \stackrel{d}{=} \xi_1 + \ldots + \xi_n$$
.

$$Var\xi = nVar\xi_1 = n(E\xi_1^2 - (E\xi_1)^2) = np(1-p).$$

• $\xi \sim U\{1,\ldots,n\}$

•
$$\xi \sim \mathcal{N}(0,1)$$

$$\nabla \text{ar} \xi \sim U\{1, \dots, n\}
\text{Var} \xi = \frac{1}{n} \sum_{i=1}^{n} i^2 - \left(\frac{n+1}{2}\right)^2 = \frac{(n+1)(2n+1)}{n} - \frac{(n+1)^2}{n} = \frac{n^2-1}{n}$$

$$\operatorname{Var}\xi = \frac{1}{n} \sum_{i=1}^{m} I^2 - \left(\frac{n+1}{2}\right) = \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4} = \frac{n^2 - 1}{12}.$$

 $\operatorname{Var} \xi = \int_{-\infty}^{\infty} \frac{x^2}{\sqrt{2\pi}} e^{-x^2/2} dx =$

$$\operatorname{Var} \xi = \frac{1}{n} \sum_{i=1}^{n} i^{2} - \left(\frac{n+1}{2}\right)^{2} = \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^{2}}{4} = \frac{n^{2}-1}{12}.$$

$$\begin{aligned}
& \text{Var}\xi = \frac{1}{n} \sum_{i=1}^{n} I^{2} - \left(\frac{n+2}{2}\right) = \frac{1}{n} \\
& \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^{2}}{4} = \frac{n^{2}-1}{12}. \\
& \blacktriangleright \xi \sim \mathcal{N}(0,1) \\
& \text{Var}\xi = \int_{-\infty}^{\infty} \frac{x^{2}}{\sqrt{2\pi}} e^{-x^{2}/2} dx = \frac{1}{n} \\
\end{aligned}$$

 $-\int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi}} de^{-x^2/2} =$

•
$$\xi \sim U\{1,\ldots,n\}$$

$$\operatorname{Var} \xi = \frac{1}{n} \sum_{i=1}^{n} i^{2} - \left(\frac{n+1}{2}\right)^{2} = \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^{2}}{4} = \frac{n^{2}-1}{12}.$$

•
$$\xi \sim \mathcal{N}(0)$$

$$\operatorname{Var}\xi = \int_{-\infty}^{\infty} \frac{x^{2}}{\sqrt{2\pi}} e^{-x^{2}/2} dx = -\int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi}} de^{-x^{2}/2} = -\frac{x}{\sqrt{2\pi}} e^{-x^{2}/2} \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^{2}/2} dx = -\frac{x}{\sqrt{2\pi}} e^{-x^{2}/2} \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^{2}/2} dx = -\frac{x}{\sqrt{2\pi}} e^{-x} e$$

•
$$\xi \sim U\{1,\ldots,n\}$$

$$\operatorname{Var} \xi = \frac{1}{n} \sum_{i=1}^{n} i^2 - \left(\frac{n+1}{2}\right)^2 = \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4} = \frac{n^2 - 1}{12}.$$

$$\epsilon \sim \mathcal{N}(0,1)$$

$$\begin{aligned}
&\text{Var}\xi = \int_{-\infty}^{\infty} \frac{x^2}{\sqrt{2\pi}} e^{-x^2/2} dx = \\
&- \int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi}} de^{-x^2/2} = \\
&- \frac{x}{\sqrt{2\pi}} e^{-x^2/2} \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = 1.
\end{aligned}$$

Cauchy-Bunyakovsky-Schwarz inequality

Theorem

Let $\mathsf{E}\xi^2<\infty$, $\mathsf{E}\eta^2<\infty$.

Cauchy-Bunyakovsky-Schwarz inequality

Theorem

Let $E\xi^2 < \infty$, $E\eta^2 < \infty$.

Then $E|\xi\eta| < \infty$ and $(E|\xi\eta|)^2 \le E\xi^2 E\eta^2$.

 $\blacktriangleright \ \text{Let} \ \mathsf{E} \xi^2 = 0.$

$$\mathsf{P}(\xi=0)=1$$

▶ Let
$$E\xi^2 = 0$$
.

$$\mathsf{P}(\xi=0)=1\Rightarrow\mathsf{P}(\xi\eta=0)=1$$

▶ Let
$$\mathsf{E}\xi^2 = 0$$
.

$$\mathsf{P}(\xi=0)=1\Rightarrow \mathsf{P}(\xi\eta=0)=1\Rightarrow \mathsf{E}|\xi\eta|=0.$$

Let
$$\mathsf{E}\xi^2=0$$
.
 $\mathsf{P}(\xi=0)=1\Rightarrow \mathsf{P}(\xi\eta=0)=1\Rightarrow$
 $\mathsf{E}|\xi\eta|=0$.

▶ Let
$$E\xi^2 > 0$$
, $E\eta^2 > 0$.

Let
$$\mathsf{E}\xi^2=0$$
.
$$\mathsf{P}(\xi=0)=1\Rightarrow \mathsf{P}(\xi\eta=0)=1\Rightarrow$$

$$\begin{aligned} & \mathsf{E}|\xi\eta| = 0. \\ & \blacktriangleright \mathsf{Let} \; \mathsf{E}\xi^2 > 0, \; \mathsf{E}\eta^2 > 0. \\ & \mathsf{Denote} \; \xi_0 = \xi/\sqrt{\mathsf{E}\xi^2}, \; \eta_0 = \eta/\sqrt{\mathsf{E}\eta^2}. \end{aligned}$$

▶ Let $E\xi^2 = 0$.

 $E|\xi\eta|=0.$

▶ Let $E\xi^2 > 0$, $E\eta^2 > 0$.

 $2|\xi_0\eta_0| \leq \xi_0^2 + \eta_0^2$

 $P(\xi = 0) = 1 \Rightarrow P(\xi \eta = 0) = 1 \Rightarrow$

Denote $\xi_0 = \xi/\sqrt{\mathsf{E}\xi^2}$, $\eta_0 = \eta/\sqrt{\mathsf{E}\eta^2}$.

 $P(\xi = 0) = 1 \Rightarrow P(\xi \eta = 0) = 1 \Rightarrow$

Denote $\xi_0 = \xi/\sqrt{\mathsf{E}\xi^2}$, $\eta_0 = \eta/\sqrt{\mathsf{E}\eta^2}$.

▶ Let $E\xi^2 = 0$.

 $E|\xi\eta|=0.$

▶ Let $E\xi^2 > 0$, $E\eta^2 > 0$.

 $2|\xi_0\eta_0| \le \xi_0^2 + \eta_0^2 \Rightarrow$

 $E(2|\xi_0\eta_0|) \leq E(\xi_0^2 + \eta_0^2) = 2$

- The proof

 $P(\xi = 0) = 1 \Rightarrow P(\xi \eta = 0) = 1 \Rightarrow$

Denote $\xi_0 = \xi/\sqrt{\mathsf{E}\xi^2}$, $\eta_0 = \eta/\sqrt{\mathsf{E}\eta^2}$.

 $E(2|\xi_0\eta_0|) \le E(\xi_0^2 + \eta_0^2) = 2 \Rightarrow$

▶ Let $E\xi^2 = 0$.

 $E|\xi \eta| = 0.$

 $E|\xi_0\eta_0| < 1$

▶ Let $E\xi^2 > 0$, $E\eta^2 > 0$.

 $2|\xi_0\eta_0| \le \xi_0^2 + \eta_0^2 \Rightarrow$

▶ Let $E\xi^2 = 0$.

 $E|\xi \eta| = 0.$

▶ Let $E\xi^2 > 0$, $E\eta^2 > 0$.

 $2|\xi_0\eta_0| \leq \xi_0^2 + \eta_0^2 \Rightarrow$

 $P(\xi = 0) = 1 \Rightarrow P(\xi \eta = 0) = 1 \Rightarrow$

Denote $\xi_0 = \xi/\sqrt{\mathsf{E}\xi^2}$, $\eta_0 = \eta/\sqrt{\mathsf{E}\eta^2}$.

 $E(2|\xi_0\eta_0|) \le E(\xi_0^2 + \eta_0^2) = 2 \Rightarrow$ $E|\xi_0\eta_0| < 1 \Rightarrow E|\xi\eta| < \sqrt{E\xi^2E\eta^2}$.

Markov inequality

Theorem

Let $\mathsf{E}|\xi|<\infty$.

Markov inequality

Theorem

Let $E|\xi| < \infty$.

Then, for every a > 0, $P(|\xi| \ge a) \le \frac{E|\xi|}{a}$.

$$\mathsf{E}|\xi| =$$

$$E|\xi| = E(|\xi|I(|\xi| < a) + |\xi|I(|\xi| \ge a))$$

$$\mathsf{E}|\xi| = \mathsf{E}(|\xi| \mathit{I}(|\xi| < a) + |\xi| \mathit{I}(|\xi| \geq a))$$

$$\geq a \mathsf{E}(I(|\xi| \geq a))$$

$$\mathsf{E}|\xi| = \mathsf{E}(|\xi|I(|\xi| < a) + |\xi|I(|\xi| \ge a))$$

$$\geq a \mathsf{E}(I(|\xi| \geq a)) = a \mathsf{P}(|\xi| \geq a).$$

$$\mathsf{E}|\xi| = \mathsf{E}(|\xi| I(|\xi| < a) + |\xi| I(|\xi| \ge a))$$

$$\geq a\mathsf{E}(I(|\xi|\geq a))=a\mathsf{P}(|\xi|\geq a).$$

Therefore, $P(|\xi \ge a|) \le \frac{E|\xi|}{a}$.

Chebyshev inequality

Theorem

Let $\mathsf{E}\xi^2<\infty$.

Chebyshev inequality

Theorem

Let $\mathsf{E}\xi^2<\infty$.

Then, for every a > 0,

$$P(|\xi - \mathsf{E}\xi| \ge a) \le \frac{\mathrm{Var}\xi}{a^2}.$$

Chebyshev inequality

Theorem

Let
$$\mathsf{E}\xi^2<\infty$$
.

Then, for every a > 0,

$$P(|\xi - E\xi| \ge a) \le \frac{Var\xi}{a^2}.$$

Proof. Follows from Markov inequality by replacing $\xi \to (\xi - \mathsf{E}\xi)^2$, $a \to a^2$.

Jensen's inequality

Theorem

Let $g : \mathbb{R} \to \mathbb{R}$ be convex and $\mathsf{E}|\xi| < \infty$.

Jensen's inequality

Theorem

Let $g : \mathbb{R} \to \mathbb{R}$ be convex and $E|\xi| < \infty$.

Then $g(\mathsf{E}\xi) \leq \mathsf{E}g(\xi)$.

Set
$$x_0 = \mathsf{E}\xi$$
, $x = \xi$.

Set
$$x_0 = \mathsf{E}\xi$$
, $x = \xi$.
Then $g(\xi) \ge g(\mathsf{E}\xi) + \lambda(\mathsf{E}\xi)(\xi - \mathsf{E}\xi)$.

Set
$$x_0 = \mathsf{E}\xi$$
, $x = \xi$.
Then $g(\xi) \geq g(\mathsf{E}\xi) + \lambda(\mathsf{E}\xi)(\xi - \mathsf{E}\xi)$.
Therefore, $\mathsf{E}g(\xi) \geq g(\mathsf{E}\xi)$.

Law of large numbers

Theorem

Let ξ_1, ξ_2, \ldots be pairwise independent identically distributed random variables. Let $\mathsf{E}\xi_1^2 < \infty$.

Law of large numbers

Theorem

Let ξ_1, ξ_2, \ldots be pairwise independent identically distributed random variables.

Let
$$\mathsf{E}\xi_1^2<\infty$$
.

Then, for all $\varepsilon > 0$

$$P(\frac{|S_n-\mathsf{E}S_n|}{n}>\varepsilon) o 0.$$

$$P\left(\frac{|S_n - ES_n|}{n} > \varepsilon\right) \le \frac{\operatorname{Var} S_n}{n^2 \varepsilon^2} =$$

$$\mathsf{P}\left(rac{|\mathcal{S}_n - \mathsf{E}\mathcal{S}_n|}{n} > arepsilon
ight) \leq rac{\mathrm{Var}\mathcal{S}_n}{n^2arepsilon^2} = rac{\mathrm{Var}\xi_1}{arepsilon^2 n}$$

$$\mathsf{P}\left(rac{|S_n - \mathsf{E}S_n|}{n} > arepsilon
ight) \leq rac{\mathrm{Var}S_n}{n^2arepsilon^2} = rac{\mathrm{Var}\xi_1}{arepsilon^2 n} o 0, \quad n o \infty.$$

A generalization

Theorem

Let ξ_1, ξ_2, \dots be a sequence of random variables such that

- - for $i \neq j$, $cov(\xi_i, \xi_i) = 0$,
 - there exists C > 0 such that, for every i, $\mathrm{Var} \xi_i < C$.

A generalization

Theorem

Let ξ_1, ξ_2, \dots be a sequence of random variables such that

• for
$$i \neq j$$
, $cov(\xi_i, \xi_j) = 0$,
• there exists $C > 0$ such that, for every i ,

$$\operatorname{Var} \xi_i < C$$
.

Then, for all $\varepsilon > 0, \delta > 0$

nen, for all
$$arepsilon>0,\delta>0$$

$$\mathsf{P}(frac{|S_n-\mathsf{E}S_n|}{n^{1/2+\delta}}>arepsilon)\to 0.$$