Probability theory

Lecture 6: Variance and covariance

Maksim Zhukovskii

MIPT

Definitions

- Variance of ξ is $\mathrm{Var}\xi = \mathsf{E}(\xi - \mathsf{E}\xi)^2$.

Definitions

- Variance of ξ is $\mathrm{Var}\xi = \mathsf{E}(\xi - \mathsf{E}\xi)^2$.
- Covariance of ξ, η is $\mathrm{cov}(\xi, \eta) = \mathsf{E}((\xi - \mathsf{E}\xi)(\eta - \mathsf{E}\eta))$.

Definitions

- Variance of ξ is $\mathrm{Var}\xi = \mathsf{E}(\xi - \mathsf{E}\xi)^2$.
- Covariance of ξ, η is $\mathrm{cov}(\xi, \eta) = \mathsf{E}((\xi - \mathsf{E}\xi)(\eta - \mathsf{E}\eta))$.

Clearly, $\mathrm{Var}\xi = \mathrm{cov}(\xi, \xi)$.

Properties of variance

1. $\mathrm{Var}\xi = \mathsf{E}\xi^2 - (\mathsf{E}\xi)^2.$

Properties of variance

1. $\mathrm{Var}\xi = \mathsf{E}\xi^2 - (\mathsf{E}\xi)^2.$

Proof.

Properties of variance

1. $\mathrm{Var}\xi = \mathsf{E}\xi^2 - (\mathsf{E}\xi)^2.$

Proof. $\mathrm{Var}\xi = \mathsf{E}(\xi - \mathsf{E}\xi)^2 =$

Properties of variance

1. $\mathrm{Var}\xi = \mathsf{E}\xi^2 - (\mathsf{E}\xi)^2$.

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

Properties of variance

1. $\mathrm{Var}\xi = \mathsf{E}\xi^2 - (\mathsf{E}\xi)^2.$

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

Properties of variance

1. $\mathrm{Var}\xi = \mathsf{E}\xi^2 - (\mathsf{E}\xi)^2$.

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

Properties of variance

1. $\mathrm{Var}\xi = \mathsf{E}\xi^2 - (\mathsf{E}\xi)^2$.

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

2. $\mathrm{Var}\xi \geq 0$.

Properties of variance

1. $\mathrm{Var}\xi = \mathsf{E}\xi^2 - (\mathsf{E}\xi)^2$.

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

2. $\mathrm{Var}\xi \geq 0$.

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

Properties of covariance

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

Properties of covariance

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

Proof.

Properties of covariance

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

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

Properties of covariance

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

Proof. $\operatorname{cov}(\xi, \eta) = \mathsf{E}(\xi - \mathsf{E}\xi)(\eta - \mathsf{E}\eta)$
$= \mathsf{E}[\xi\eta - \xi\mathsf{E}\eta - \eta\mathsf{E}\xi + \mathsf{E}\xi\mathsf{E}\eta] =$

Properties of covariance

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

Proof. $\mathrm{cov}(\xi,\eta) = \mathsf{E}(\xi - \mathsf{E}\xi)(\eta - \mathsf{E}\eta)$
$= \mathsf{E}[\xi\eta - \xi\mathsf{E}\eta - \eta\mathsf{E}\xi + \mathsf{E}\xi\mathsf{E}\eta] =$
$\mathsf{E}\xi\eta - 2\mathsf{E}\xi\mathsf{E}\eta + \mathsf{E}\xi\mathsf{E}\eta =$

Properties of covariance

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

   Proof. $\mathrm{cov}(\xi,\eta) = \mathsf{E}(\xi - \mathsf{E}\xi)(\eta - \mathsf{E}\eta)$
   $= \mathsf{E}[\xi\eta - \xi\mathsf{E}\eta - \eta\mathsf{E}\xi + \mathsf{E}\xi\mathsf{E}\eta] =$
   $\mathsf{E}\xi\eta - 2\mathsf{E}\xi\mathsf{E}\eta + \mathsf{E}\xi\mathsf{E}\eta = \mathsf{E}\xi\eta - \mathsf{E}\xi\mathsf{E}\eta$.

Properties of covariance

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

   Proof. $\mathrm{cov}(\xi,\eta) = \mathsf{E}(\xi - \mathsf{E}\xi)(\eta - \mathsf{E}\eta)$
   $= \mathsf{E}[\xi\eta - \xi\mathsf{E}\eta - \eta\mathsf{E}\xi + \mathsf{E}\xi\mathsf{E}\eta] =$
   $\mathsf{E}\xi\eta - 2\mathsf{E}\xi\mathsf{E}\eta + \mathsf{E}\xi\mathsf{E}\eta = \mathsf{E}\xi\eta - \mathsf{E}\xi\mathsf{E}\eta$.

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

Properties of covariance

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

   Proof. $\mathrm{cov}(\xi, \eta) = \mathsf{E}(\xi - \mathsf{E}\xi)(\eta - \mathsf{E}\eta)$
   $= \mathsf{E}[\xi\eta - \xi\mathsf{E}\eta - \eta\mathsf{E}\xi + \mathsf{E}\xi\mathsf{E}\eta] =$
   $\mathsf{E}\xi\eta - 2\mathsf{E}\xi\mathsf{E}\eta + \mathsf{E}\xi\mathsf{E}\eta = \mathsf{E}\xi\eta - \mathsf{E}\xi\mathsf{E}\eta$.

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

   Since $\mathsf{E}\xi\eta = \mathsf{E}\xi\mathsf{E}\eta$.

Properties of covariance

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

Properties of covariance

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

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

Properties of covariance

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

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

Proof.

Properties of covariance

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

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

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

Properties of covariance

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

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

   Proof. $\mathrm{cov}(a_1\xi_1 + a_2\xi_2, \eta) =$
   $\mathsf{E}(a_1\xi_1\eta + a_2\xi_2\eta) - \mathsf{E}(a_1\xi_1 + a_2\xi_2)\mathsf{E}\eta =$

Properties of covariance

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

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

Proof. $\mathrm{cov}(a_1\xi_1 + a_2\xi_2, \eta) = \mathsf{E}(a_1\xi_1\eta + a_2\xi_2\eta) - \mathsf{E}(a_1\xi_1 + a_2\xi_2)\mathsf{E}\eta = a_1\mathsf{E}\xi_1\eta + a_2\mathsf{E}\xi_2\eta - a_1\mathsf{E}\xi_1\mathsf{E}\eta - a_2\mathsf{E}\xi_2\mathsf{E}\eta$

Properties of covariance

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

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

Proof. $\mathrm{cov}(a_1\xi_1 + a_2\xi_2, \eta) =$
$\mathsf{E}(a_1\xi_1\eta + a_2\xi_2\eta) - \mathsf{E}(a_1\xi_1 + a_2\xi_2)\mathsf{E}\eta =$
$a_1\mathsf{E}\xi_1\eta + a_2\mathsf{E}\xi_2\eta - a_1\mathsf{E}\xi_1\mathsf{E}\eta - a_2\mathsf{E}\xi_2\mathsf{E}\eta$
$= a_1\mathrm{cov}(\xi_1, \eta) + a_2\mathrm{cov}(\xi_2, \eta)$.

Variance of a sum of random variables

Claim

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

Variance of a sum of random variables

Claim

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

Proof.

Variance of a sum of random variables

Claim

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

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

Variance of a sum of random variables

Claim

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

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

Variance of a sum of random variables

Claim

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

Proof. $\mathrm{Var}(\xi_1 + \ldots + \xi_n) =$
$\mathrm{cov}(\xi_1 + \ldots + \xi_n, \xi_1 + \ldots + \xi_n) =$
$\sum_{i=1}^{n} \mathrm{cov}(\xi_i, \xi_i) + \sum_{i \neq j} \mathrm{cov}(\xi_i, \xi_j) =$

Variance of a sum of random variables

Claim

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

Proof. $\mathrm{Var}(\xi_1 + \ldots + \xi_n) =$
$\mathrm{cov}(\xi_1 + \ldots + \xi_n, \xi_1 + \ldots + \xi_n) =$
$\sum_{i=1}^{n} \mathrm{cov}(\xi_i, \xi_i) + \sum_{i \neq j} \mathrm{cov}(\xi_i, \xi_j) =$
$\mathrm{Var}\xi_1 + \ldots + \mathrm{Var}\xi_n + \sum_{i \neq j} \mathrm{cov}(\xi_i, \xi_j)$.

Variance of a sum of random variables

Claim

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

Proof. $\mathrm{Var}(\xi_1 + \ldots + \xi_n) =$
$\mathrm{cov}(\xi_1 + \ldots + \xi_n, \xi_1 + \ldots + \xi_n) =$
$\sum_{i=1}^n \mathrm{cov}(\xi_i, \xi_i) + \sum_{i \neq j} \mathrm{cov}(\xi_i, \xi_j) =$
$\mathrm{Var}\xi_1 + ... + \mathrm{Var}\xi_n + \sum_{i \neq j} \mathrm{cov}(\xi_i, \xi_j)$.

If $\xi_1, \ldots \xi_n$ are pairwise independent, then
$\mathrm{Var}(\xi_1 + \ldots + \xi_n) = \mathrm{Var}(\xi_1) + \ldots + \mathrm{Var}(\xi_n)$.

Variance of a binomial random variable

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

Variance of a binomial random variable

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

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

Variance of a binomial random variable

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

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

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

Variance of a binomial random variable

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

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

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

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

Variance of a binomial random variable

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

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

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

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

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

Variance of a binomial random variable

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

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

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

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

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

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

Variance of a binomial random variable

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

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

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

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

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

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

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

Variance of a binomial random variable

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

Variance of a binomial random variable

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

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

Variance of a binomial random variable

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

Let $\xi_1, ..., \xi_n$ are independent Bernoulli random variables.
Then $\xi \stackrel{d}{=} \xi_1 + \ldots + \xi_n$.

Variance of a binomial random variable

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

Let $\xi_1, ..., \xi_n$ are independent Bernoulli random variables.
Then $\xi \overset{d}{=} \xi_1 + \ldots + \xi_n$.

$$\mathrm{Var}\xi = n\mathrm{Var}\xi_1 = n(\mathsf{E}\xi_1^2 - (\mathsf{E}\xi_1)^2) = np(1-p).$$

Other examples

- $\xi \sim U\{1, \ldots, n\}$

Other examples

- $\xi \sim U\{1, \ldots, n\}$

  $\mathrm{Var}\xi = \frac{1}{n}\sum_{i=1}^{n} i^2 - \left(\frac{n+1}{2}\right)^2 =$

Other examples

- $\xi \sim U\{1, \ldots, n\}$

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

Other examples

- $\xi \sim U\{1, \dots, n\}$

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

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

Other examples

- $\xi \sim U\{1, \dots, n\}$

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

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

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

Other examples

- $\xi \sim U\{1, \ldots, n\}$

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

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

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

Other examples

- $\xi \sim U\{1, \ldots, n\}$

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

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

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

Other examples

- $\xi \sim U\{1, \ldots, n\}$

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

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

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

Cauchy–Bunyakovsky–Schwarz inequality

Theorem

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

Cauchy–Bunyakovsky–Schwarz inequality

Theorem

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

Then $\mathsf{E}|\xi\eta| < \infty$ *and* $(\mathsf{E}|\xi\eta|)^2 \leq \mathsf{E}\xi^2\mathsf{E}\eta^2$.

The proof

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

The proof

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

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

The proof

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

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

The proof

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

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

The proof

- 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 $\mathsf{E}\xi^2 > 0$, $\mathsf{E}\eta^2 > 0$.

The proof

- 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 $\mathsf{E}\xi^2 > 0$, $\mathsf{E}\eta^2 > 0$.

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

The proof

- 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 $\mathsf{E}\xi^2 > 0$, $\mathsf{E}\eta^2 > 0$.

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

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

The proof

- 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 $\mathsf{E}\xi^2 > 0$, $\mathsf{E}\eta^2 > 0$.

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

  $2|\xi_0\eta_0| \leq \xi_0^2 + \eta_0^2 \Rightarrow$
  $\mathsf{E}(2|\xi_0\eta_0|) \leq \mathsf{E}(\xi_0^2 + \eta_0^2) = 2$

The proof

- 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 $\mathsf{E}\xi^2 > 0$, $\mathsf{E}\eta^2 > 0$.

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

  $2|\xi_0\eta_0| \le \xi_0^2 + \eta_0^2 \Rightarrow$
  $\mathsf{E}(2|\xi_0\eta_0|) \le \mathsf{E}(\xi_0^2 + \eta_0^2) = 2 \Rightarrow$
  $\mathsf{E}|\xi_0\eta_0| \le 1$

The proof

- Let $E\xi^2 = 0$.

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

- Let $E\xi^2 > 0$, $E\eta^2 > 0$.

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

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

Markov inequality

Theorem

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

Markov inequality

Theorem

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

Then, for every $a > 0$, $\mathsf{P}(|\xi| \geq a) \leq \frac{\mathsf{E}|\xi|}{a}$.

The proof

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

The proof

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

The proof

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

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

The proof

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

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

The proof

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

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

Therefore, $\mathrm{P}(|\xi \geq a|) \leq \frac{\mathrm{E}|\xi|}{a}$.

Chebyshev inequality

Theorem

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

Chebyshev inequality

Theorem

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

Then, for every $a > 0$,

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

Chebyshev inequality

Theorem

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

Then, for every $a > 0$*,*

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

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

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

The proof

For every x_0, there exists $\lambda(x_0)$ such that $g(x) \geq g(x_0) + \lambda(x_0)(x - x_0)$ for all $x \in \mathbb{R}$.

The proof

For every x_0, there exists $\lambda(x_0)$ such that $g(x) \geq g(x_0) + \lambda(x_0)(x - x_0)$ for all $x \in \mathbb{R}$.

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

The proof

For every x_0, there exists $\lambda(x_0)$ such that $g(x) \geq g(x_0) + \lambda(x_0)(x - x_0)$ for all $x \in \mathbb{R}$.

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

Then $g(\xi) \geq g(\mathsf{E}\xi) + \lambda(\mathsf{E}\xi)(\xi - \mathsf{E}\xi)$.

The proof

For every x_0, there exists $\lambda(x_0)$ such that $g(x) \geq g(x_0) + \lambda(x_0)(x - x_0)$ for all $x \in \mathbb{R}$.

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

Then $g(\xi) \geq g(\mathrm{E}\xi) + \lambda(\mathrm{E}\xi)(\xi - \mathrm{E}\xi)$.

Therefore, $\mathrm{E}g(\xi) \geq g(\mathrm{E}\xi)$.

Law of large numbers

Theorem

Let $\xi_1, \xi_2, \ldots$ *be pairwise independent identically distributed random variables.*
Let $\mathsf{E}\xi_1^2 < \infty$*.*

Law of large numbers

Theorem

Let $\xi_1, \xi_2, \dots$ be pairwise independent identically distributed random variables. Let $\mathsf{E}\xi_1^2 < \infty$.

Then, for all $\varepsilon > 0$

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

The proof

$$\mathrm{P}\left(\frac{|S_n - \mathrm{E}S_n|}{n} > \varepsilon\right) \leq \frac{\mathrm{Var}\, S_n}{n^2 \varepsilon^2} =$$

The proof

$$\mathsf{P}\left(\frac{|S_n - \mathsf{E}S_n|}{n} > \varepsilon\right) \leq \frac{\mathrm{Var}\, S_n}{n^2\varepsilon^2} =$$

$$\frac{\mathrm{Var}\xi_1}{\varepsilon^2 n}$$

The proof

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

$$
\frac{\mathrm{Var}\xi_1}{\varepsilon^2 n} \to 0, \quad n \to \infty.
$$

A generalization

Theorem

Let $\xi_1, \xi_2, \ldots$ *be a sequence of random variables such that*

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

A generalization

Theorem

Let $\xi_1, \xi_2, \ldots$ *be a sequence of random variables such that*

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

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

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