MATH E-156 Mathematical Statistics

Harvard Extension School

Dmitry Kurochkin

Fall 2020
Lecture 6

Contents

Contents

Beta

$X \sim$ Beta(a, b), where $a, b > 0$, if X is a continuous random variable with the following probability density function (pdf):

$$f_X(x) = \begin{cases} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}, & \text{for all } x \in [0,1], \\ 0, & \text{otherwise.} \end{cases}$$

Here,

$$\Gamma(x) \doteq \int_0^{+\infty} u^{x-1} e^{-u} du \ \text{ for } \ x > 0.$$

<u>Claim</u>:
If $X \sim$ Beta(a, b) with some $a, b > 0$ then

- Expected value
$$\mathrm{E}[X] = \frac{a}{a+b}.$$
- Variance
$$\mathrm{Var}[X] = \frac{ab}{(a+b)^2(a+b+1)}.$$

Contents

Chauchy

$X \sim$ Chauchy(θ, σ), where $\theta \in \mathbb{R}$, $\sigma > 0$, if X is a continuous random variable with the following probability density function (pdf):

$$f_X(x) = \frac{1}{\pi\sigma} \frac{1}{1 + \left(\frac{x-\theta}{\sigma}\right)^2}, \quad \text{for all } x \in \mathbb{R},$$

Claim:
If $X \sim$ Chauchy(θ, σ) with some $\theta \in \mathbb{R}$, $\sigma > 0$ then

$$E[X] \text{ and } \mathrm{Var}(X) \text{ do not exist.}$$

Contents

Chebyshev's Inequality

Thm. (Chebyshev's Inequality)
Let X be a random variable with mean $\mathrm{E}[X] = \mu$ and variance $\mathrm{Var}(X) = \sigma^2$. Then, for any $t > 0$,

$$P(|X - \mu| > t) \leq \frac{\sigma^2}{t^2}.$$

Chebyshev's Inequality

Thm. (Chebyshev's Inequality)
Let X be a random variable with mean $\mathrm{E}[X] = \mu$ and variance $\mathrm{Var}(X) = \sigma^2$. Then, for any $t > 0$,

$$P(|X - \mu| > t) \leq \frac{\sigma^2}{t^2}.$$

Note:
For $k > 0$, let $t = k\sigma$ then

$$P(|X - \mu| > k\sigma) \leq \frac{1}{k}.$$

Corollary:
If $\mathrm{Var}(X) = 0$, then $P(X = \mu) = 1$.

Contents

Bias and Mean Squared Error

<u>Def.</u>

Let X be a measurement (random variable) of some true value x_0. Then

- *Bias* is defined as

$$\text{Bias} = \mathrm{E}[X - x_0].$$

- *Mean Squared Error* (MSE) is defined as

$$\text{MSE} = \mathrm{E}\left[(X - x_0)^2\right].$$

Bias and Mean Squared Error

Def.

Let X be a measurement (random variable) of some true value x_0. Then

- *Bias* is defined as

$$\text{Bias} = \mathrm{E}[X - x_0].$$

- *Mean Squared Error* (MSE) is defined as

$$\text{MSE} = \mathrm{E}\left[(X - x_0)^2\right].$$

Thm.

Let X be a measurement of some value x_0. Then

$$\text{MSE} = \text{Bias}^2 + \mathrm{Var}(X).$$

Contents

Covariance and Correlation

Def.

Let X and Y be two random variables. Then, provided the expectations exist,

- *Covariance* of X and Y is defined as

$$\mathrm{Cov}(X) = \mathrm{E}\left[(X - \mathrm{E}[X])\,(Y - \mathrm{E}[Y])\right].$$

- *Correlation* of X and Y is defined as

$$\rho_{X,Y} = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X) Var(Y)}},$$

provided covariance exists and $\mathrm{Var}(X), Var(Y) \neq 0$.

Contents

Properties of Covariance

Let X, Y, Z, and W be random variables and $a, b, c \in \mathbb{R}$ be some constants. Assuming covariances exist,

1. $\mathrm{Cov}(X, Y) = \mathrm{E}[XY] - E[X]E[Y]$

Properties of Covariance

Let X, Y, Z, and W be random variables and $a, b, c \in \mathbb{R}$ be some constants. Assuming covariances exist,

1. $\mathrm{Cov}(X, Y) = \mathrm{E}[XY] - E[X]E[Y]$
2. $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$

Properties of Covariance

Let X, Y, Z, and W be random variables and $a, b, c \in \mathbb{R}$ be some constants. Assuming covariances exist,

1. $\mathrm{Cov}(X, Y) = \mathrm{E}[XY] - E[X]E[Y]$
2. $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$
3. $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$

Properties of Covariance

Let X, Y, Z, and W be random variables and $a, b, c \in \mathbb{R}$ be some constants. Assuming covariances exist,

1. $\mathrm{Cov}(X, Y) = \mathrm{E}[XY] - E[X]E[Y]$
2. $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$
3. $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$
4. $\mathrm{Cov}(X, a) = 0$

Properties of Covariance

Let X, Y, Z, and W be random variables and $a, b, c \in \mathbb{R}$ be some constants. Assuming covariances exist,

1. $\text{Cov}(X, Y) = \text{E}[XY] - E[X]E[Y]$
2. $\text{Cov}(X, X) = \text{Var}(X)$
3. $\text{Cov}(X, Y) = \text{Cov}(Y, X)$
4. $\text{Cov}(X, a) = 0$
5. $\text{Cov}(X, Y) = 0$ if X and Y are independent

Properties of Covariance

Let X, Y, Z, and W be random variables and $a, b, c \in \mathbb{R}$ be some constants. Assuming covariances exist,

1. $\mathrm{Cov}(X, Y) = \mathrm{E}[XY] - E[X]E[Y]$
2. $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$
3. $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$
4. $\mathrm{Cov}(X, a) = 0$
5. $\mathrm{Cov}(X, Y) = 0$ if X and Y are independent
6. $\mathrm{Cov}(aX, bY) = ab\,\mathrm{Cov}(X, Y)$

Properties of Covariance

Let X, Y, Z, and W be random variables and $a, b, c \in \mathbb{R}$ be some constants. Assuming covariances exist,

1. $\mathrm{Cov}(X, Y) = \mathrm{E}[XY] - E[X]E[Y]$
2. $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$
3. $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$
4. $\mathrm{Cov}(X, a) = 0$
5. $\mathrm{Cov}(X, Y) = 0$ if X and Y are independent
6. $\mathrm{Cov}(aX, bY) = ab\,\mathrm{Cov}(X, Y)$
7. $\mathrm{Cov}(X, Y + Z) = \mathrm{Cov}(X, Y) + \mathrm{Cov}(X, Z)$

Properties of Covariance

Let X, Y, Z, and W be random variables and $a, b, c \in \mathbb{R}$ be some constants. Assuming covariances exist,

1. $\mathrm{Cov}(X, Y) = \mathrm{E}[XY] - E[X]E[Y]$
2. $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$
3. $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$
4. $\mathrm{Cov}(X, a) = 0$
5. $\mathrm{Cov}(X, Y) = 0$ if X and Y are independent
6. $\mathrm{Cov}(aX, bY) = ab\,\mathrm{Cov}(X, Y)$
7. $\mathrm{Cov}(X, Y + Z) = \mathrm{Cov}(X, Y) + \mathrm{Cov}(X, Z)$
8. $\mathrm{Cov}(aX + bY, cZ + dW) =$ $ac\,\mathrm{Cov}(X, Z) + ad\,\mathrm{Cov}(X, W) + bc\,\mathrm{Cov}(Y, Z) + bd\,\mathrm{Cov}(Y, W)$

Properties of Covariance

THEOREM **A**

Suppose that $U = a + \sum_{i=1}^{n} b_i X_i$ and $V = c + \sum_{j=1}^{m} d_j Y_j$. Then

$$\mathrm{Cov}(U, V) = \sum_{i=1}^{n} \sum_{j=1}^{m} b_i d_j \mathrm{Cov}(X_i, Y_j) \qquad \blacksquare$$

COROLLARY **A**

$\mathrm{Var}(a + \sum_{i=1}^{n} b_i X_i) = \sum_{i=1}^{n} \sum_{j=1}^{n} b_i b_j \mathrm{Cov}(X_i, X_j)$. $\blacksquare$

COROLLARY **B**

$\mathrm{Var}(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} \mathrm{Var}(X_i)$, if the X_i are independent. $\blacksquare$

Source: *Mathematical Statistics and Data Analysis* by John Rice

Properties of Covariance

Thm.

Let X and Y be two random variables. Then

$$-1 \leq \rho_{X,Y} \leq 1,$$

provided the correlation exists. Furthermore,

$$\rho_{X,Y} = \pm 1 \ \text{ if and only if } \ P(Y = a + bX) = 1$$

for some constants a and b.

Contents

Covariance of Discrete Random Variables

Example:

Let X and Y be two discrete random variables with the following joint probability mass function (joint pmf):

	$X=1$	$X=2$	$X=3$
$Y=1$	.10	.08	.02
$Y=2$	.04	.22	.04
$Y=3$	.02	.18	.30

Then $E[X]=2.2$, $E[Y]=2.3$, $E[XY]=5.36$, $\mathrm{Var}(X)=0.48$, and $\mathrm{Var}(Y)=0.61$ and hence

- $\mathrm{Cov}(X,Y)=E[XY]-E[X]E[Y]=5.36-2.2\cdot 2.3=0.3$
- $\rho_{X,Y}=\frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}=\frac{0.3}{\sqrt{0.48\cdot 0.61}}=0.554416$

Covariance of Continuous Random Variables

Example:

Let X and Y be two continuous random variables that are jointly Normally distributed, that is,

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2} + \frac{(x-\mu_Y)^2}{\sigma_Y^2} - \frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}\right]},$$

where $x \in \mathbb{R}$.

Then $E[X] = \mu_X$, $E[Y] = \mu_Y$, $\mathrm{Var}(X) = \sigma_X^2$, $\mathrm{Var}(Y) = \sigma_Y^2$, and

- $\mathrm{Cov}(X,Y) = \mathrm{E}\left[(X - \mathrm{E}[X])\,(Y - \mathrm{E}[Y])\right]$
  $= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (x-\mu_X)\,(y-\mu_Y)\, f_{X,Y}(x,y)dxdy = \rho\,\sigma_X\sigma_Y$
- $\rho_{X,Y} = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}} = \frac{\rho\,\sigma_X\sigma_Y}{\sqrt{\sigma_X^2\sigma_Y^2}} = \rho$

Contents

Conditional Expectation

Def.

Let X and Y be two random variables. *Conditional expectation* of X given $Y = y$ is defined as follows:

- if X and Y are discrete random variables,

$$E[X|Y = y] \doteq \sum_{x} x p_{X|Y}(x|y)$$

- if X and Y are continuous random variables,

$$E[X|Y = y] \doteq \int_{-\infty}^{+\infty} x f_{X|Y}(x|y) dx$$

Conditional Expectation

Def.
Let X and Y be two random variables. *Conditional expectation* of X given $Y = y$ is defined as follows:

- if X and Y are discrete random variables,

$$E[X|Y=y] \doteq \sum_x x p_{X|Y}(x|y)$$

- if X and Y are continuous random variables,

$$E[X|Y=y] \doteq \int_{-\infty}^{+\infty} x f_{X|Y}(x|y) dx$$

Note:
Using conditional expectation $\mathrm{E}[\cdot|Y=y]$, one can now define *conditional variance*:

$$\begin{aligned}\mathrm{Var}(X|Y=y) &\doteq \mathrm{E}\left[(X - E[X|Y=y])^2 \,|Y=y\right] \\ &= \mathrm{E}[X^2|Y=y] - (E[X|Y=y])^2 .\end{aligned}$$

Contents

Properties of Conditional Expectation

Thm.
Let X and Y be two random variables. Assuming expectations exist, then

$$E[X] = E\left[E[X|Y]\right].$$

Properties of Conditional Expectation

Thm.
Let X and Y be two random variables. Assuming expectations and variances exist, then

$$Var(X) = Var\left(E[X|Y]\right) + E\left[Var(X|Y)\right].$$

Contents

Prediction

Thm.

Let X and Y be two random variables. Then $h(X)$ that minimizes

$$\mathrm{E}\left[(Y - h(X))^2\right]$$

is given by

$$h(X) = \mathrm{E}[Y|X],$$

provided the expectations exist.