Probability and Statistics

UG2, Core course, IIIT,H

Pawan Kumar

IIIT, Hyderabad

November 23, 2021

Outline

Definition of Covariance...

Definition of Covariance...

Definition of Covariance

Let X and Y be two random variables.

Definition of Covariance...

Definition of Covariance

Let X and Y be two random variables. The covariance between X and Y is defined as

Definition of Covariance...

Definition of Covariance

Let X and Y be two random variables. The covariance between X and Y is defined as

Remarks

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

Derivation

$$E\left[XY - X\underbrace{E[Y]} - \underbrace{E[X]}Y + \underbrace{E[X]E[Y]}\right]$$

Using Lin.

$$= E[XY] - E[Y]E[X] - E[X]\cancel{E[Y]} + E[X]\cancel{E[Y]}$$

$$= E[XY] - E[X]E[Y]$$

Solved Example

Solved Example

Recall $X \sim U(a,b) \Rightarrow E[X] = \frac{a+b}{2}$

Example (Solved Example)

Suppose $X \sim \text{Uniform}(1,2)$ and given $X = x$, Y is exponential with parameter $\lambda = x$. Find $\text{Cov}(X,Y)$.

$\frac{1}{b-a} = \frac{1}{2-1} = 1$

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

$X \sim \text{Uniform}(1,2) \Rightarrow E[X] = 3/2$

$E[Y] = E[\underbrace{E[Y|X]}]$ (by L.O.T.E)

$= E[1/X] = \int_1^2 \frac{1}{x} \cdot 1 \, dx = [\ln x]_1^2 = \ln 2 - \ln 1 = \ln 2.$

$E[XY] = E[E[XY|X]]$ [L.O.T.E]

$= E[X\,E[Y|X]] = E[X \cdot \frac{1}{X}] = E[1] = 1$

$\boxed{\text{Cov}(X,Y) = 1 - 3/2 \ln 2}$

Properties of Covariance...

Properties of Covariance...

Properties of Covariance

1. $\text{Cov}(X, X) = \text{Var}(X)$

Properties of Covariance...

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

$= E[X]E[Y]$ (if X, Y ind).

Properties of Covariance

1. $\text{Cov}(X, X) = \text{Var}(X)$
2. If X and Y are independent, then $\text{Cov}(X, Y) = 0$. $\boxed{\text{[Note: converse is not true!]}}$

(Look for example)

Properties of Covariance...

Properties of Covariance

1. $\text{Cov}(X, X) = \text{Var}(X)$
2. If X and Y are independent, then $\text{Cov}(X, Y) = 0$. [Note: converse is not true!]
3. $\text{Cov}(X, Y) = \text{Cov}(Y, X)$

Properties of Covariance...

Properties of Covariance

1. $\text{Cov}(X, X) = \text{Var}(X)$ Ex
2. If X and Y are independent, then $\text{Cov}(X, Y) = 0$. [Note: converse is not true!]
3. $\text{Cov}(X, Y) = \text{Cov}(Y, X)$ Ex
4. $\text{Cov}(a\,X, Y) = a\,\text{Cov}(X, Y)$ Ex

Properties of Covariance...

$\mathrm{Cov}(X, Y+c) = \mathrm{Cov}(X, Y)$

$\| \quad \mathrm{Cov}(Y+c, X) = \mathrm{Cov}(Y, X) = \mathrm{Cov}(X, Y)$

Properties of Covariance

1. $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$
2. If X and Y are independent, then $\mathrm{Cov}(X, Y) = 0$. [Note: converse is not true!]
3. $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$
4. $\mathrm{Cov}(aX, Y) = a\,\mathrm{Cov}(X, Y)$
5. $\mathrm{Cov}(X + c, Y) = \mathrm{Cov}(X, Y)$ Eg

Properties of Covariance...

Properties of Covariance

1. $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$
2. If X and Y are independent, then $\mathrm{Cov}(X, Y) = 0$. [Note: converse is not true!]
3. $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$
4. $\mathrm{Cov}(a\,X, Y) = a\,\mathrm{Cov}(X, Y)$
5. $\mathrm{Cov}(X + c, Y) = \mathrm{Cov}(X, Y)$
6. $\mathrm{Cov}(X + Y, Z) = \mathrm{Cov}(X, Z) + \mathrm{Cov}(Y, Z)$

Properties of Covariance...

Properties of Covariance

1. $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$
2. If X and Y are independent, then $\mathrm{Cov}(X, Y) = 0$. [Note: converse is not true!]
3. $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$
4. $\mathrm{Cov}(a\,X, Y) = a\,\mathrm{Cov}(X, Y)$
5. $\mathrm{Cov}(X + c, Y) = \mathrm{Cov}(X, Y)$
6. $\mathrm{Cov}(X + Y, Z) = \mathrm{Cov}(X, Z) + \mathrm{Cov}(Y, Z)$
7. More generally, we have

Properties of Covariance...

Properties of Covariance

1. $\text{Cov}(X, X) = \text{Var}(X)$
2. If X and Y are independent, then $\text{Cov}(X, Y) = 0$. [Note: converse is not true!]
3. $\text{Cov}(X, Y) = \text{Cov}(Y, X)$
4. $\text{Cov}(aX, Y) = a\,\text{Cov}(X, Y)$
5. $\text{Cov}(X + c, Y) = \text{Cov}(X, Y)$
6. $\text{Cov}(X + Y, Z) = \text{Cov}(X, Z) + \text{Cov}(Y, Z)$
7. More generally, we have

$$\text{Cov}\left(\sum_{i=1}^{m} a_i X_i, \sum_{j=1}^{n} b_j Y_j\right) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_i b_j \,\text{Cov}(X_i, Y_j)$$

Use 3, 4, 6 ✓

Solved Example...

Solved Example...

$Var(X) = 1 = E[X^2] - E[X]^2$ $X \sim N(0,1)$ $N(0, 1)$

Example (Solved example)

Let X and Y be two independent random variables following standard normal distribution, and

$$Z = 1 + X + XY^2$$

$$W = 1 + X$$

Find $\text{Cov}(Z, W)$.

$$\text{Cov}(Z, W) = \text{Cov}(1+X+XY^2, 1+X) = \text{Cov}(X+XY^2, X)$$

$$= \text{Cov}(X, X) + \text{Cov}(XY^2, X) = Var(X) + \text{Cov}(XY^2, X)$$

$X \sim N(0,1)$

$$= 1 + E[X^2Y^2] - E[XY^2]\,E[X]$$

$$= 1 + E[X^2]\,E[Y^2] - E[X]\,E[Y^2]\,E[X]$$

$$= 1 + 1 \cdot 1 - 0 = 2$$

Variance of a Sum...

Variance of a Sum...

Example (Variance of a Sum)

Let $Z = X + Y$, then

Variance of a Sum...

Example (Variance of a Sum)

Let $Z = X + Y$, then

$$\mathrm{Var}(Z) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X, Y)$$

Variance of a Sum...

Example (Variance of a Sum)

Let $Z = X + Y$, then

$$\mathrm{Var}(Z) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X, Y)$$

More generally, for $a, b \in \mathbb{R}$, and $Z = aX + bY$, we have

Variance of a Sum...

Example (Variance of a Sum)

Let $Z = X + Y$, then

$$\text{Var}(Z) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$$

More generally, for $a, b \in \mathbb{R}$, and $Z = aX + bY$, we have

$$\text{Var}(Z) = a^2\,\text{Var}(X) + b^2\,\text{Var}(Y) + 2ab\,\text{Cov}(X, Y) \leftarrow \text{Ex}$$

Solution

$$\text{Var}(Z) = \text{Cov}(Z, Z) = \text{Cov}(X+Y, X+Y)$$

$$= \text{Cov}(X, X) + \text{Cov}(X, Y) + \text{Cov}(Y, X) + \text{Cov}(Y, Y)$$

$$= \text{Var}(X) + \text{Var}(Y) + 2\,\text{Cov}(X, Y)$$

Correlation Coefficient...

Correlation Coefficient...

Correlation coefficient

The correlation coefficient ρ_{XY} or $\rho(X, Y)$ is defined as follows

Correlation Coefficient...

Correlation coefficient

The correlation coefficient ρ_{XY} or $\rho(X, Y)$ is defined as follows

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

Correlation Coefficient...

Correlation coefficient

The correlation coefficient ρ_{XY} or $\rho(X, Y)$ is defined as follows

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

- Given two RVs X and Y, define standardized versions of these as follows

Correlation Coefficient...

Correlation coefficient

The correlation coefficient ρ_{XY} or $\rho(X,Y)$ is defined as follows

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

- Given two RVs X and Y, define standardized versions of these as follows

$$U = \frac{X - E[X]}{\rho_X}, \quad V = \frac{Y - E[Y]}{\rho_Y}$$

Correlation Coefficient...

Correlation coefficient

The correlation coefficient ρ_{XY} or $\rho(X, Y)$ is defined as follows

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

- Given two RVs X and Y, define standardized versions of these as follows

$$U = \frac{X - E[X]}{\rho_X}, \quad V = \frac{Y - E[Y]}{\rho_Y}$$

- With this by computing the covariance $\mathrm{Cov}(U, V)$, we have

Correlation Coefficient...

Correlation coefficient

The correlation coefficient ρ_{XY} or $\rho(X, Y)$ is defined as follows

$$\rho_{XY} = \rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\,\text{Var}(Y)}} = \frac{\text{Cov}(X, Y)}{\rho_X\, \rho_Y}$$

- Given two RVs X and Y, define standardized versions of these as follows

$$U = \frac{X - E[X]}{\rho_X}, \quad V = \frac{Y - E[Y]}{\rho_Y}$$

- With this by computing the covariance $\text{Cov}(U, V)$, we have

$$\rho_{XY} = \text{Cov}(U, V) = \text{Cov}\left(\frac{X - E[X]}{\rho_X}, \frac{Y - E[Y]}{\rho_Y}\right)$$
$$= \text{Cov}\left(\frac{X}{\rho_X}, \frac{Y}{\rho_Y}\right) = \frac{\text{Cov}(X, Y)}{\rho_X \rho_Y}$$

Properties of Correlation Coefficient...

Properties of Correlation Coefficient...

Properties of correlation coefficient

Let X, Y be two RVs.

Properties of Correlation Coefficient...

Properties of correlation coefficient

Let X, Y be two RVs. These are some properties of correlation coefficient

Properties of Correlation Coefficient...

Properties of correlation coefficient

Let X, Y be two RVs. These are some properties of correlation coefficient

1. $-1 \leq \rho(X, Y) \leq 1$ $\leftarrow$ Try (C-S)

Properties of Correlation Coefficient...

Properties of correlation coefficient

Let X, Y be two RVs. These are some properties of correlation coefficient

1. $-1 \leq \rho(X, Y) \leq 1$
2. If $\rho(X, Y) = 1$,

Properties of Correlation Coefficient...

$Y = f(X) = X^2 + X^3$

Properties of correlation coefficient

Let X, Y be two RVs. These are some properties of correlation coefficient

1. $-1 \leq \rho(X, Y) \leq 1$
2. If $\rho(X, Y) = 1$, then $\boxed{Y = aX + b}$, where $a > 0$

Properties of Correlation Coefficient...

Properties of correlation coefficient

Let X, Y be two RVs. These are some properties of correlation coefficient

1. $-1 \leq \rho(X, Y) \leq 1$
2. If $\rho(X, Y) = 1$, then $Y = aX + b$, where $a > 0$
3. If $\rho(X, Y) = 1$,

Properties of Correlation Coefficient...

Properties of correlation coefficient

Let X, Y be two RVs. These are some properties of correlation coefficient

1. $-1 \leq \rho(X, Y) \leq 1$
2. If $\rho(X, Y) = 1$, then $Y = aX + b$, where $a > 0$
3. If $\rho(X, Y) = -1$, then $Y = aX + b$, where $a < 0$

Properties of Correlation Coefficient...

Properties of correlation coefficient

Let X, Y be two RVs. These are some properties of correlation coefficient

1. $-1 \leq \rho(X, Y) \leq 1$
2. If $\rho(X, Y) = 1$, then $Y = aX + b$, where $a > 0$
3. If $\rho(X, Y) = 1$, then $Y = aX + b$, where $a < 0$
4. $\rho(aX + b, cY + d) = \rho(X, Y)$ for $a, c > 0$

Cov( )

Answer to previous problem...

Positive Correlation, Negative Correlation, Uncorrelation...

Positive Correlation, Negative Correlation, Uncorrelation...

Definition of positive, negative correlation

Let X and Y be two RVs.

Positive Correlation, Negative Correlation, Uncorrelation...

Definition of positive, negative correlation

Let X and Y be two RVs.

1. If $\rho(X, Y) = 0$, we say that X and Y are uncorrelated

Positive Correlation, Negative Correlation, Uncorrelation...

Definition of positive, negative correlation

Let X and Y be two RVs.

1. If $\rho(X, Y) = 0$, we say that X and Y are uncorrelated
2. If $\rho(X, Y) > 0$, we say that X and Y are positively correlated

Positive Correlation, Negative Correlation, Uncorrelation...

Definition of positive, negative correlation

Let X and Y be two RVs.

1. If $\rho(X, Y) = 0$, we say that X and Y are uncorrelated
2. If $\rho(X, Y) > 0$, we say that X and Y are positively correlated
3. If $\rho(X, Y) < 0$, we say that X and Y are negatively correlated

Positive Correlation, Negative Correlation, Uncorrelation...

Definition of positive, negative correlation

Let X and Y be two RVs.

1. If $\rho(X, Y) = 0$, we say that X and Y are uncorrelated
2. If $\rho(X, Y) > 0$, we say that X and Y are positively correlated
3. If $\rho(X, Y) < 0$, we say that X and Y are negatively correlated

0

Pairwise uncorrelation and Variance

$2\,\mathrm{Cov}(X,Y)$

If X and Y are uncorrelated, then

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$$

More generally, if $X_1, X_2, \ldots, X_n$, are pairwise uncorrelated, i.e., $\rho(X_i, X_j) = 0$ when $i \neq j$, then

Positive Correlation, Negative Correlation, Uncorrelation...

Definition of positive, negative correlation

Let X and Y be two RVs.

1. If $\rho(X, Y) = 0$, we say that X and Y are uncorrelated
2. If $\rho(X, Y) > 0$, we say that X and Y are positively correlated
3. If $\rho(X, Y) < 0$, we say that X and Y are negatively correlated

Pairwise uncorrelation and Variance

If X and Y are uncorrelated, then

$$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$$

More generally, if $X_1, X_2, \ldots, X_n$, are pairwise uncorrelated, i.e., $\rho(X_i, X_j) = 0$ when $i \neq j$, then

$$\mathrm{Var}(X_1 + X_2 + \cdots + X_n) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \cdots + \mathrm{Var}(X_n) + \cancel{2\mathrm{Cov}(X_i, X_j)} \to 0$$

Solved Example

Example (Solved Example)

Consider a point (X, Y) chosen uniformly at random from the following disc

Solved Example

Example (Solved Example)

Consider a point (X, Y) chosen uniformly at random from the following disc

$$D = \{(x, y) \mid x^2 + y^2 \leq 1\}$$

Solved Example

Example (Solved Example)

Consider a point (X, Y) chosen uniformly at random from the following disc

$$D = \{(x, y) \mid x^2 + y^2 \leq 1\}$$

The joint PDF of X and Y is given by

Solved Example

Example (Solved Example)

Consider a point (X, Y) chosen uniformly at random from the following disc

$$D = \{(x, y) \mid x^2 + y^2 \leq 1\}$$

The joint PDF of X and Y is given by

$$f_{XY}(x, y) = \begin{cases} c & (x, y) \in D \\ 0 & \text{otherwise} \end{cases}$$

Solved Example

Example (Solved Example)

Consider a point (X, Y) chosen uniformly at random from the following disc

$$D = \{(x, y) \mid x^2 + y^2 \leq 1\}$$

The joint PDF of X and Y is given by

$$f_{XY}(x, y) = \begin{cases} c & (x, y) \in D \\ 0 & \text{otherwise} \end{cases}$$

Are X and Y uncorrelated?

Skipped

Answer to previous problem...

Bivariate Normal Distribution...

Bivariate Normal Distribution...

Example (Sum of Two Normal Distribution May Not be Normal)

Let RVs $X \sim N(0,1)$ and $W \sim \text{Bernoulli}\left(\frac{1}{2}\right)$ be two independent RVs.

Bivariate Normal Distribution...

Example (Sum of Two Normal Distribution May Not be Normal)

Let RVs $X \sim N(0,1)$ and $W \sim \text{Bernoulli}\left(\frac{1}{2}\right)$ be two independent RVs. Let random variable Y be a function of X and W as follows

Bivariate Normal Distribution...

Example (Sum of Two Normal Distribution May Not be Normal)

Let RVs $X \sim N(0,1)$ and $W \sim \text{Bernoulli}\left(\frac{1}{2}\right)$ be two independent RVs. Let random variable Y be a function of X and W as follows

$$Y = h(X, W) = \begin{cases} X & \text{if } W = 0 \\ -X & \text{if } W = 1 \end{cases}$$

Bivariate Normal Distribution...

Example (Sum of Two Normal Distribution May Not be Normal)

Let RVs $X \sim N(0,1)$ and $W \sim$ Bernoulli $\left(\frac{1}{2}\right)$ be two independent RVs. Let random variable Y be a function of X and W as follows

$$Y = h(X, W) = \begin{cases} X & \text{if } W = 0 \\ -X & \text{if } W = 1 \end{cases}$$

Find the PDF of Y

Answer to previous problem...

Bivariate Normal Distribution and Properties...

Bivariate Normal Distribution and Properties...

Definition (Bivariate Normal Distribution)

Two RVs X and Y are called bivariate normal,

Bivariate Normal Distribution and Properties...

Definition (Bivariate Normal Distribution)

Two RVs X and Y are called bivariate normal, or jointly normal,

Bivariate Normal Distribution and Properties...

Definition (Bivariate Normal Distribution)

Two RVs X and Y are called bivariate normal, or jointly normal, if $aX + bY$ has a normal distribution for all $a, b \in \mathbb{R}$.

Bivariate Normal Distribution and Properties...

Definition (Bivariate Normal Distribution)

Two RVs X and Y are called bivariate normal, or jointly normal, if $aX + bY$ has a normal distribution for all $a, b \in \mathbb{R}$.

1. If $a = b = 0$,

Bivariate Normal Distribution and Properties...

Definition (Bivariate Normal Distribution)

Two RVs X and Y are called bivariate normal, or jointly normal, if $aX + bY$ has a normal distribution for all $a, b \in \mathbb{R}$.

1. If $a = b = 0$, then $aX + bY = 0$ is a normal distribution with mean and variance 0

Bivariate Normal Distribution and Properties...

Definition (Bivariate Normal Distribution)

Two RVs X and Y are called bivariate normal, or jointly normal, if $aX + bY$ has a normal distribution for all $a, b \in \mathbb{R}$.

1. If $a = b = 0$, then $aX + bY = 0$ is a normal distribution with mean and variance 0
2. If X and Y are bivariate normal,

Bivariate Normal Distribution and Properties...

Definition (Bivariate Normal Distribution)

Two RVs X and Y are called bivariate normal, or jointly normal, if $aX + bY$ has a normal distribution for all $a, b \in \mathbb{R}$.

1. If $a = b = 0$, then $aX + bY = 0$ is a normal distribution with mean and variance 0
2. If X and Y are bivariate normal, then by letting $a = 1, b = 0$,

Bivariate Normal Distribution and Properties...

Definition (Bivariate Normal Distribution)

Two RVs X and Y are called bivariate normal, or jointly normal, if $\underbrace{aX + bY}$ has a normal distribution for all $a, b \in \mathbb{R}$.

1. If $a = b = 0$, then $aX + bY = 0$ is a normal distribution with mean and variance 0
2. If X and Y are bivariate normal, then by letting $a = 1, b = 0$, X must be normal

Bivariate Normal Distribution and Properties...

Definition (Bivariate Normal Distribution)

Two RVs X and Y are called bivariate normal, or jointly normal, if $aX + bY$ has a normal distribution for all $a, b \in \mathbb{R}$.

1. If $a = b = 0$, then $aX + bY = 0$ is a normal distribution with mean and variance 0
2. If X and Y are bivariate normal, then by letting $a = 1, b = 0$, X must be normal
3. If X and Y are bivariate normal,

Bivariate Normal Distribution and Properties...

Definition (Bivariate Normal Distribution)

Two RVs X and Y are called bivariate normal, or jointly normal, if $aX + bY$ has a normal distribution for all $a, b \in \mathbb{R}$.

1. If $a = b = 0$, then $aX + bY = 0$ is a normal distribution with mean and variance 0
2. If X and Y are bivariate normal, then by letting $a = 1, b = 0$, X must be normal
3. If X and Y are bivariate normal, then by letting $a = 0, b = 1$,

Bivariate Normal Distribution and Properties...

Definition (Bivariate Normal Distribution)

Two RVs X and Y are called bivariate normal, or jointly normal, if $aX + bY$ has a normal distribution for all $a, b \in \mathbb{R}$.

1. If $a = b = 0$, then $aX + bY = 0$ is a normal distribution with mean and variance 0
2. If X and Y are bivariate normal, then by letting $a = 1, b = 0$, X must be normal
3. If X and Y are bivariate normal, then by letting $a = 0, b = 1$, Y must be normal

Bivariate Normal Distribution and Properties...

Definition (Bivariate Normal Distribution)

Two RVs X and Y are called bivariate normal, or jointly normal, if $aX + bY$ has a normal distribution for all $a, b \in \mathbb{R}$.

1. If $a = b = 0$, then $aX + bY = 0$ is a normal distribution with mean and variance 0
2. If X and Y are bivariate normal, then by letting $a = 1, b = 0$, X must be normal
3. If X and Y are bivariate normal, then by letting $a = 0, b = 1$, Y must be normal
4. If $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$ are independent,

Bivariate Normal Distribution and Properties...

Definition (Bivariate Normal Distribution)

Two RVs X and Y are called bivariate normal, or jointly normal, if $aX + bY$ has a normal distribution for all $a, b \in \mathbb{R}$.

1. If $a = b = 0$, then $aX + bY = 0$ is a normal distribution with mean and variance 0
2. If X and Y are bivariate normal, then by letting $a = 1, b = 0$, X must be normal
3. If X and Y are bivariate normal, then by letting $a = 0, b = 1$, Y must be normal
4. If $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma^2)$ are independent, then they are jointly normal

Bivariate Normal Distribution and Properties...

Definition (Bivariate Normal Distribution)

Two RVs X and Y are called bivariate normal, or jointly normal, if $aX + bY$ has a normal distribution for all $a, b \in \mathbb{R}$.

1. If $a = b = 0$, then $aX + bY = 0$ is a normal distribution with mean and variance 0
2. If X and Y are bivariate normal, then by letting $a = 1, b = 0$, X must be normal
3. If X and Y are bivariate normal, then by letting $a = 0, b = 1$, Y must be normal
4. If $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma^2)$ are independent, then they are jointly normal
5. If $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$ are jointly normal,

Bivariate Normal Distribution and Properties...

Definition (Bivariate Normal Distribution)

Two RVs X and Y are called bivariate normal, or jointly normal, if $aX + bY$ has a normal distribution for all $a, b \in \mathbb{R}$.

1. If $a = b = 0$, then $aX + bY = 0$ is a normal distribution with mean and variance 0
2. If X and Y are bivariate normal, then by letting $a = 1, b = 0$, X must be normal
3. If X and Y are bivariate normal, then by letting $a = 0, b = 1$, Y must be normal
4. If $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma^2)$ are independent, then they are jointly normal
5. If $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$ are jointly normal, then

$$X + Y \sim N(\mu_X + \mu_Y, \quad \sigma_X^2 + \sigma_Y^2 + 2\rho(X, Y)\sigma_X\sigma_Y)$$

Bivariate Normal Distribution and Properties...

Definition (Bivariate Normal Distribution)

Two RVs X and Y are called bivariate normal, or jointly normal, if $aX + bY$ has a normal distribution for all $a, b \in \mathbb{R}$.

1. If $a = b = 0$, then $aX + bY = 0$ is a normal distribution with mean and variance 0
2. If X and Y are bivariate normal, then by letting $a = 1, b = 0$, X must be normal
3. If X and Y are bivariate normal, then by letting $a = 0, b = 1$, Y must be normal
4. If $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma^2)$ are independent, then they are jointly normal
5. If $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$ are jointly normal, then

$$X + Y \sim N(\mu_X + \mu_Y, \quad \sigma_X^2 + \sigma_Y^2 + 2\rho(X, Y)\sigma_X\sigma_Y)$$

6. Can we provide a simple way to generate jointly normal random variables?

Bivariate Normal Distribution and Properties...

Definition (Bivariate Normal Distribution)

Two RVs X and Y are called bivariate normal, or jointly normal, if $\underbrace{aX + bY}$ has a normal distribution for all $a, b \in \mathbb{R}$.

1. If $a = b = 0$, then $aX + bY = 0$ is a normal distribution with mean and variance 0
2. If X and Y are bivariate normal, then by letting $a = 1, b = 0$, X must be normal
3. If X and Y are bivariate normal, then by letting $a = 0, b = 1$, Y must be normal
4. If $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma^2)$ are independent, then they are jointly normal
5. If $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$ are jointly normal, then

$$X + Y \sim N(\mu_X + \mu_Y, \quad \sigma_X^2 + \sigma_Y^2 + 2\rho(X, Y)\sigma_X\sigma_Y)$$

6. Can we provide a simple way to generate jointly normal random variables?
7. We first introduce standard bivariate normal distribution

Example of Bivariate Normal...

Example of Bivariate Normal...

Example (Bivariate Normal Variable)

Let Z_1 and Z_2 be two independent $N(0, 1)$ RVs.

Example of Bivariate Normal...

Example (Bivariate Normal Variable)

Let Z_1 and Z_2 be two independent $N(0, 1)$ RVs. We define

Example of Bivariate Normal...

Example (Bivariate Normal Variable)

Let Z_1 and Z_2 be two independent $N(0, 1)$ RVs. We define

$$X = Z_1$$

$$Y = \rho Z_1 + \sqrt{1 - \rho^2} Z_2,$$

Example of Bivariate Normal...

Example (Bivariate Normal Variable)

Let Z_1 and Z_2 be two independent $N(0,1)$ RVs. We define

$$X = Z_1$$
$$Y = \rho Z_1 + \sqrt{1-\rho^2} Z_2,$$

where ρ is a real number in $(-1,1)$.

Example of Bivariate Normal...

Example (Bivariate Normal Variable)

Let Z_1 and Z_2 be two independent $N(0,1)$ RVs. We define

$$X = Z_1$$

$$Y = \rho Z_1 + \sqrt{1-\rho^2} Z_2,$$

where ρ is a real number in $(-1,1)$.

1 Is X and Y bivariate normal?

Given that z_1, z_2 are ind.

$$f_{Z_1 Z_1}(z_1, z_2) = \underline{f_{Z_1}}(z_1) f_{Z_2}(z_2)$$

$$= \frac{1}{\sqrt{2\pi}} e^{-z_1^2/2} \cdot \frac{1}{\sqrt{2\pi}} e^{-z_2^2/2} = \frac{1}{\sqrt{2\pi}} e^{-\frac{(z_1^2+z_2^2)}{2}}$$

Example of Bivariate Normal...

Example (Bivariate Normal Variable)

Let Z_1 and Z_2 be two independent $N(0,1)$ RVs. We define

$$X = Z_1$$
$$Y = \rho Z_1 + \sqrt{1-\rho^2} Z_2,$$

where ρ is a real number in $(-1,1)$.

1. Is X and Y bivariate normal?
2. What is the joint PDF of X and Y?

Example of Bivariate Normal...

Example (Bivariate Normal Variable)

Let Z_1 and Z_2 be two independent $N(0,1)$ RVs. We define

$$X = Z_1$$
$$Y = \rho Z_1 + \sqrt{1-\rho^2} Z_2,$$

where ρ is a real number in $(-1, 1)$.

1. Is X and Y bivariate normal?
2. What is the joint PDF of X and Y?
3. Find $\rho(X, Y)$

Answer to previous problem...

Claim : $ax+by$ is normal

$$ax+by = aZ_1 + b\left(\rho Z_1 + \sqrt{1-\rho^2}\, Z_2\right)$$

$$= (a+b\rho)Z_1 + b\sqrt{1-\rho^2}\, Z_2 .$$

is a linear combination of two standard normal $\Rightarrow$ $ax+by$ is normal.

$\Rightarrow$ $ax+by$ is bivariate normal

Recall

$aZ_1 + bZ_2$ is normal

$Z_1 \sim N(0,1)$

$Z_2 \sim N(0,1)$

Answer to previous problem...

Standard Bivariate Normal Distribution...

Standard Bivariate Normal Distribution...

Definition of Standard Bivariate Normal Distribution

Two RVs X and Y are said to have the standard bivariate normal distribution with correlation coefficient ρ if their joint PDF is given by

Standard Bivariate Normal Distribution...

Definition of Standard Bivariate Normal Distribution

Two RVs X and Y are said to have the standard bivariate normal distribution with correlation coefficient ρ if their joint PDF is given by

$$f_{XY}(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}[x^2 - 2\rho xy + y^2]\right\}$$

Standard Bivariate Normal Distribution...

$\frac{1}{2\pi} e^{-\frac{(x^2+y^2)}{2}}$ ← $\boxed{X+Y}$

Definition of Standard Bivariate Normal Distribution

Two RVs X and Y are said to have the standard bivariate normal distribution with correlation coefficient ρ if their joint PDF is given by

$$f_{XY}(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}[x^2 - 2\rho xy + y^2,]\right\}$$

where $\rho \in (-1,1)$. If $\rho = 0$, then we call X and Y to have standard normal bivariate normal distribution.

Definition of Bivariate Normal Distribution...

Definition of Bivariate Normal Distribution...

Std. Normal

Norm

Definition of Bivariate Normal Distribution

Two random variables X and Y are said to have a bivariate normal distribution with parameters $\mu_x, \sigma_X^2, \mu_Y, \sigma_Y^2$, and ρ,

Definition of Bivariate Normal Distribution...

Definition of Bivariate Normal Distribution

Two random variables X and Y are said to have a bivariate normal distribution with parameters $\mu_x, \sigma_X^2, \mu_Y, \sigma_Y^2$, and ρ, if their joint PDF is given by

$$f_{XY} = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \cdot$$

$$\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2 - 2\rho\frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}\right]\right\},$$

Definition of Bivariate Normal Distribution...

Definition of Bivariate Normal Distribution

Two random variables X and Y are said to have a bivariate normal distribution with parameters $\mu_x, \sigma_X^2, \mu_Y, \sigma_Y^2$, and ρ, if their joint PDF is given by

$$f_{XY} = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \cdot$$

$$\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2 - 2\rho\frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}\right]\right\},$$

where $\mu_X, \mu_Y \in \mathbb{R}, \sigma_X, \sigma_Y > 0$ and $\rho \in (-1, 1)$ are all constant.

Creating a Jointly Normal Distribution...

Creating a Jointly Normal Distribution...

Creation of Jointly Normal Random Variables

If we want two jointly normal random variables X and Y such that

Creating a Jointly Normal Distribution...

Creation of Jointly Normal Random Variables

If we want two jointly normal random variables X and Y such that

$$X \sim N(\mu_X, \sigma_X^2), \quad Y \sim N(\mu_Y, \sigma_Y^2), \quad \text{and } \rho(X, Y) = \rho$$

Creating a Jointly Normal Distribution...

Creation of Jointly Normal Random Variables

If we want two jointly normal random variables X and Y such that

$$X \sim N(\mu_X, \sigma_X^2), \quad Y \sim N(\mu_Y, \sigma_Y^2), \quad \text{and } \rho(X, Y) = \rho$$

we start with two independent standard normal RVs Z_1 and Z_2 and define

Creating a Jointly Normal Distribution...

Creation of Jointly Normal Random Variables

If we want two jointly normal random variables X and Y such that

$$X \sim N(\mu_X, \sigma_X^2), \quad Y \sim N(\mu_Y, \sigma_Y^2), \quad \text{and } \rho(X, Y) = \rho$$

we start with two independent standard normal RVs Z_1 and Z_2 and define

$$\begin{aligned} X &= \sigma_X Z_1 + \mu_X \\ Y &= \sigma_Y(\rho Z_1 + \sqrt{1-\rho^2} Z_2) + \mu_Y \end{aligned}$$

and follow the above procedure: solve for Z_1, Z_2, and apply method of transformation

Bivariate Normal and Uncorrelation implies Independence...

Bivariate Normal and Uncorrelation implies Independence...

Theorem

If X and Y are bivariate normal and uncorrelated, then they are independent.

Try!

Solved Problem 1

Example (Solved Problem 1)

Let X and Y be two jointly continuous random variables with joint PDF

Solved Problem 1

Example (Solved Problem 1)

Let X and Y be two jointly continuous random variables with joint PDF

$$f_{XY}(x,y) = \begin{cases} 2 & y + x \leq 1,\ x > 0,\ y > 0 \\ 0 & \text{otherwise} \end{cases}$$

Solved Problem 1

Example (Solved Problem 1)

Let X and Y be two jointly continuous random variables with joint PDF

$$f_{XY}(x,y) = \begin{cases} 2 & y + x \leq 1,\ x > 0,\ y > 0 \\ 0 & \text{otherwise} \end{cases}$$

Find $\text{Cov}(X,Y)$ and $\rho(X,Y)$.

Answer to previous problem...

Answer to previous problem...

Solved Problem 2...

Solved Problem 2...

Example (Solved Problem 2)

I roll a fair die n times.

Solved Problem 2...

Example (Solved Problem 2)

I roll a fair die n times. Let X be the number of 1's that I observe and let Y be the number of 2's that I observe.

Solved Problem 2...

Example (Solved Problem 2)

I roll a fair die n times. Let X be the number of 1's that I observe and let Y be the number of 2's that I observe. Find $\text{Cov}(X, Y)$ and $\rho(X, Y)$.

Answer to previous problem...

Outline

Union bound and extension...

Union bound and extension...

Union Bound

Recall the inclusion exclusion principle:

Union bound and extension...

Union Bound

Recall the inclusion exclusion principle:

$$P((\cup_{i=1}^{n} A_i)) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) - \cdots + (-1)^{n-1} P\left(\cap_{i=1}^{n} A_i\right)$$

Union bound and extension...

Union Bound

Recall the inclusion exclusion principle:

$$P((\cup_{i=1}^{n} A_i)) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j)$$
$$+ \sum_{i<j<k} P(A_i \cap A_j \cap A_k) - \cdots + (-1)^{n-1} P\left(\cap_{i=1}^{n} A_i\right)$$

- union bound states that probability of union of events is smaller than the sum of first term.

Union bound and extension...

Union Bound

Recall the inclusion exclusion principle:

$$P((\cup_{i=1}^{n} A_i)) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j)$$
$$+ \sum_{i<j<k} P(A_i \cap A_j \cap A_k) - \cdots + (-1)^{n-1} P\left(\cap_{i=1}^{n} A_i\right)$$

- union bound states that probability of union of events is smaller than the sum of first term. That is for $n = 2$, we have

Union bound and extension...

Union Bound

Recall the inclusion exclusion principle:

A_i, A_j

$$P((\cup_{i=1}^{n} A_i)) = \sum_{i=1}^{n} P(A_i) - \boxed{\sum_{i<j} P(A_i \cap A_j)}$$

$$+ \sum_{i<j<k} P(A_i \cap A_j \cap A_k) - \cdots + (-1)^{n-1} P\left(\cap_{i=1}^{n} A_i\right)$$

- union bound states that probability of union of events is smaller than the sum of first term. That is for $n = 2$, we have

$$P(A \cup B) \leq P(A) + P(B), \quad - \cancel{P(A \cap B)}$$

Union bound and extension...

Union Bound

Recall the inclusion exclusion principle:

$$P((\cup_{i=1}^{n} A_i)) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j)$$

$$+ \sum_{i<j<k} P(A_i \cap A_j \cap A_k) - \cdots + (-1)^{n-1} P\left(\cap_{i=1}^{n} A_i\right)$$

- union bound states that probability of union of events is smaller than the sum of first term. That is for $n = 2$, we have

$$P(A \cup B) \leq P(A) + P(B),$$

For any events $A_1, A_2, \ldots, A_n$, we have

Union bound and extension...

Union Bound

Recall the inclusion exclusion principle:

$$P((\cup_{i=1}^{n} A_i)) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j)$$
$$+ \sum_{i<j<k} P(A_i \cap A_j \cap A_k) - \cdots + (-1)^{n-1} P\left(\cap_{i=1}^{n} A_i\right)$$

- union bound states that probability of union of events is smaller than the sum of first term. That is for $n = 2$, we have

$$P(A \cup B) \leq P(A) + P(B),$$

  For any events $A_1, A_2, \ldots, A_n$, we have

$$P\left(\cup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i) \quad \text{(Union Bound)}$$

Generalized Union Bounds and Bonferroni Inequalities...

Generalized Union Bounds and Bonferroni Inequalities...

Generalized Union Bounds

Let $A_1, A_2, \ldots, A_n$ be events, then

$$P((\cup_{i=1}^{n} A_i)) \leq \sum_{i=1}^{n} P(A_i)$$

$$P((\cup_{i=1}^{n} A_i)) \geq \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j)$$

$$P((\cup_{i=1}^{n} A_i)) \geq \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k)$$

$$\cdots\cdots \quad \cdots \quad \cdots\cdots \quad \cdots\cdots$$

Generalized Union Bounds and Bonferroni Inequalities...

Generalized Union Bounds

Let $A_1, A_2, \ldots, A_n$ be events, then

$$P((\cup_{i=1}^{n} A_i)) \leq \sum_{i=1}^{n} P(A_i)$$

$$P((\cup_{i=1}^{n} A_i)) \geq \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j)$$

$$P((\cup_{i=1}^{n} A_i)) \leq \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k)$$

$$\ldots\ldots \quad \ldots \quad \ldots\ldots \quad \ldots\ldots$$

1 If we stop at the second term, we obtain a lower bound

Generalized Union Bounds and Bonferroni Inequalities...

Generalized Union Bounds

Let $A_1, A_2, \dots, A_n$ be events, then

$$P((\cup_{i=1}^{n} A_i)) \leq \sum_{i=1}^{n} P(A_i)$$

$$P((\cup_{i=1}^{n} A_i)) \geq \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j)$$

$$P((\cup_{i=1}^{n} A_i)) \leq \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k)$$

$$\dots\dots \quad \dots \quad \dots\dots \quad \dots\dots$$

1. If we stop at the second term, we obtain a lower bound
2. If we stop at the third term, we obtain an upper bound, etc

Generalized Union Bounds and Bonferroni Inequalities...

Generalized Union Bounds

Let $A_1, A_2, \ldots, A_n$ be events, then

$$P((\cup_{i=1}^{n} A_i)) \leq \sum_{i=1}^{n} P(A_i)$$

$$P((\cup_{i=1}^{n} A_i)) \geq \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j)$$

$$P((\cup_{i=1}^{n} A_i)) \geq \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k)$$

$$\ldots\ldots \quad \ldots \quad \ldots\ldots \quad \ldots\ldots$$

1. If we stop at the second term, we obtain a lower bound
2. If we stop at the third term, we obtain an upper bound, etc
3. In general, if we write an odd number of terms, we get an upper bound

Generalized Union Bounds and Bonferroni Inequalities...

Generalized Union Bounds

Let $A_1, A_2, \ldots, A_n$ be events, then

$$P((\cup_{i=1}^{n} A_i)) \leq \sum_{i=1}^{n} P(A_i)$$

$$P((\cup_{i=1}^{n} A_i)) \geq \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j)$$

$$P((\cup_{i=1}^{n} A_i)) \geq \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k)$$

$$\ldots\ldots \quad \ldots \quad \ldots\ldots \quad \ldots\ldots$$

1. If we stop at the second term, we obtain a lower bound
2. If we stop at the third term, we obtain an upper bound, etc
3. In general, if we write an odd number of terms, we get an upper bound
4. If we write an even number of terms, we get a lower bound

Markov and Chebyshev Inequalities...

Markov and Chebyshev Inequalities...

Markov and Chebyshev Inequalities

If X is any nonnegative random variable, then

Markov and Chebyshev Inequalities

If X is any nonnegative random variable, then

$$P(X \geq a) \leq \frac{E[X]}{a}.$$

Markov and Chebyshev Inequalities...

Markov and Chebyshev Inequalities

If X is any nonnegative random variable, then

$$P(X \geq a) \leq \frac{E[X]}{a}.$$

This inequality is called Markov inequality.

Markov and Chebyshev Inequalities...

Markov and Chebyshev Inequalities

If X is any nonnegative random variable, then

$$P(X \geq a) \leq \frac{E[X]}{a}.$$

This inequality is called Markov inequality. Moreover, let $b > 0$, then

Markov and Chebyshev Inequalities...

Markov and Chebyshev Inequalities

If X is any nonnegative random variable, then

$$P(X \geq a) \leq \frac{E[X]}{a}.$$

This inequality is called Markov inequality. Moreover, let $b > 0$, then

$$P(\,|X - E[X]| \geq b\,) \leq \frac{\mathrm{Var}(X)}{b^2}$$

Markov and Chebyshev Inequalities...

Markov and Chebyshev Inequalities

If X is any nonnegative random variable, then

$$P(X \geq a) \leq \frac{E[X]}{a}.$$

This inequality is called Markov inequality. Moreover, let $b > 0$, then

$$P(\,|X - E[X]| \geq b\,) \leq \frac{\mathrm{Var}(X)}{b^2}$$

The above inequality is called Chebyshev inequality.

- Chebyshev inequality states that the difference between X and $E[X]$ is bounded by $\mathrm{Var}(X)$

$$E[Y] = E[(X - E[X])^2] \overset{\text{def of var}}{=} \mathrm{Var}(X)$$

$$\Rightarrow P(Y \geq b^2) \leq \frac{\mathrm{Var}(X)}{b^2}$$

$$\Rightarrow P((X - E[X])^2 \geq b^2) \leq \frac{\mathrm{Var}(X)}{b^2}$$

$Y = (X - E[X])^2$ is non-neg.

$$P(Y \geq b^2) \leq \frac{E[Y]}{b^2} \quad [\text{Markov}]$$

$$\Rightarrow P(|X - E[X]| \geq b) \leq \frac{\mathrm{Var}(X)}{b^2}$$

Answer to previous problem...

Example of Markov Inequality...

Example of Markov Inequality...

Example (Markov Inequality)

Let $X \sim \text{Binomial}(n, p)$.

Example of Markov Inequality...

Example (Markov Inequality)

Let $X \sim \text{Binomial}(n, p)$. Using Markov inequality, find an upper bound on $P(X \geq \alpha n)$, where $p < \alpha < 1$.

Example of Markov Inequality...

Example (Markov Inequality)

Let $X \sim \text{Binomial}(n, p)$. Using Markov inequality, find an upper bound on $P(X \geq \alpha n)$, where $p < \alpha < 1$. Verify this bound for $p = 1/2$ and $\alpha = 3/4$.

Answer to previous problem...

Example of Chebychev Inequality...

Example of Chebychev Inequality...

Example (Chebychev Inequality)

Let $X \sim \text{Binomial}(n, p)$.

Example of Chebychev Inequality...

Example (Chebychev Inequality)

Let $X \sim \text{Binomial}(n, p)$. Using Chebyshev inequality, find an upper bound on $P(X \geq \alpha n)$, where $p < \alpha < 1$.

Example of Chebychev Inequality...

Example (Chebychev Inequality)

Let $X \sim \text{Binomial}(n, p)$. Using Chebyshev inequality, find an upper bound on $P(X \geq \alpha n)$, where $p < \alpha < 1$. Verify this bound for $p = 1/2$ and $\alpha = 3/4$.

Answer to previous problem...

Chernoff Bounds...

Chernoff Bounds...

Chernoff Bound

Let X be a random variable and $a \in \mathbb{R}$.

Chernoff Bounds...

Chernoff Bound

Let X be a random variable and $a \in \mathbb{R}$. Let $M_X(s) = E[e^{sX}]$ be the moment generating function.

Chernoff Bounds...

Chernoff Bound

Let X be a random variable and $a \in \mathbb{R}$. Let $M_X(s) = E[e^{sX}]$ be the moment generating function. Then the following holds

Chernoff Bounds...

Chernoff Bound

Let X be a random variable and $a \in \mathbb{R}$. Let $M_X(s) = E[e^{sX}]$ be the moment generating function. Then the following holds

$$P(X \geq a) \leq e^{-sa} M_X(s), \quad \text{for all } s > 0$$

$$P(X \leq a) \leq e^{-sa} M_X(s), \quad \text{for all } s < 0$$

Chernoff Bounds...

Chernoff Bound

Let X be a random variable and $a \in \mathbb{R}$. Let $M_X(s) = E[e^{sX}]$ be the moment generating function. Then the following holds

$$P(X \geq a) \leq e^{-sa} M_X(s), \quad \text{for all } s > 0$$
$$P(X \leq a) \leq e^{-sa} M_X(s), \quad \text{for all } s < 0$$

Since, the above holds for any s, we have the following

Chernoff Bounds...

Chernoff Bound

Let X be a random variable and $a \in \mathbb{R}$. Let $M_X(s) = E[e^{sX}]$ be the moment generating function. Then the following holds

$$P(X \geq a) \leq e^{-sa} M_X(s), \quad \text{for all } s > 0$$
$$P(X \leq a) \leq e^{-sa} M_X(s), \quad \text{for all } s < 0$$

Since, the above holds for any s, we have the following

$$P(X \geq a) \leq \min_{s>0} e^{-sa} M_X(s)$$
$$P(X \leq a) \leq \min_{s<0} e^{-sa} M_X(s)$$

Answer to previous problem...

Cauchy Schwarz Inequality...

Cauchy Schwarz Inequality...

Linear Alg

$\langle u, v \rangle \le \|u\| \, \|v\|$

Cauchy Schwarz Inequality

For any two random variables X and Y we have

$$E[XY] \leq \sqrt{E[X^2]E[Y^2]}$$

where equality holds if and only if $X = \alpha Y$, for some constant $\alpha \in \mathbb{R}$.

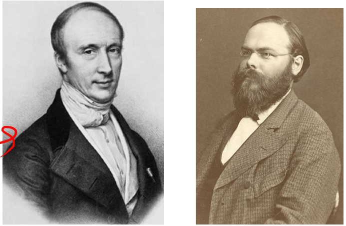


Figure: Left: Cauchy, Right: Schwarz

Answer to previous problem...

Answer to previous problem...

Example of Cauchy Schwarz...

Example of Cauchy Schwarz...

Example (Application of Cauchy Schwarz)

For any two random variables X and Y, show that

Example of Cauchy Schwarz...

Example (Application of Cauchy Schwarz)

For any two random variables X and Y, show that

$$|\rho(X, Y)| \leq 1$$

Example of Cauchy Schwarz...

Example (Application of Cauchy Schwarz)

For any two random variables X and Y, show that

$$|\rho(X, Y)| \leq 1$$

$$-1 \leq \rho(X, Y) \leq 1$$

using Cauchy Schwarz inequality.

Can you try?

Example of Cauchy Schwarz...

Example (Application of Cauchy Schwarz)

For any two random variables X and Y, show that

$$|\rho(X, Y)| \leq 1$$

using Cauchy Schwarz inequality. Furthermore, show that $|\rho(X, Y)| = 1$ if and only if $Y = aX + b$ for some constants $a, b \in \mathbb{R}$.

Answer to previous problem...

Convex Functions and Jensen's Inequality...

Convex Functions and Jensen's Inequality...

Definition of convex function

A function $g : I \to \mathbb{R}$ is convex if for any two points x and y in I and any $\alpha \in [0, 1]$, we have

Convex Functions and Jensen's Inequality...

Definition of convex function

A function $g : I \to \mathbb{R}$ is convex if for any two points x and y in I and any $\alpha \in [0, 1]$, we have

$$g(\alpha x + (1 - \alpha)y) \leq \alpha g(x) + (1 - \alpha)g(y)$$

Convex Functions and Jensen's Inequality...

Definition of convex function

A function $g : I \to \mathbb{R}$ is convex if for any two points x and y in I and any $\alpha \in [0, 1]$, we have

$$g(\alpha x + (1 - \alpha)y) \leq \alpha g(x) + (1 - \alpha)g(y)$$

If the above inequality is $\geq$, then the function g is concave.

- Here $\alpha x + (1 - \alpha)y$ is the weighted average of x and y

Convex Functions and Jensen's Inequality...

Definition of convex function

A function $g : I \to \mathbb{R}$ is convex if for any two points x and y in I and any $\alpha \in [0, 1]$, we have

$$g(\alpha x + (1 - \alpha)y) \leq \alpha g(x) + (1 - \alpha)g(y)$$

If the above inequality is $\geq$, then the function g is concave.

- Here $\alpha x + (1 - \alpha)y$ is the weighted average of x and y
- Here $\alpha g(x) + (1 - \alpha)g(y)$ is the weighted average of x and y

Convex Functions and Jensen's Inequality...

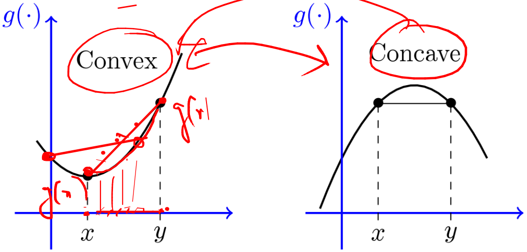
Definition of convex function

A function $g : I \to \mathbb{R}$ is convex if for any two points x and y in I and any $\alpha \in [0,1]$, we have

$$g(\alpha x + (1-\alpha)y) \leq \alpha g(x) + (1-\alpha) g(y)$$

If the above inequality is $\geq$, then the function g is concave.

- Here $\alpha x + (1-\alpha)y$ is the weighted average of x and y
- Here $\alpha g(x) + (1-\alpha)g(y)$ is the weighted average of x and y



not convex

$tx + (1-t)y$

line segment

betw x, y.

Convex Functions and Jensen's Inequality...

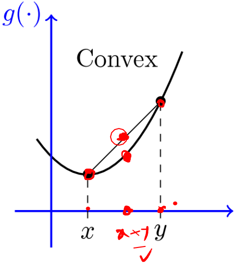
Definition of convex function

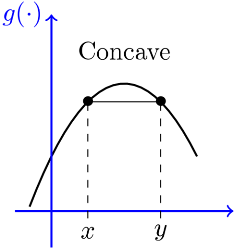
A function $g : I \to \mathbb{R}$ is convex if for any two points x and y in I and any $\alpha \in [0, 1]$, we have

$$g(\alpha x + (1-\alpha)y) \leq \alpha g(x) + (1-\alpha)g(y)$$

If the above inequality is $\geq$, then the function g is concave.

- Here $\alpha x + (1-\alpha)y$ is the weighted average of x and y
- Here $\alpha g(x) + (1-\alpha)g(y)$ is the weighted average of x and y





- From the definition of convexity on left, we conclude

$$E[g(X)] \geq g(E[X])$$

Jensen's Inequality...

Jensen's Inequality...

Jensen's inequality

If $g(x)$ is a convex function on R_X, and $E[g(X)]$ and $g(E[X])$ are finite, then

Jensen's Inequality...

Jensen's inequality

If $g(x)$ is a convex function on R_X, and $E[g(X)]$ and $g(E[X])$ are finite, then

$$E[g(X)] \geq g(E[X]).$$

Jensen's Inequality...

Jensen's inequality

If $g(x)$ is a convex function on R_X, and $E[g(X)]$ and $g(E[X])$ are finite, then

$$E[g(X)] \geq g(E[X]).$$

- To know whether a function is convex,

Jensen's Inequality...

Jensen's inequality

If $g(x)$ is a convex function on R_X, and $E[g(X)]$ and $g(E[X])$ are finite, then

$$E[g(X)] \geq g(E[X]).$$

- To know whether a function is convex, a useful method for differentiable function is second derivative test:

Jensen's Inequality...

Jensen's inequality

If $g(x)$ is a convex function on R_X, and $E[g(X)]$ and $g(E[X])$ are finite, then

$$E[g(X)] \geq g(E[X]).$$

- To know whether a function is convex, a useful method for differentiable function is second derivative test: A twice differentiable function $g : I \to \mathbb{R}$ is convex if and only if $g''(x) \geq 0$ for all $x \in I$

Jensen's Inequality...

$E[x^2] \geq (E[x])^2$ $\boxed{g = x^2}$

$g''(x) \geq 0$

$g''(x) = 2 \geq 0$

$\Rightarrow g$ is convex

Jensen's inequality

If $g(x)$ is a convex function on R_X, and $E[g(X)]$ and $g(E[X])$ are finite, then

$$E[g(X)] \geq g(E[X]).$$

- To know whether a function is convex, a useful method for differentiable function is second derivative test: A twice differentiable function $g : I \to \mathbb{R}$ is convex if and only if $g''(x) \geq 0$ for all $x \in I$
- For example, $g(x) = x^2$ is convex in $\mathbb{R}$

$E[-\log x] \geq -\log(E[x])$

$\log x \rightarrow$ concave

$-\log x \rightarrow$ convex

Application of Jensen's Inequality...

Application of Jensen's Inequality...

Example (Jensen's Inequality)

Consider a random variable X with $E[X] = 10$, and X being positive.

Application of Jensen's Inequality...

Example (Jensen's Inequality)

Consider a random variable X with $E[X] = 10$, and X being positive. Estimate the following quantities

Application of Jensen's Inequality...

Example (Jensen's Inequality)

Consider a random variable X with $E[X] = 10$, and X being positive. Estimate the following quantities

1. $E[\frac{1}{X+1}]$

Application of Jensen's Inequality...

Example (Jensen's Inequality)

Consider a random variable X with $E[X] = 10$, and X being positive. Estimate the following quantities

1. $E[\frac{1}{X+1}]$
2. $E[e^{\frac{1}{X+1}}]$

Application of Jensen's Inequality...

Example (Jensen's Inequality)

Consider a random variable X with $E[X] = 10$, and X being positive. Estimate the following quantities

1. $E[\frac{1}{X+1}]$
2. $E[e^{\frac{1}{X+1}}]$
3. $E[\ln \sqrt{X}]$

Answer to previous problem...

Answer to previous problem...

Outline

Definition of Sample Mean

i.i.d = independent + Identical

Same mean

Definition of Sample Mean

Let $X_1, X_2, \ldots, X_n$ be n i.i.d. random variables,

Definition of Sample Mean

Definition of Sample Mean

Let $X_1, X_2, \ldots, X_n$ be n i.i.d. random variables, then the sample mean $\overline{X}$ is defined as follows

Definition of Sample Mean

Definition of Sample Mean

Let $X_1, X_2, \ldots, X_n$ be n i.i.d. random variables, then the sample mean $\overline{X}$ is defined as follows

$$\overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

Definition of Sample Mean

Definition of Sample Mean

Let $X_1, X_2, \ldots, X_n$ be n i.i.d. random variables, then the sample mean $\overline{X}$ is defined as follows

$$\overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

It is easy to establish the following:

Definition of Sample Mean

Definition of Sample Mean

Let $X_1, X_2, \ldots, X_n$ be n i.i.d. random variables, then the sample mean $\overline{X}$ is defined as follows

$$\overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

It is easy to establish the following:

1. $E[\bar{X}] = E[X]$

Definition of Sample Mean

Definition of Sample Mean

Let $X_1, X_2, \ldots, X_n$ be n i.i.d. random variables, then the sample mean $\overline{X}$ is defined as follows

$$\overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

It is easy to establish the following:

1. $E[\bar{X}] = E[X]$ Sample
2. $\mathrm{Var}(\bar{X}) = \dfrac{\mathrm{Var}(X)}{n}$

Answer to previous problem...

Weak Law of Large Numbers...

Weak Law of Large Numbers...

Weak Law of Large Numbers

Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with mean $E[X_i] = \mu < \infty$. Then for any $\epsilon > 0$,

Weak Law of Large Numbers...

Weak Law of Large Numbers

Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with mean $E[X_i] = \mu < \infty$. Then for any $\epsilon > 0$,

$$\underbrace{\lim_{n\to\infty} P(|\bar{X} - \mu| \geq \epsilon) = 0}$$

$\epsilon = 10^{-10}$

Central Limit Theorem...

Central Limit Theorem...

Central Limit Theorem

Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with expected value $E[X_i] = \mu < \infty$ and variance $0 < \mathrm{Var}(X_i) = \sigma^2 < \infty$. Then, the random variable

Central Limit Theorem...

Central Limit Theorem

Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with expected value $E[X_i] = \mu < \infty$ and variance $0 < \text{Var}(X_i) = \sigma^2 < \infty$. Then, the random variable

$$Z_n = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sqrt{n}\sigma}$$

Central Limit Theorem...

Central Limit Theorem

Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with expected value $E[X_i] = \mu < \infty$ and variance $0 < \mathrm{Var}(X_i) = \sigma^2 < \infty$. Then, the random variable

$$Z_n = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sqrt{n}\sigma}$$

converges in distribution to the standard normal random variable as n goes to infinity, that is

Central Limit Theorem...

Central Limit Theorem

Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with expected value $E[X_i] = \mu < \infty$ and variance $0 < \mathrm{Var}(X_i) = \sigma^2 < \infty$. Then, the random variable

$$Z_n = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sqrt{n}\sigma}$$

converges in distribution to the standard normal random variable as n goes to infinity, that is

$$\lim_{n\to\infty} P(X_n \le x) = \Phi(x), \quad \text{for all } x \in \mathbb{R},$$

where $\Phi(x)$ is the standard normal CDF.

Central Limit Theorem...

Central Limit Theorem

Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with expected value $E[X_i] = \mu < \infty$ and variance $0 < \mathrm{Var}(X_i) = \sigma^2 < \infty$. Then, the random variable

$$Z_n = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sqrt{n}\sigma}$$

converges in distribution to the standard normal random variable as n goes to infinity, that is

$$\lim_{n\to\infty} P(X_n \leq x) = \Phi(x), \quad \text{for all } x \in \mathbb{R},$$

where $\Phi(x)$ is the standard normal CDF.

- It does not matter what the distribution of X_i is

Central Limit Theorem...

Central Limit Theorem

Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with expected value $E[X_i] = \mu < \infty$ and variance $0 < \mathrm{Var}(X_i) = \sigma^2 < \infty$. Then, the random variable

$$Z_n = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sqrt{n}\sigma}$$

converges in distribution to the standard normal random variable as n goes to infinity, that is

$$\lim_{n\to\infty} P(X_n \leq x) = \Phi(x), \quad \text{for all } x \in \mathbb{R},$$

where $\Phi(x)$ is the standard normal CDF.

- It does not matter what the distribution of X_i is
- The X_i can be discrete, continuous, or mixed random variables

1. Let X_i be Bernoulli(p)
2. Then $E[X_i] = p, \mathrm{Var}(X_i) = p(1-p)$
3. $Y_n = X_1 + X_2 + \cdots + X_n$ has Binomial((n, p))
4. Hence,

$$Z_n = \frac{Y_n - np}{\sqrt{np(1-p)}}$$

5. The figure on the right shows PMF of Z_n for different values of n
6. As we observe, the shape of PMF gets closer to a normal PDF

1 Let X_i be Bernoulli(p)

2 Then $E[X_i] = p, \text{Var}(X_i) = p(1-p)$

3 $Y_n = X_1 + X_2 + \cdots + X_n$ has Binomial$((n, p))$

4 Hence,

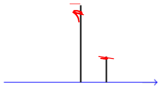
$$Z_n = \frac{Y_n - np}{\sqrt{np(1-p)}}$$

5 The figure on the right shows PMF of Z_n for different values of n

6 As we observe, the shape of PMF gets closer to a normal PDF

$$Z_1 = \frac{X_1 - p}{\sqrt{p(1-p)}}$$

PMF of Z_1

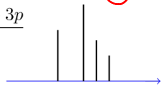


$$Z_2 = \frac{X_1 + X_2 - 2p}{\sqrt{2p(1-p)}}$$

PMF of Z_2

$$Z_3 = \frac{X_1 + X_2 + X_3 - 3p}{\sqrt{3p(1-p)}}$$

PMF of Z_3



$e^{-x^2/2}$

$$Z_{30} = \frac{\sum_{i=1}^{30} X_i - 30p}{\sqrt{30p(1-p)}}$$

PMF of Z_{30}