Introduction to Statistics (MAT 283)

Dipti Dubey

Department of Mathematics
Shiv Nadar University

Feburary 8, 2021

$$\mathrm{Var}(x) = E(x^2) - (E(x))^2 > 0$$

$$\Rightarrow \quad E(x^2) > \begin{array}{c}(E(x))^2 \\ \vee \;\; \geq \\ E(x)\end{array}$$

$$Y = X^5$$

$$E(X^{10}) \leq E(X^5)$$

Table of Contents

Conditional Variance

Special Bivariate Distributions

Transformation of Random Variables

Theorem: Let X and Y be two random variables with mean μ_X and μ_Y , and standard deviation X and Y , respectively. If the conditional expectation of Y given $X = x$ is linear in x, then

$$E(Var(Y|X)) = (1 - \rho^2)Var(Y).$$

where ρ denotes the correlation coefficient of X and Y .

EXAMPLE: Let $E(Y|X = x) = 2x$ and $Var(Y|X = x) = 4x^2$, and let X have a uniform distribution on the interval from 0 to 1. What is the variance of Y ?

Table of Contents

Conditional Variance

Special Bivariate Distributions

Transformation of Random Variables

Bivariate Bernoulli Distribution:

A discrete bivariate random variable (X, Y) is said to have the bivariate Bernoulli distribution if its joint probability density is of the form

$$f(x,y) = \begin{cases} \frac{1}{x!y!(1-x-y)!} p_1^x p_2^y (1-p_1-p_2)^{1-x-y} & \text{if } x, y = 0, 1; \\ 0 & \text{otherwise,} \end{cases}$$

where $p_1, p_2 > 0, p_1 + p_2 < 1$ and $x + y \leq 1$. We denote a bivariate Bernoulli random variable by writing $(X, Y) \sim BER(p_1, p_2)$.

Theorem: Let $(X, Y) \sim BER(p_1, p_2)$, where p_1 and p_2 are parameters. Then

$$E(X) = p_1$$

$$E(Y) = p_2$$

$$Var(X) = p_1(1 - p_1)$$

$$Var(Y) = p_2(1 - p_2)$$

$$Cov(X, Y) = -p_1 p_2$$

$$M(s, t) = 1 - p_1 - p_2 + p_1 e^s + p_2 e^t.$$

Bivariate Binomial Distribution:

A discrete bivariate random variable (X, Y) is said to have the bivariate binomial distribution with parameters n, p_1, p_2 if its joint probability density is of the form

$$f(x,y) = \begin{cases} \frac{n}{x!y!(n-x-y)!} p_1^x p_2^y (1-p_1-p_2)^{n-x-y} & \text{if } x, y = 0, 1, \ldots, n; \\ 0 & \text{otherwise,} \end{cases}$$

where $p_1, p_2 > 0, p_1 + p_2 < 1$ and $x + y \leq n$. We denote a bivariate Bernoulli random variable by writing $(X, Y) \sim BIN(n, p_1, p_2)$.

EXAMPLE: A certain game involves rolling a fair die and watching the numbers of rolls of 4 and 5. What is the probability that in 10 rolls of the die one 4 and three 5 will be observed?

Theorem: Let $(X, Y) \sim BER(n, p_1, p_2)$, where n,p_1 and p_2 are parameters. Then

$$E(X) = np_1$$

$$E(Y) = np_2$$

$$Var(X) = np_1(1 - p_1)$$

$$Var(Y) = np_2(1 - p_2)$$

$$Cov(X, Y) = -np_1 p_2$$

$$M(s, t) = (1 - p_1 - p_2 + p_1 e^s + p_2 e^t)^n.$$

Bivariate Normal Distribution:

A continuous bivariate random variable (X, Y) is said to have the bivariate normal distribution if its joint probability density function is of the form

$$f(x,y)=\begin{cases}\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}e^{\frac{1}{2}Q(x,y)} & \text{if } 0<x,y<\infty;\\ 0 & \text{otherwise,}\end{cases}$$

where $\mu_1,\ \mu_2 \in \mathbb{R},\ \sigma_1,\ \sigma_2 \in (0,1)$ and $\rho \in (-1,1)$ are parameters, and

$$Q(x,y)=\frac{1}{1-\rho^2}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2-2\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right)+\left(\frac{y-\mu_2}{\sigma_2}\right)^2\right].$$

We denote this bivariate normal random variable by writing $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$.

- The graph of f(x, y) has a shape of a mountain.

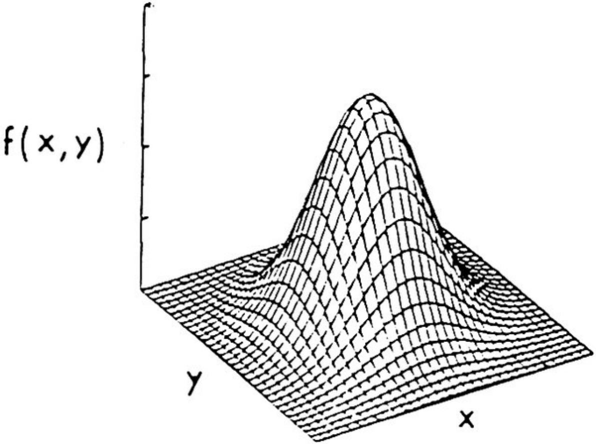


Table of Contents

Conditional Variance

Special Bivariate Distributions

Transformation of Random Variables

$(X, \underset{pdf}{f})$

$\underline{Y} = g(X)$

$\boxed{f_Y\,?}$

Suppose X is a random variable with pdf (or pmf) f and CDF F and $Y = r(X)$ be a function of X (for example $Y = X^2$, $Y = e^X$).

EXAMPLE: Let X be a random variable such that $R_X = \{-1, 0, 1\}$ and

$$P(-1) = 1/4, \; P(0) = 1/2 \text{ and } P(1) = 1/4.$$

Find pmf of $Y = X^2$.

Suppose X is a random variable with pdf (or pmf) f and CDF F and $Y = r(X)$ be a function of X (for example $Y = X^2,\ Y = e^X$).

Q. How to find pdf and CDF of random variable Y?

Suppose X is a random variable with pdf (or pmf) f and CDF F and $Y = r(X)$ be a function of X (for example $Y = X^2$, $Y = e^X$).

Q. How to find pdf and CDF of random variable Y?

EXAMPLE: Let X be a random variable such that $R_X = \{-1, 0, 1\}$ and

$$P(-1) = 1/4, \; P(0) = 1/2 \text{ and } P(1) = 1/4.$$

Find pmf of $Y = X^2$.

$R_Y = \{0, 1\}$

$P(Y=0) = P(X=0) = 1/2$, $P(Y=1) = P(X=1) + P(X=-1)$

$= 1/4 + 1$

More generally, given a set of random variables $X_1, X_2, ..., X_n$ and their joint pdf, we shall be interested in finding the pdf of some random variable

$$Y = u(X_1, X_2, ..., X_n).$$

This means that the values of Y are related to those of the Xs by means of the equation

$$y = u(x_1, x_2, ..., x_n)$$

More generally, given a set of random variables $X_1, X_2, ..., X_n$ and their joint pdf, we shall be interested in finding the pdf of some random variable

$$Y = u(X_1, X_2, ..., X_n).$$

This means that the values of Y are related to those of the Xs by means of the equation

$$y = u(x_1, x_2, ..., x_n)$$

METHOD FOR FINDING PDF of Y:

- Distribution Function Technique
- Transformation Technique
- MGF Technique

DISTRIBUTION FUNCTION TECHNIQUE :

To find the probability density function of a transformation of continuous random variables is to determine its CDF and then pdf by differentiation.

EXAMPLE: If the pdf of X is given by

$$f(x) = \begin{cases} 6x(1-x) & \text{for} \quad 0 < x < 1; \\ 0 & \text{otherwise,} \end{cases}$$

find the pdf of $Y = X^3$.

Letting $G(y)$ denote the value of the CDF of Y at y, we can write

$$\begin{aligned}
G(y) &= P(Y \leq y) \\
&= P(X^3 \leq y) \\
&= P(X \leq y^{1/3}) \\
&= \int_0^{y^{1/3}} 6x(1-x)dx \\
&= 3y^{2/3} - 2y,
\end{aligned}$$

Letting $G(y)$ denote the value of the CDF of Y at y, we can write

$$\begin{aligned} G(y) &= P(Y \leq y) \\ &= P(X^3 \leq y) \\ &= P(X \leq y^{1/3}) \\ &= \int_0^{y^{1/3}} 6x(1-x)dx \\ &= 3y^{2/3} - 2y, \end{aligned}$$

and hence $g(y) = \frac{dG}{dy} = 2(y^{-1/3} - 1)$ for $0 < y < 1$; elsewhere, $g(y) = 0$.

EXAMPLE: If the joint density of X_1 and X_2 is given by

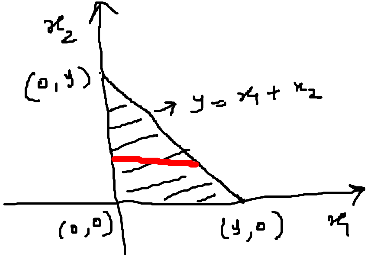
$$f(x_1, x_2) = \begin{cases} 6e^{-3x_1-2x_2} & \text{for} \quad x_1 > 0; x_2 > 0 \\ 0 & \text{otherwise,} \end{cases}$$

$y = x_1 + x_2$

find the pdf of $Y = X_1 + X_2$.

$$F_Y(y) = P(Y \leq y)$$

$$= P(X_1 + X_2 \leq y)$$



$$= \int_0^y \int_0^{y-x_2} 6\, e^{-3x_1 - 2x_2}\, dx_1\, dx_2$$

$$= 1 + 2e^{-3y} - 3e^{-2y}$$

$$f_Y(y) = \frac{dF_Y}{dy} = \begin{cases} 6\left(e^{-2y} - e^{-3y}\right), & y > 0 \\ 0 & \text{otherwise} \end{cases}$$