Mathematics 3 (SM 211): Probability and Statistics

Amit Chattopadhyay

IIIT-Bangalore

Ch. 6: Mathematical Expectation - II

Syllabus: Outline

Probability:

1. The Concept of Probability
2. Compound or Joint Experiment
3. Probability Distributions-I
4. Mathematical Expectation-I
5. Probability Distributions-II
6. Mathematical Expectation-II
7. Some Important Continuous Univariate Distributions
8. Convergence of a Sequence of Random Variables and Limit Theorems

Statistics:

1. Random Samples
2. Sampling Distributions
3. Estimation of Parameters
4. Testing of Hypothesis

Reference Books

1. Mathematical Probability by A. Banerjee, S.K. De and S. Sen

2. Mathematical Statistics by S.K. De and S. Sen

3. Groundwork of Mathematical Probability and Statistics by Amritava Gupta

4. Introduction to Probability and Statistics for Engineers and Scientists by S.M. Ross

5. Introduction to Probability Models, by S.M. Ross

6. Probability and Statistics, (Schaum's Outlines) by Murray R Spiegel, John J Schiller and R Alu Srinivasan

Mathematical Expectation - II

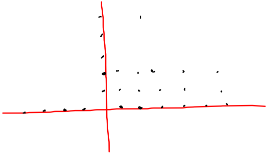
Objective

- Expectation - Moments
- Correlation coefficient - Properties
- Conditional Expectation, Regression
- Independence
- Joint characteristic function, Reproductive property

Expectation: Definition

Discrete Case

Let (X, Y) be a bivariate discrete random variable with spectrum $\{(x_i, y_j) : i = \pm 1, \pm 2, \ldots; j = \pm 1, \pm 2, \ldots\}$ and $g(X, Y)$ be a continuous real-valued function of X and Y. Then

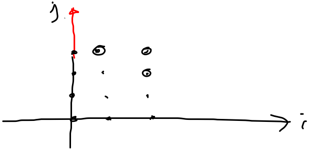


$$\begin{aligned} E(g(X,Y)) &= \sum_{j=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} g(x_i, y_j) P(X = x_i, Y = y_j) \\ &= \sum_{j=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} g(x_i, y_j) f_{X,Y}(x_i, y_j) \end{aligned}$$

provided the infinite series is absolutely convergent. Here $f_{X,Y}$ is the joint p.m.f. of (X, Y).

Expectation: Example

Ex 1: Let (X, Y) be a random variable with spectrum points $\{(i,j) : i = 0,1,2; j = 0,1,2,3\}$ and $f_{ij} = \frac{1}{9}$ for all i,j except $f_{13} = f_{22} = f_{23} = 0$. Compute $E\{|X - Y|\}$.



$$E\{|X-Y|\} = \sum_{j=0}^{3}\sum_{i=0}^{2} |i-j|\, p_{ij}$$

$$= \frac{11}{9}. \quad (\text{check!})$$

Expectation: Definition

Continuous Case

continuous

Let (X, Y) be a bivariate ~~discrete~~ random variable and $g(X, Y)$ be a continuous real-valued function of X and Y.

$$E(g(X,Y)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\, f_{X,Y}(x,y)dxdy$$

provided the improper integration is absolutely convergent. Here, $f_{X,Y}$ is joint p.d.f. of (X, Y).

PS-9, Prob 1: Let X, Y are independent standard normal variates, find the mean value of the greater of $|X|$ and $|Y|$.

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad -\infty < x < \infty$$

$$f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}, \quad -\infty < x < \infty$$

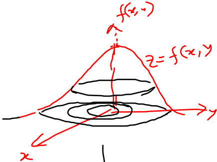
$$f_{X,Y}(x,y) = f_X(x) f_Y(y) = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}}, \quad \begin{array}{l} -\infty < x < \infty \\ -\infty < y < \infty \end{array}$$

H.W

$$E\{\max(|X|, |Y|)\} = ?$$

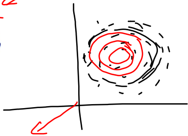
Moments

1. **Moments about origin of order $k+l$:** $\alpha_{kl} = E(X^k Y^l)$, k, l are non-negative integers.



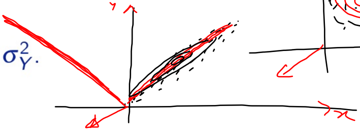
$\alpha_{00} = 1$, $\alpha_{10} = m_X$, $\alpha_{01} = m_Y$.

2. **Central moments of order $k+l$:** $\mu_{kl} = E\{(X - m_X)^k (Y - m_Y)^l\}$, k, l are non-negative integers.



$\mu_{00} = 1$, $\mu_{10} = 0$, $\mu_{01} = 0$, $\mu_{20} = \sigma_X^2$, $\mu_{02} = \sigma_Y^2$.

$\| \; E\{(X - m_X)\} = 0$



3. **Covariance:**

$$\mu_{11} = E\{(X - m_X)(Y - m_Y)\} = Cov(X, Y)$$

$$= E(XY - m_X Y - X m_Y + m_X m_Y)$$

Significance: Measure of the jointness or tendency of having linear relationship of the bivaraite distribution.

4. $\mu_{11} = \alpha_{11} - m_X m_Y$

Definition $E\left\{(X^* - m_{X^*})(Y^* - m_{Y^*})\right\} = E\left\{X^* Y^*\right\} = E\left\{\left(\frac{X - m_X}{\sigma_X}\right)\left(\frac{Y - m_Y}{\sigma_Y}\right)\right\}$

$\|$

$$\rho(X, Y) = Cov(X^*, Y^*) = \frac{Cov(X, Y)}{\sigma_X \sigma_Y}, \; (\sigma_X, \sigma_Y \neq 0)$$

where $X^* = \frac{X - m_X}{\sigma_X}$ and $Y^* = \frac{Y - m_Y}{\sigma_Y}$ are the standardised variates of X and Y, respectively.

1. If $a_1 (\neq 0)$, $a_2 (\neq 0)$, b_1, b_2 are constants, then

$$\rho(a_1 X + b_1, a_2 Y + b_2) = \frac{a_1 a_2}{|a_1||a_2|} \rho(X, Y).$$

Note: ρ is independent of choice of origin & units of measurements of r.v.s,

Pf: $\rho(a_1X+b_1, a_2Y+b_2) = E\{(a_1X+b_1)^*(a_2Y+b_2)^*\} = E\left\{\frac{a_1}{|a_1|}X^* \frac{a_2}{|a_2|}Y^*\right\}$

$= \frac{a_1 a_2}{|a_1||a_2|}\rho(X,Y)$.

$(a_1X+b_1)^* = \frac{a_1}{|a_1|}X^* \leftarrow \frac{a_1X+b_1-(a_1m_X+b_1)}{|a_1|\sigma_X}$,

$Var(a_1X+b_1) = E\{(a_1X+\cancel{b_1}-a_1m_X-\cancel{b_1})^2\}$

$= a_1^2 Var(X)$

1. If $a_1 (\neq 0)$, $a_2 (\neq 0)$, b_1, b_2 are constants, then

$$\rho(a_1 X + b_1, a_2 Y + b_2) = \frac{a_1 a_2}{|a_1||a_2|} \rho(X, Y).$$

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

Pf: $\rho(X,Y) = E\{X^* Y^*\}$, $\quad E(X^{*2}) = 1$, $E(Y^{*2}) = 1$.

$$(X^* \pm Y^*)^2 \geq 0$$

$$\Rightarrow E\{X^{*2} + Y^{*2} \pm 2X^* Y^*\} \geq 0$$

$$\Rightarrow 2 \pm 2\rho(X,Y) \geq 0$$

$$\Rightarrow 1 \geq \rho(X,Y) \geq -1.$$

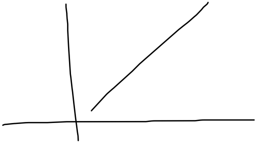
Correlation Coefficient: Properties

1. If $a_1 (\neq 0)$, $a_2 (\neq 0)$, b_1, b_2 are constants, then

$$\rho(a_1 X + b_1, a_2 Y + b_2) = \frac{a_1 a_2}{|a_1||a_2|} \rho(X, Y).$$

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

3. $\rho(X, Y) = \pm 1$ iff $X = aY + b$ for some constants a and b, i.e. if there's a linear relationship between X and Y



Pf: $\Rightarrow$ $\rho(X,Y) = 1 \Rightarrow E(X^* Y^*) = 1$

$$\Rightarrow E\{(X^* - Y^*)^2\} = 0$$

$$\Rightarrow X^* - Y^* = 0$$

$$\Rightarrow \frac{X - m_X}{\sigma_X} = \frac{Y - m_Y}{\sigma_Y}$$

$\Leftarrow$ Let, $X = aY + b$, $\rho(X,Y) = \rho(aY+b, Y) = \frac{a}{|a|}\rho(Y,Y) = \frac{a}{|a|} E(Y^{*2}) = \pm 1$.

Correlation Coefficient: Properties

1. If $a_1\,(\neq 0)$, $a_2\,(\neq 0)$, b_1, b_2 are constants, then
$$\rho(a_1X + b_1, a_2Y + b_2) = \frac{a_1 a_2}{|a_1||a_2|}\rho(X, Y).$$
2. $-1 \leq \rho(X, Y) \leq 1$
3. $\rho(X, Y) = \pm 1$ iff $X = aY + b$ for some constants a and b, i.e. if there's a linear relationship between X and Y
4. **Definition:** If $\rho(X, Y) = 0$, X, Y are called **uncorrelated**

~~X~~, Y are independent

$$\rho(X,Y) = \frac{E\{(X-m_X)(Y-m_Y)\}}{\sigma_X \sigma_Y}$$

(Not yet discussed!)

Correlation Coefficient: Properties

1. If $a_1\,(\neq 0)$, $a_2\,(\neq 0)$, b_1, b_2 are constants, then
$$\rho(a_1X+b_1, a_2Y+b_2) = \frac{a_1a_2}{|a_1||a_2|}\rho(X,Y).$$
2. $-1 \leq \rho(X,Y) \leq 1$
3. $\rho(X,Y) = \pm 1$ iff $X = aY + b$ for some constants a and b, i.e. if there's a linear relationship between X and Y
4. **Definition:** If $\rho(X,Y) = 0$, X, Y are called **uncorrelated**
5. If X, Y are independent, then X, X are uncorrelated. **Converse** is not true.

Converse is not true:

Consider random variable (X, Y) with spectrum $\{(0,0),(1,1),(-1,1)\}$ and $P(X=0, Y=0) = \frac{1}{2}$, $P(X=1, Y=1) = P(X=-1, Y=1) = \frac{1}{4}$.

$$P(X=0) = \frac{1}{2} \qquad P(Y=0) = \frac{1}{2}$$
$$P(X=1) = \frac{1}{4} \qquad P(Y=1) = \frac{1}{2}$$
$$P(X=-1) = \frac{1}{4}$$

$$P(X=1, Y=1) \neq P(X=1)\,P(Y=1)$$

$\Rightarrow$ X, Y are not independent.

$$\rho(X,Y) = \frac{E\{(X-m_X)(Y-m_Y)\}}{\sigma_X \sigma_Y} = \frac{\alpha_{11} - m_X m_Y}{\sigma_X \sigma_Y} = \frac{0 - 0}{\sigma_X \sigma_Y} = 0$$

$\Rightarrow$ X, Y are uncorrelated.

$$m_X = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{4} + (-1)\frac{1}{4} = 0$$

$$m_Y = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{1}{2}.$$

$$\alpha_{11} = E(XY) = \sum_{j=0,1} \sum_{i=0,1,-1} i \cdot j \; p_{ij}$$
$$= 0 \cdot 0\, p_{00} + 1 \cdot 1 \cdot p_{11} + (-1) \cdot 1 \cdot p_{-1,1}$$
$$= 0$$

Correlation Coefficient: Examples

HW.

PS-9 Problems 2, 3

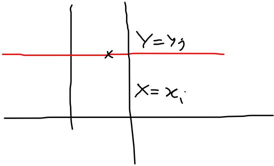
Conditional Expectations: Discrete Case

Definition

Let (X, Y) be a bivariate discrete random variable with joint p.m.f. f_{ij} at the spectrum point (x_i, y_j) for $i = \pm 1, \pm 2, \ldots; j = \pm 1, \pm 2, \ldots$. Then the *conditional expectation* or *conditional mean* of a continuous real-valued function $g(X, Y)$ of X and Y on the hypothesis $Y = y_j$ is defined as:

$$E\{g(X,Y)|Y = y_j\} = \sum_{j=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} g(x_i, y_j) f_{i|j}$$

$$= \frac{\sum\limits_{i=-\infty}^{\infty} g(x_i, y_j) f_{ij}}{f_{\bullet j}}.$$



$E\{g(X,Y)|X = x_i\}$ can be defined, similarly.

Definition

Let (X, Y) be a bivariate continuous random variable with joint p.d.f $f_{X,Y}(x, y)$. Then the conditional expectation or conditional mean of a continuous real-valued function $g(X, Y)$ of X and Y on the hypothesis $Y = y$ is defined as:

$$E(g(X, Y)|Y = y) = \int_{-\infty}^{\infty} g(x, y)\, f_X(x|y) dx.$$

$E(g(X, Y)|X = x)$ can be defined, similarly.

Conditional Expectations: Conditional Mean of X

Discrete Case:

- **Conditional Mean of X given $Y = y_j$:**

$$m_{X|j} = E\{X|Y = y_j\} = \sum_i x_i f_{i|j} = \frac{\sum_i x_i f_{ij}}{f_{\bullet j}}$$

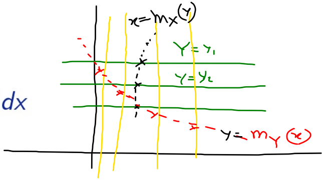
- **Conditional Mean of Y given $X = x_i$:**

$$m_{Y|i} = E\{Y|X = x_i\} = \sum_j y_j f_{j|i} = \frac{\sum_j y_j f_{ij}}{f_{i\bullet}}$$

Conditional Expectations: In Particular

Continuous Case:

- **Conditional Mean of X given $Y = y$:**



$$m_X(y) = E\{X|Y = y\} = \int_{-\infty}^{\infty} x\, f_X(x|y)dx$$

- **Conditional Mean of Y given $X = x$:**

$$m_Y(x) = E\{Y|X = x\} = \int_{-\infty}^{\infty} y\, f_Y(y|x)dy$$

Conditional Expectations: In Particular

Discrete Case:

- **Conditional Variance of X given $Y = y_j$:**

$$\sigma^2_{X|j} = Var\{X|Y = y_j\} = E\{(X - m_{X|j})^2|Y = y_j\}$$

- **Conditional Variance of Y given $X = x_i$:**

$$\sigma^2_{Y|i} = Var\{Y|X = x_i\} = E\{(Y - m_{Y|i})^2|X = x_i\}$$

Conditional Expectations: In Particular

Continuous Case:

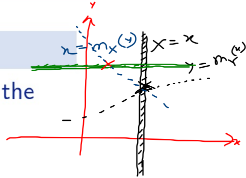
- **Conditional Variance of X given $Y = y$:**

$$\sigma_X^2(y) = Var(X|Y = y) = E\{(X - m_X(y))^2|Y = y\}\cancel{dx}$$

- **Conditional Variance of Y given $X = x$:**

$$\sigma_Y^2(x) = Var(Y|X = x) = E\{(Y - m_Y(x))^2|X = x\}.$$

Regression Curves



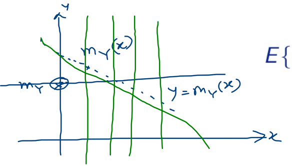
Regression Curve of Y on X

The conditional mean $m_Y(x)$ for a continuous distribution is called the *regression function* of Y on X.

The curve $y = m_Y(x)$ is called the *regression curve* of Y on X.

Geometrically, $m_Y(x)$ represents the y-coordinate of the centre of mass of the bivariate probability mass in the infinitesimal vertical strip bounded by x and $x + dx$.
The regression curve is the locus of this centre of mass as x varies.

(similarly, define for Regression Curve of X on Y)



(*) $E\{E\{Y|X\}\}$

$$E\{m_Y(X)\} = E\{E\{Y|X = x\}\}$$
$$= \int_{-\infty}^{\infty} f_X(x) \left\{ \int_{-\infty}^{\infty} y\, f_Y(y|x)dy \right\} dx = \int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty} y \cdot \frac{f_{X,Y}(x,y)}{f_X(x)}\, dy\right\} f_X(x)\, dx$$
$$= \int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty} y\, f_{X,Y}(x,y)\, dy\right\} dx$$
$$= \int_{-\infty}^{\infty} y\left\{\int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dx\right\} dy$$
$$= \int_{-\infty}^{\infty} y f_Y(y)dy$$
$$= m_Y$$

$$Var(m_Y(X)) = E\{\{m_Y(X) - m_Y\}^2\}$$

$\Longrightarrow$ measure of dispersion of the regression curve $y = m_Y(x)$ from the horizontal line $y = m_Y$.

Expectation of $\sigma_Y^2(X)$

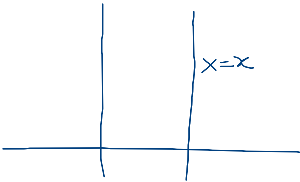
Show that: $\sigma_{yx}^2 = E\{\sigma_Y^2(X)\} = E\{(Y - m_Y(X))^2\}$

$$= \int_{-\infty}^{\infty} \sigma_Y^2(x)\, f_X(x)\, dx$$

$$= \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} (y - m_Y(x))^2 f_Y(y|x)\, dy \right\} f_X(x)\, dx$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (y - m_Y(x))^2 f_{X,Y}(x,y)\, dy\, dx$$

$$= E\left\{ (Y - m_Y(x))^2 \right\}$$

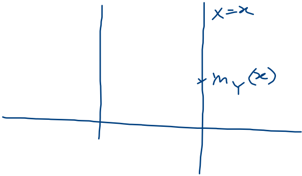


Interpretation: Measure of dispersion (or variance) of the bivariate distribution about the regression curve $y = m_Y(x)$

Minimum Property

Recall: Conditional second moment is minimum when taken about the corresponding conditional mean.

Theorem: For any real-valued continuous function $g(x)$, $E\{(Y-g(X))^2\}$ is minimum when $\boxed{g(x) = m_Y(x).}$



Pf:

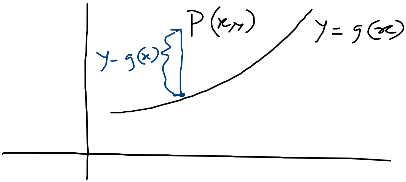
$$E\{(Y-g(X))^2\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (y-g(x))^2 f_{X,Y}(x,y)\,dx\,dy$$

$$= \int_{-\infty}^{\infty} f_X(x)\left\{\int_{-\infty}^{\infty} (y-g(x))^2 f_Y(y|x)\,dy\right\} dx$$

$$= \int_{-\infty}^{\infty} f_X(x)\, E\{(Y-g(x))^2 \mid X=x\}\, dx$$

↳ minimum when $g(x) = m_Y(x)$.

$E\{(Y - g(X))^2\}$: Measure of dispersion about the curve $y = g(x)$.
Among all continuous curves, this mean value is minimum for $y = m_Y(x)$.



Best Fitting Curve/ Least Square Regression Curve

Let $y = g(x; c_0, c_1, \ldots, c_n)$ be a family of curves, $c_0, c_1, \ldots, c_n$ being parameters of the family.

Let $S = E\{(Y - g(X; c_0, c_1, \ldots, c_n))^2\}$.

If S is minimum for $c_0 = c_0^*, c_1 = c_1^*, \ldots, c_n = c_n^*$, then the curve $y = g(x; c_0^*, c_1^*, \ldots, c_n^*)$ is called the *best fitting curve/ least square regression curve* to probability mass function.

Least Square Regression Lines

Let $y = c_0 + c_1 x$ be the family of lines

To minimize: $S(c_0, c_1) = E\left\{ (Y - c_0 - c_1 X)^2 \right\}$

$$= \frac{\partial}{\partial c_0} \iint (y - c_0 - c_1 x)^2 f_{X,Y}(x,y)\,dx\,dy$$

$$= -\iint 2(y - c_0 - c_1 x) f_{X,Y}(x,y)\,dx\,dy$$

$$= -2E(Y - c_0 - c_1 X)$$

Least Square Reg. line of Y on X

$y = \left(m_Y - \rho \frac{\sigma_Y}{\sigma_X} m_X\right) + \rho \frac{\sigma_Y}{\sigma_X} \cdot x$

$= m_Y + \rho \frac{\sigma_Y}{\sigma_X}(x - m_X)$

$$\frac{\partial S}{\partial c_0} = 0, \quad \frac{\partial S}{\partial c_1} = 0$$

$$\Rightarrow \begin{cases} E(Y - c_0 - c_1 X) = 0, \\ E\{X(Y - c_0 - c_1 X)\} = 0 \end{cases}$$

$$\Rightarrow \begin{cases} c_0 + c_1 m_X = m_Y \\ c_0 m_X + c_1 \alpha_{20} = \alpha_{11} \end{cases}$$

$$\Rightarrow \begin{aligned} c_0 m_X + c_1 m_X^2 &= m_X m_Y \\ c_0 m_X + c_1 \alpha_{20} &= \alpha_{11} \end{aligned}$$

$$\Rightarrow c_1 (m_X^2 - \alpha_{20}) = m_X m_Y - \alpha_{11}$$

$$\Rightarrow c_1^* = \frac{m_X m_Y - \alpha_{11}}{m_X^2 - \alpha_{20}} = \frac{\text{cov}(X,Y)}{\sigma_X^2}$$

$$= \frac{\rho \sigma_X \sigma_Y}{\sigma_X^2} = \rho \frac{\sigma_Y}{\sigma_X}$$

$$c_0^* = m_Y - \rho \frac{\sigma_Y}{\sigma_X} m_X$$

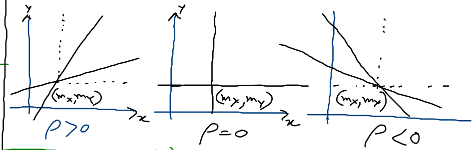
$(E(X))^2 - E(X^2)$

$= -\text{Var}(X) = -\sigma_X^2$

$E(X)E(Y) - E(XY)$

$= -\text{cov}(X,Y)$

Least Sq. Reg. line of Y on X :



$$Y = m_Y + \rho \frac{\sigma_Y}{\sigma_X}(x - m_X)$$

$$\Rightarrow \boxed{Y - m_Y = \rho \frac{\sigma_Y}{\sigma_X}(x - m_X)}$$

$$\beta_{yx} = \rho \frac{\sigma_Y}{\sigma_X} \quad \text{(ii)}$$

(Reg. coefficient of Y on X)

$$\Leftrightarrow \boxed{\frac{Y - m_Y}{\sigma_Y} = \rho \frac{x - m_X}{\sigma_X}} \quad \text{(i)}$$

Note:

$$\boxed{\beta_{yx}\beta_{xy} = \rho^2}$$

Similarly Check! Least Sq. Reg. Line of X on Y :

$$x = d_0^* + d_1^* y \qquad \begin{cases} d_0^* = m_X - \rho \dfrac{\sigma_X}{\sigma_Y} m_Y \\ d_1^* = \rho \dfrac{\sigma_X}{\sigma_Y} \end{cases}$$

$$\boxed{\frac{Y - m_Y}{\sigma_Y} = \frac{1}{\rho} \frac{x - m_X}{\sigma_X}} \quad \text{(iii)}$$

$$\beta_{xy} = \text{Reg. coeff of X on Y} = \rho \frac{\sigma_X}{\sigma_Y} \quad \text{(iv)}$$

PS-9
Prob 6

Reg. lines of dist. of (X,Y):

$$x+6y=6 \quad \text{—(i)}$$
$$3x+2y=10 \quad \text{—(ii)}$$

Find (i) m_X, m_Y & (ii) $\rho(X,Y)$.

Sol. Intersection of (i) & (ii), $x=3,\ y=\frac{1}{2}$

$$\Rightarrow m_X=3,\quad m_Y=\frac{1}{2}$$

(i) $\Rightarrow$ $6y-3=3-x$

$\Rightarrow$ $y-\frac{1}{2}=-\frac{1}{6}(x-3)$: Reg line of Y on X

$$\Rightarrow \rho\frac{\sigma_Y}{\sigma_X}=-\frac{1}{6}$$

(ii) $\Rightarrow$ $x-3=-\frac{2}{3}\left(y-\frac{1}{2}\right)$

$$\Rightarrow \rho\frac{\sigma_X}{\sigma_Y}=-\frac{2}{3}$$

$$\Rightarrow \rho^2=\frac{1}{9}$$

$$\Rightarrow \rho=-\frac{1}{3}$$ [since here, $\rho<0$]

$$S = E\left\{\left(Y - c_0^* - c_1^* X\right)^2\right\}$$

$$= E\left\{\left(Y - m_Y + \rho\frac{\sigma_Y}{\sigma_X} m_X - \rho\frac{\sigma_Y}{\sigma_X} X\right)^2\right\}$$

$$= E\left\{\left(Y - m_Y - \rho\frac{\sigma_Y}{\sigma_X}\left(X - m_X\right)\right)^2\right\}$$

$$= E\left\{(Y - m_Y)^2 + \rho^2\frac{\sigma_Y^2}{\sigma_X^2}(X - m_X)^2 - 2\rho\frac{\sigma_Y}{\sigma_X}(X - m_X)(Y - m_Y)\right\}$$

$$= \sigma_Y^2 + \rho^2\frac{\sigma_Y^2}{\cancel{\sigma_X^2}}\cancel{\sigma_X^2} - 2\rho\frac{\sigma_Y}{\cancel{\sigma_X}}\cdot\rho\cancel{\sigma_X}\sigma_Y$$

$$= \sigma_Y^2\left(1 - \rho^2\right), \quad \text{—— (v)}$$

Illy, $S = E\left\{\left(X - d_0^* - d_1^* Y\right)^2\right\} = \sigma_X^2\left(1 - \rho^2\right)$ —— (vi) (check!)

(v) & (vi) $\Rightarrow$ measure of dispersion is proportional to $1 - \rho^2$

$|\rho|$ is a measure of "goodness of fit"

Theorem

Let X and Y are independent random variables and $g_1(X), g_2(Y)$ are continuous functions of X and Y, respectively whose expectations exist. Then

$$E\{g_1(X)g_2(Y)\} = E\{g_1(X)\}E\{g_2(Y)\}.$$

Pf:

$$E\{g_1(X)\,g_2(Y)\} = \int_{y=-\infty}^{\infty}\int_{x=-\infty}^{\infty} g_1(x)\,g_2(y)\,\underbrace{f_{X,Y}(x,y)}\,dx\,dy$$

$$= \iint g_1(x)\,g_2(y)\,f_X(x)\,f_Y(y)\,dx\,dy$$

Since X, Y indp.

$\Rightarrow f_{X,Y}(x,y) = f_X(x)\,f_Y(y)$

$$= \int g_1(x)\,f_X(x)\,dx \int g_2(y)\,f_Y(y)\,dy$$

$$= E(g_1(X))\,E(g_2(Y)).$$

Mutually Independent Random Variables

Definition

1. The random variables X, Y, Z are mutually independent if

$$F_{X,Y,Z}(x,y,z) = F_X(x)F_Y(y)F_Z(z)$$

where $F_{X,Y,Z}(x,y,z)$ is the joint d.f. of (X,Y,Z); $F_X(x)$, $F_Y(y)$ and $F_Z(z)$ are the marginal d.f.s of X, Y and Z, respectively.

2. If (X,Y,Z) is continuous r.v., then

$$f_{X,Y,Z}(x,y,z) = f_X(x)f_Y(y)f_Z(z).$$

(using density functions)

Mutually Independent Random Variables

Theorem

If $X_1, X_2, \ldots, X_n$ are mutually independent r.v.s then

(a) any group of $m(< n)$ of these variates are mutually independent.

(b) the variates $(X_1, X_2, \ldots, X_{k_1})$, $(X_{k_1+1}, \underset{k_1+}{X_2}, \ldots, X_{k_2})$, $\ldots$, $(X_{k_m+1}, \underset{k_m+}{X_2}, \ldots, X_n)$, where $1 < k_1 < k_2 < \ldots < k_m < n$ are mutually independent.

(c) the variates $g_1(X_1, X_2, \ldots, X_{k_1})$, $g_2(X_{k_1+1}, \underset{k_1+}{X_2}, \ldots, X_{k_2})$, $\ldots$, $g_{m+1}(X_{k_m+1}, \underset{k_m+}{X_2}, \ldots, X_n)$ (where g_i's are continuous functions) are mutually independent.

Mutually Independent Random Variables

Theorem

If $X_1, X_2, \ldots, X_n$ are mutually independent then

(a) any group of $m(< n)$ of these variates are mutually independent.

(b) the variates $(X_1, X_2, \ldots, X_{k_1})$, $(X_{k_1+1}, X_2, \ldots, X_{k_2})$, …, $(X_{k_m+1}, X_2, \ldots, X_n)$, where $1 < k_1 < k_2 < \ldots < k_m < n$ are mutually independent.

(c) the variates $g_1(X_1, X_2, \ldots, X_{k_1})$, $g_2(X_{k_1+1}, X_2, \ldots, X_{k_2})$, …, $g_{m+1}(X_{k_m+1}, X_2, \ldots, X_n)$ (where g_i's are continuous functions) are mutually independent.

- For 3 mutually independent random variables X, Y, Z: (a) X, Y are independent, (b) (X, Y) and Z are independent and (c) $U = g_1(X, Y)$ and $W = g_2(Z)$ are independent.

Mutually Independent Random Variables

Theorem

If X, Y, Z are mutually independent, then

1. $E\{g_1(X,Y)g_2(Z)\} = E\{g_1(X,Y)\}E\{g_2(Z)\}$
2. $E\{g_1(X)g_2(Y)g_3(Z)\} = E\{g_1(X)\}E\{g_2(Y)\}E\{g_3(Z)\}$

Joint Characteristic Function

Definition

Let (X, Y) be a two-dimensional random variable. We define the joint characteristic function of X and Y as

$$\chi_{X,Y}(t,u) = E\{e^{itX+iuY}\}. \qquad i = \sqrt{-1}$$

- If X, Y are independent,

$$\chi_{X,Y}(t,u) = E\{e^{itX}\}E\{e^{iuY}\} = \chi_X(t)\chi_Y(u).$$

- **Theorem:** A necessary and sufficient condition that n random variables $X_1, X_2, \ldots, X_n$ are mutually independent is

$$\chi_{X_1,X_2,\ldots,X_n}(t_1,t_2,\ldots,t_n) = \chi_{X_1}(t_1)\chi_{X_2}(t_2)\ldots\chi_{X_n}(t_n).$$

(**Proof of converse is not part of the syllabus**)

Sum of n mutually independent random variables

$x_1, x_2, \ldots, x_n$

$$\frac{x_1+x_2+\cdots+x_n}{n}$$

Let $X_1, X_2, \ldots, X_n$ are n mutually independent random variables and $S_n = X_1 + X_2 + \ldots + X_n$. Then

$$\chi_{S_n}(t) = \chi_{X_1}(t)\chi_{X_2}(t)\ldots\chi_{X_n}(t).$$

$$\chi_{S_n}(t) = E\left(e^{itS_n}\right) = E\left(e^{itX_1}\, e^{itX_2} \ldots e^{itX_n}\right)$$

Since

$$= E\left(e^{itX_1}\right) E\left(e^{itX_2}\right) \ldots E\left(e^{itX_n}\right)$$

$$= \chi_{X_1}(t)\, \chi_{X_2}(t) \ldots \chi_{X_n}(t).$$

Reproductive Properties

1. Let $X_1, X_2, \ldots, X_n$ are mutually independent binomial variates having parameters $(\nu_1, p), (\nu_2, p), \ldots, (\nu_n, p)$, respectively. Then their sum $S_n = X_1 + X_2 + \ldots + X_n$ is binomial distributed with parameters (ν, p) with $\nu = \nu_1 + \nu_2 + \ldots + \nu_n$.

Pf:

$$X_1 \sim \text{binomial}(\nu_1, p) \Rightarrow \chi_{X_1}(t) = (pe^{it} + q)^{\nu_1}$$

$$X_2 \sim \text{binomial}(\nu_2, p) \Rightarrow \chi_{X_2}(t) = (pe^{it} + q)^{\nu_2}$$

$$\vdots \qquad \qquad \vdots$$

$$X_n \sim \text{binomial}(\nu_n, p) \Rightarrow \chi_{X_n}(t) = (pe^{it} + q)^{\nu_n}$$

$$\Rightarrow \chi_{S_n}(t) = \chi_{X_1} \cdots \chi_{X_n} = (pe^{it} + q)^{\nu_1 + \nu_2 + \cdots + \nu_n}$$

$$\Rightarrow S_n \sim \text{binomial}(\nu_1 + \nu_2 + \cdots + \nu_n, p).$$

Reproductive Properties

2. Let $X_1, X_2, \ldots, X_n$ are mutually independent Poisson variates having parameters $\mu_1, \mu_2, \ldots, \mu_n$, respectively. Then their sum $S_n = X_1 + X_2 + \ldots + X_n$ is Poisson-$(\mu_1 + \mu_2 + \ldots + \mu_n)$ distributed.

Pf.: Try yourself!

3. Let $X_1, X_2, \ldots, X_n$ are mutually independent Normal variates having parameters $(m_1, \sigma_1), (m_2, \sigma_2), \ldots, (m_n, \sigma_n)$, respectively. Then any linear combination $X = a_1X_1 + a_2X_2 + \ldots + a_nX_n$ is Normal-(m_X, σ_X) distributed where

$$m_X = a_1m_1 + a_2m_2 + \ldots + a_nm_n$$

and

$$\sigma_X^2 = a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \ldots + a_n^2\sigma_n^2.$$

Pf: Try yourself!

4 If $X_1, X_2, \ldots, X_n$ are mutually indp. gamma variates with parameters $l_1, l_2, \ldots, l_n$. Then $S_n = X_1 + X_2 + \cdots + X_n$ is gamma variate with paramtr $l_1 + l_2 + \cdots + l_n$.