Probability theory 1

12/11/21

Routine before the mid sem 1 :

Nov 12 (Today) – Class

Nov 16 (Tuesday) – Class [Mid sem syllabus is upto what is taught till Nov 16]

Nov 17 (Wednesday) – Tutorial (Arijit sir)

Nov 19 (Friday) — Holiday

Nov 23 (Tuesday) — Tutorial (Arijit sir)

Nov 24 (Wednesday) – Exam

$\forall\ i, j$ assign $\sum_{i,j} P(X = x_i, Y = y_j) = 1$

X' , Y' $\in \Omega : \to \mathbb{R}$.

$$\boxed{P(X = x_i) = \sum_j P\underbrace{(X = x_i, Y = y_j)}}$$

$$\geq P(X = x_i, Y = y_{j_0})$$

$$P\left((X' = x_i) \cap (Y' = y_j)\right)$$

$\cdot\ X \perp Y$, $E(X), E(Y) < \infty \Rightarrow E(XY) \neq \infty$

$E(|XY|) < \infty$

$E(XY) < \infty$ & $X \perp Y$

$E(X), E(Y) < \infty$.

$X \perp Y \Rightarrow |X| \perp |Y|$.

$f(X) \perp g(Y)$

$E(X) < \infty \not\Rightarrow E(X^2) = \infty$ $\qquad X = n \quad w.p. \ \frac{c}{n^3}$

$E(|X|) < \infty \quad E(|Y|) < \infty$ $\qquad Y = n \quad w.p. \ \frac{c}{n^3}$.

$$E(|XY|) = E(|X||Y|) = \sum_{m, n \geq 1} |x_m||y_n| \, P(X = x_m, Y = y_n)$$

$$= \sum_m \sum_n |x_m||y_n| \, P(X = x_m) P(Y = y_n)$$

$$= \sum_m |x_m| P(X = x_m) \sum_n |y_n| P(Y = y_n).$$

- $E(XY) < \infty \overset{?}{\Rightarrow} E(X), E(Y) < \infty$. [$X \sim$ Cauchy $\quad f(x) = \frac{1}{\pi} \frac{1}{1+x^2}$] Take X & $\frac{1}{X}$.

no

Q: If $E(XY) < \infty \Rightarrow E(X) < \infty$?

Counterexample 1:

$$T = n \quad \text{w.p.} \quad \frac{c}{n^3} \quad \forall n \geq 1$$

$$\frac{1}{T} = \frac{1}{n} \quad \text{w.p.} \quad \frac{c}{n^3} \quad \forall n \geq 1$$

Take $X = T^2$ and $Y = \frac{1}{T}$

Then $E(XY) = E\left(T^2 \cdot \frac{1}{T}\right) = E(T) = \sum_n n \frac{c}{n^3} < \infty$.

and $E\left(\frac{1}{T}\right) < \infty$ but $E(X) = E(T^2) = \sum_n n^2 \frac{c}{n^3} = \infty$.

Counter example 2:

$X =$ any random variable taking values from $\{1, 2, 3, \cdots\}$

such that $E(X)$ does not exists. [T^2 is such a random variable

previous example.]

$$Y = \mathbb{1}(X = 1)$$

$$X\mathbb{1}(X=1) = 1 \text{ w.p. } p_1$$
$$= 0 \text{ w.p. } 1-p_1$$

Then $E(XY) = E\left(X\mathbb{1}(X=1)\right)$

$$= E\left(\mathbb{1}(X=1)\right) = p_1 < \infty,$$

but $E(X)$ does not exist.

<u>Remark:</u> If X and Y are independent random variables such that $E(XY)$ exists finitely (i.e. $E(|XY|)<\infty$) then both $E(X)$ and $E(Y)$ exists finitely.

<u>Theorem:</u> If $X, Y : \Omega \to \mathbb{R}$ be two random variables & $f, g : \mathbb{R} \to \mathbb{R}$. If X and Y are independent then so is $f(X)$ and $g(Y)$.

<u>Pf:</u> $$\mathbb{P}(f(X)=x, g(Y)=y) = \mathbb{P}\Big(X \in \underbrace{f^{-1}(x)}_{= \{x_{i_1}, x_{i_2}, \ldots\}},\ Y \in g^{-1}(y) \Big)$$ [note that $f^{-1}(x)$ & $g^{-1}(y)$ are subsets of $\mathbb{R}$]

$$= \mathbb{P}(X \in f^{-1}(x))\, \mathbb{P}(Y \in g^{-1}(y))$$

$$= \mathbb{P}(f(X)=x)\ \mathbb{P}(g(Y)=y) \quad . \qquad \square .$$

<u>Corollary</u>: X and Y independent $\Rightarrow$ $|X|$ and $|Y|$ are independent.

<u>Theorem</u>: If A and B are two independent events then $\mathbb{1}_A$ and $\mathbb{1}_B$ are independent.

<u>Proof</u>: $\mathbb{P}(\mathbb{1}_A = 0,\ \mathbb{1}_B = 0) = \mathbb{P}(A^c \cap B^c) = \mathbb{P}(A^c)\,\mathbb{P}(B^c) = \mathbb{P}(\mathbb{1}_A = 0)\,\mathbb{P}(\mathbb{1}_B = 0)$

$\mathbb{P}(\mathbb{1}_A = 0,\ \mathbb{1}_B = 1) = \mathbb{P}(A^c \cap B)$ - - - -

(fill up)

$\mathbb{P}(\mathbb{1}_A = 1,\ \mathbb{1}_B = 0) = \mathbb{P}(A \cap B^c)$ - - - .

$\mathbb{P}(\mathbb{1}_A = 1,\ \mathbb{1}_B = 1) = \mathbb{P}(A \cap B)$ - - - .

Definition (Covariance):

For two random variables X and Y with expectations μ_X and μ_Y respectively, the covariance between X and Y is

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

whenever the expectation on the RHS exists finitely.

Theorem:

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

2) $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$ whenever RHS exists.

3) $\mathrm{Cov}(aX+c, bY+d) = ab\,\mathrm{Cov}(X,Y)$

$\mathrm{Var}(aX+b) = a^2\,\mathrm{Var}(X)$.

4) $\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) + 2\,\text{Cov}(X,Y)$ whenever $\text{Var}(X), \text{Var}(Y) < \infty$.

X & Y independent $\Rightarrow$ $\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y)$.

5) $\text{Cov}(aX, bY+cZ) = ab\,\text{Cov}(X,Y) + ac\,\text{Cov}(X,Z)$.

Pf: H.W. (follows from definition).

Exercise: (HW): Let X be a random variable such that $E(X^2)$ is defined. Suppose that Y takes the values ± 1 with probability $\frac{1}{2}$, independently of X. Define $Z = XY$. Calculate $\text{Cov}(X,Z)$. Assume furthermore that $|X|$ is non-degenerate. Are X and Z independent?

Theorem: Show that for random variables $X_1, X_2, \dots, X_m, Y_1, Y_2, \dots, Y_n$ and constants $\alpha_1, \dots, \alpha_m, \beta_1, \dots, \beta_n$,

$$\mathrm{Cov}\left(\sum_{i=1}^{m} \alpha_i X_i, \sum_{j=1}^{n} \beta_j Y_j\right) = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_i \beta_j \mathrm{Cov}(X_i, Y_j)$$

Proof: **Step 1:** Prove the part 5 of "previous theorem" — *

Step 2: $\mathrm{Cov}\left(\alpha_i X_i, \sum_{j=1}^{m} \beta_j Y_j\right) = \sum_{j=1}^{m} \alpha_i \beta_j \mathrm{Cov}(X_i, Y_j)$ for any fixed i.

(induction on m)

Step 3: $\mathrm{Cov}\left(\sum_{i=1}^{n} \alpha_i X_i, \sum_{j=1}^{m} \beta_j Y_j\right)$

$$= \sum_{i=1}^{n} \alpha_i \operatorname{Cov}\left(X_i, \sum_{j=1}^{n} \beta_j Y_j \right) \quad (\text{using step } 2)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_i \beta_j \operatorname{Cov}\left(X_i, Y_j \right) \quad (\text{using step} 2)$$

One fact is needed in this proof. $\left(\operatorname{Cov}(X,Y) = \operatorname{Cov}(Y, X)\right)$. $\square$.

<u>Corollary:</u> Let $X_1, \ldots, X_n$ be random variables each having a variance. Then,

$$\operatorname{Var}\left(\sum_{i=1}^{n} X_i \right) = \sum_{i=1}^{n} \operatorname{Var}(X_i) + 2 \sum_{1 \le i < j \le n} \operatorname{Cov}(X_i, X_j)$$

<u>Pf:</u> Follows trivially from previous theorem. (choose $Y_j = X_j$).

Corollary: If for every $1 \le i < j \le n$, X_i and X_j are independent. Then,

$$\operatorname{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \operatorname{Var}(X_i).$$

Proof: Part 4 of the "previous theorem" ✪ and previous corollary gives us the result.

Remark: It follows from the Cauchy Schwartz inequality that If X and Y have variances then $\operatorname{Cov}(X,Y)$ is defined.

$$|\operatorname{Cov}(X,Y)| = \left|E\left[(X-\mu_X)(Y-\mu_Y)\right]\right| \le \sqrt{E\left((X-\mu_X)^2\right)E\left((Y-\mu_Y)^2\right)} = \sqrt{\operatorname{Var}(X)\operatorname{Var}(Y)}. \quad \square$$

①

Definition: (Correlation coefficient)

Let X and Y be non degenerate random variables having variances.

Then the correlation coefficient between X and Y is defined as

$$\operatorname{Corr}(X,Y) := \frac{\operatorname{Cov}(X,Y)}{\sqrt{\operatorname{Var}(X)\operatorname{Var}(Y)}} .$$

Corollary: $|\operatorname{Corr}(X,Y)| \leq 1$.

Pf: Follows form the inequality ①.

Definition! (Non-degenerate random variable)

X is non-degenerate random variable if it takes at least 2 values with strictly positive probabilities. (in case of discrete random variables).

OR

In general, a random variable is degenerate if $\exists$ a $c \in \mathbb{R}$ such that

$$\mathbb{P}(X=c) = 1.$$

A random variable is non-degenerate if it is not degenerate.

If $X \sim \text{Bern}(1)$

$$X = \begin{cases} 1 & \text{w.p. } 1 \\ 0 & \text{w.p. } 0 \end{cases}$$

$$\frac{\overset{n}{\cancel{\log n}}}{n} \qquad \frac{\#\text{ occurrences of ''0''}}{\#\text{ word}} \to 0.$$

Theorem: If $X \geq 0$ and $E(X) = 0$. Then $P(X=0) = 1$.

(In other words we call $X=0$ almost surely).

Proof:

$$0 = E(X) = \sum_{n \geq 1}^{\infty} x_n p_n \quad \left(\Rightarrow x_n p_n = 0 \quad \forall \; n \geq 1. \right)$$

$$= \sum_{\substack{n \geq 1: \\ \{p_n = 0\}}} x_n p_n + \sum_{\substack{n \geq 1 \\ \{p_n > 0\}}} x_n p_n$$

$$= \sum_{n : p_n > 0} x_n p_n \quad \Rightarrow \quad x_n p_n = 0 \quad \forall \{n : p_n > 0\}$$

$$\Rightarrow x_n = 0 \quad \forall \{n : p_n > 0\}$$

$$P(X=0) = \sum_{n: p_n > 0} P(X = x_n) = \sum_{n: p_n > 0} p_n = 1 \qquad \square$$

<u>Theorem:</u> If $\mathrm{Corr}(X,Y) = 1$, then

$$Y = E(Y) + \sqrt{\frac{\mathrm{Var}(Y)}{\mathrm{Var}(X)}}\,(X - E(X)) \text{ almost surely.} \quad \left[\frac{Y - E(Y)}{X - E(X)} = \sqrt{\frac{\mathrm{Var}(Y)}{\mathrm{Var}(X)}}\right]$$

If $\mathrm{Corr}(X,Y) = -1$, then

$$Y = E(Y) - \sqrt{\frac{\mathrm{Var}(Y)}{\mathrm{Var}(X)}}\,(X - E(X)) \text{ almost surely.}$$

Proof: Suppose $\mathrm{Corr}(X,Y)=1 \Rightarrow \mathrm{Cov}(X,Y)=\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}$

$X=Y$. a.s.
a.e.

$X-Y=0$ a.s.

$P(X-Y=0)=1$
$\parallel$
$P(X=Y)=1$

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

$$=E\left(\left(Y-E(Y)\right)^2\right)-2\sqrt{\frac{\mathrm{Var}(Y)}{\mathrm{Var}(X)}}E\left((Y-E(Y))(X-E(X))\right)+E\left((X-E(X))^2\right)\frac{\mathrm{Var}(Y)}{\mathrm{Var}(X)}$$

$$=\mathrm{Var}(Y)-2\sqrt{\frac{\mathrm{Var}(Y)}{\mathrm{Var}(X)}}\mathrm{Cov}(X,Y)+\mathrm{Var}(Y)$$

$$=2\mathrm{Var}(Y)-2\sqrt{\frac{\mathrm{Var}(Y)}{\mathrm{Var}(X)}}\sqrt{\mathrm{Var}(Y)\mathrm{Var}(X)}=0.$$

Hence using the previous theorem we get,

$$\left[Y - E(Y) - \sqrt{\frac{Var(Y)}{Var(X)}}\,(X - E(X))\right]^2 = 0 \quad \text{almost surely.}$$

Similarly the other case (when $Corr(X,Y) = -1$) can be shown. (HW).

__Exercise:__ If X and Y have variances σ_X^2 and σ_Y^2 respectively. Then

$$(\sigma_X - \sigma_Y)^2 \le Var(X+Y) \le (\sigma_X + \sigma_Y)^2$$

__Pf:__ (HW). __Hint:__ $|Corr(X,Y)| \le 1$.