Probability-2 Lecture-1

08 January 2024 11:11

X - discrete r.v.

D_X - set of values X can take.

$p(x) = P(X=x)$, $x \in D_X$

Expectation: $E(X) := \sum_{x \in D_X} x\, p(x)$, provided $\sum x\, p(x)$ is "well-defined".

($\Leftrightarrow \sum x^{+} \cdot p(x)$ & $\sum x^{-} \cdot p(x)$ are not both infinite)

$E(X)$ is finite $\Leftrightarrow \sum x^{+} \cdot p(x)$ & $\sum x^{-} \cdot p(x)$ are both finite.

$\Leftrightarrow \sum |x| \cdot p(x)$ is finite.

Variance:

Suppose $E(X)$ is finite.

Define $V(X) := E(X - E(X))^2$

By definition, if $E(X)$ is finite, then $V(X)$ is always defined, and $0 \leq V(X) < \infty$.

Some facts:

① $V(X) < \infty \Leftrightarrow E(X^2) < \infty$.

$$\sqrt{E(X - E(X))^2} \leq \sqrt{E(X^2)} + |E(X)|.$$

[Minkowski's inequality:

$p \geq 1$. X, Y - random variables.

$$\left(E(|X+Y|^p)\right)^{1/p} \leq \left(E(|X|^p)\right)^{1/p} + \left(E(|Y|^p)\right)^{1/p}$$

Proof: for $p=1$, nothing to prove.

Assume $p > 1$.

$E(|X+Y|^p) \leq E(|X| \cdot (|X+Y|^{p-1}) + E(|Y|)(|X+Y|^{p-1})$

Assume $p>1$.

$$E(|X+Y|^p) \le E(|X|)\cdot(|X+Y|^{p-1}) + E(|Y|)(|X+Y|^{p-1})$$

[Δ ineq., linearity of expectation]

$$E(|X|)\cdot(|X+Y|^{p-1}) \le E(|X|^p)^{1/p} \cdot E\left(|X+Y|^{(p-1)q}\right)^{\frac{1}{q}}$$

p

[Holder's inequality]

$$= E(|X|^p)^{1/p} \cdot E(|X+Y|^p)^{1/q}.$$

$$\le E(|X|^p)^{1/p} + E(|Y|^p)^{1/p}.$$

$\frac{1}{p}+\frac{1}{q}=1$

$\Rightarrow q = \frac{p}{(p-1)}$

$\Rightarrow (p-1)\cdot q = p$.

[in case of finite variance, we can simply use algebra to get the inequality, i.e., $\sqrt{a^2-b^2} \le |a|+|b|$.]

② $V(X) = \min\limits_{\alpha} \left(E(X-\alpha)^2\right)$

Idea: suppose we wish to "guess" the value of a r.v. by just some constant predictor. Then, there will be some difference between the actual value & predicted value. The idea is to minimize the difference.

"Expected value, or expectation is the best predictor in respect of minimizing the mean square error."

Proof: WLOG, assume, $E(X^2)<\infty$ [Why? Because, otherwise both sides are $+\infty$]

$$E(X-\alpha)^2 = E(X-E(X)+E(X)-\alpha)^2$$
$$= E(X-E(X))^2 + E(\underbrace{E(X)}_{\text{constant}}-\alpha)^2 - 2\cdot \overset{0}{E(X-E(X))}\cdot E(E(X)-\alpha)$$

$$= E(X-E(X))^2 + (E(X)-\alpha)^2$$

$\because V(X) \geq 0.$

$$V(X) = 0 \iff E(X-E(X))^2 = 0$$
$$\iff P(X - E(X) = 0) = 1.$$
$$\iff P(X = E(X)) = 1$$

③ $V(cX) = c^2 \cdot V(X)$

④ $V(X+Y) = E(X+Y-E(X)-E(Y))^2$
$$= V(X) + V(Y) + 2\underbrace{E\left((X-E(X))\cdot(Y-E(Y))\right)}_{\substack{Cov(X,Y). \\ (\text{covariance})}}$$

Defn: X, Y - random variables.
$E(X^2) < \infty, \quad E(Y^2) < \infty.$

Then, Covariance,
$$Cov(X,Y) := E\left((X-E(X))\cdot(Y-E(Y))\right)$$

Q: will this be finite?

$$\because \left|E(X-E(X))(Y-E(Y))\right| \leq \sqrt{\underbrace{E(X-E(X))^2}_{<\infty}} + \sqrt{\underbrace{E(Y-E(Y))^2}_{<\infty}}$$

$$< \infty$$

[By Cauchy Schwarz inequality].

⑤ $|Cov(X,Y)| \leq \sqrt{V(X)} \cdot \sqrt{V(Y)}$

$$\Rightarrow \left|\frac{Cov(X,Y)}{\sqrt{V(X)}\cdot\sqrt{V(Y)}}\right| \leq 1$$

$$\left| \sqrt{V(X)} \cdot \sqrt{V(Y)} \right|$$

$\hookrightarrow$ $Corr(X,Y)$ - "correlation coefficient between X and Y."

Note:

$\therefore$ $V(X+Y) = V(X) + V(Y)$

$\iff Cov(X,Y) = 0$.

Definition: (Independence of Random Variables).

Random variables X, Y are said to be independent

if $P(X=x, Y=y) = P(X=x) \cdot P(Y=y)$

$\forall\ x \in X,\ y \in Y$.

Generalisation (this holds for continuous X & Y as well)

$\iff P(X \in B_1,\ Y \in B_2) = P(X \in B_1) \cdot P(Y \in B_2)$.

(Exc: prove this equivalence for discrete X & Y).

Fact 1:

X, Y are independent $\Rightarrow$ $h(X)$, $g(Y)$ are independent

2 functions.

$$P(h(X)=\alpha,\ g(Y)=\beta) = P(X \in B_1,\ Y \in B_2) = P(X \in B_1) \cdot P(Y \in B_2)$$

where $B_1 = \{x : h(x) = \alpha\}$

$B_2 = \{y : g(y) = \beta\}$.

Very important fact:

$\left.\begin{array}{l} X, Y \text{ - independent,} \\ E(X), E(Y) \text{ finite} \end{array}\right\} \Rightarrow E(XY)$ is finite.

& $E(XY) = E(X) \cdot E(Y)$

$E(X), E(Y)$ finite |
& $E(XY) = E(X) \cdot E(Y)$

Eg. $\Omega = \{1, 2, \ldots\}$.
$p(n) = \left(\frac{5}{6^n}\right)$. X & Y are defined as
$X(n) := 3^n$, $Y(n) := 2^n$

Here, $E(X) = \sum \left(\frac{5}{6^n}\right) \cdot 3^n < \infty$, $E(Y) = \sum \left(\frac{5}{6^n}\right) \cdot 2^n < \infty$

But, $E(XY) = \sum \left(\frac{5}{6^n}\right) \cdot 3^n \cdot 2^n = \sum 5 \to \infty$.
(NOT finite)

- $E(XY) = \sum xy \cdot P(X = x, Y = y) = \sum_x x \cdot P(X=x) \sum_y y \cdot P(Y=y)$
$= E(X) \cdot E(Y)$ (for X, Y ind)

Note: $\because X, Y$ independent
$\Rightarrow X - E(X), Y - E(Y)$ are independent

Now, define $h(X) = X - E(X)$, $g(Y) = Y - E(Y)$.
then, $Cov(X, Y) = 0$.

But, the converse is not true.
ie, X, Y independent $\Rightarrow Cov(X, Y) = 0$
BUT, $Cov(X, Y) = 0 \not\Rightarrow X, Y$ independent.
(example given below).

- X, Y - independent.
$\Rightarrow Cov(X, Y) = 0 \Rightarrow V(X+Y) = V(X) + V(Y)$

Eg.
$\Omega = \{1, 2, 3\}$
$p(1) = \frac{1}{2}$, $p(2) = p(3) = \frac{1}{4}$
$X(1) = -1$, $X(2) = X(3) = 1$
$Y(1) = 0$, $Y(2) = 1$, $Y(3) = -1$

$P(X = -1) = \frac{1}{2}$
$P(Y = 0) = \frac{1}{2}$
$P(X = -1,) = \ldots$

$X(1) = 1, \quad X(2) = X(3) = -1$

$Y(1) = 0, \quad Y(2) = 1, \quad Y(3) = -1$

$P(Y=0) = \frac{1}{2}$

Here, $Cov(X,Y) = 0$, X & Y are <u>not</u> independent. $P(X=-1, Y=0) = \frac{1}{2} \neq P(X=-1) \cdot P(Y=0) = \frac{1}{4}$

this example illustrates that,

$Cov(X,Y) = 0$ is a <u>necessary condition</u> for independence of X and Y, but <u>NOT sufficient</u>.

<u>Properties of Covariance:</u>

- $Cov(X,Y) = E\big((X - E(X)) \cdot (Y - E(Y))\big)$
  $= E(XY) - E(X) \cdot E(Y)$

- $Cov(\alpha_1 X_1 + \alpha_2 X_2, \beta_1 Y_1 + \beta_2 Y_2) =$
  $\alpha_1 \beta_1 \, cov(X_1, Y_1) + \alpha_1 \beta_2 \cdot cov(X_1, Y_2) + \alpha_2 \beta_1 \cdot Cov(X_2, Y_1) + \alpha_2 \beta_2 \cdot Cov(X_2, Y_2)$

- $Cov(X,X) = V(X)$.

- For X, Y - r.v with <u>finite 2^{nd} Moments</u>.
  define $\langle X, Y \rangle = E(XY)$.
  $\langle \cdot, \cdot \rangle$ is "<u>symmetric</u>, <u>bilinear</u> and
  $\langle X, X \rangle \geq 0$ and $\langle X, X \rangle = 0 \Leftrightarrow P(X = 0) = 1$