# Probability theory 1

9/11/21

- We have known joint distribution of two random variables, the notion of independence between two random variables.

$X, Y \sim$ random variables. They are independent if

$$P(X = x_i, Y = y_j) = P(X = x_i)\, P(Y = y_j) \quad \text{for any } i \text{ and } j$$

where $X \in \{x_1, x_2, \dots\}$ and $Y \in \{y_1, y_2, \dots\}$.

- We saw that if $X$ and $Y$ are independent then

$$E(XY) = E(X)\, E(Y).$$

Theorem: If $X$ and $Y$ are independent then for any $A \subseteq \{x_1, x_2, \cdots\}$ and $B \subseteq \{y_1, y_2, \cdots\}$,

$$P(X \in A, Y \in B) = P(X \in A)\, P(Y \in B),$$

where $X \in \{x_1, x_2, \cdots\}$ and $Y \in \{y_1, y_2, \cdots\}$.

Proof:
$$P(X \in A, Y \in B) = \sum_{x_i \in A} \sum_{y_j \in B} P(X = x_i, Y = y_j)$$

$$= \sum_{x_i \in A} \sum_{y_j \in B} P(X = x_i)\, P(Y = y_j) \quad \text{(by independence)}$$

$$= \sum_{x_i \in A} P(X = x_i) \sum_{y_j \in B} P(Y = y_j) = P(X \in A)\, P(Y \in B). \quad \square$$

Example: Suppose that the random variables $X$ and $Y$ follow Binomial $(m, p)$ and Binomial $(n, p)$ respectively, independent of each other. Then what is the PMF of $X+Y$?

$\Rightarrow X \in \{0, 1, \ldots, m\}$, $Y \in \{0, 1, \ldots, n\}$. So, $X+Y \in \{0, 1, \ldots, m+n\}$. For $k \in \{0, 1, \ldots, m+n\}$,

$$P(X+Y=k) = \sum_{i=0}^{k} P(X=i, Y=k-i) = \sum_{i=0}^{k} P(X=i)\, P(Y=k-i) \quad \text{(by previous theorem)}$$

$$= \sum_{i=0}^{k} \binom{m}{i} p^{i} (1-p)^{m-i} \binom{n}{k-i} p^{k-i} (1-p)^{n-k+i}$$

$$= p^{k} (1-p)^{m+n-k} \sum_{i=0}^{k} \binom{m}{i} \binom{n}{k-i}$$

$$= \underbrace{\binom{n+m}{k}} p^k (1-p)^{m+n-k} \qquad \text{(by Vandermonde's identity).}$$

Hence, $X+Y \sim \text{Binomial}(m+n, p)$.

<u>Theorem (Cauchy-Schwartz inequality):</u>

If $X$ and $Y$ are random variables such that $E(X^2)$ and $E(Y^2)$ are defined, then so is $E(XY)$ and

$$|E(XY)| \le \sqrt{E(X^2)\,E(Y^2)} < \infty \quad \text{—— ①}$$

In particular $E(|XY|) \le \sqrt{E(|X|^2)\,E(|Y|^2)}$. —— ②

① follows from ② since $|E(XY)| \le E(|XY|)$. $\left( -XY \le XY \le XY \right)$ $XY \le |XY|$

$E(X)$
$= \sum x_n p_n$
$E(XY)$
$= \sum r_n P(XY = r_n)$

$= \sum_{\omega \in \Omega} XY(\omega) P(\{\omega\})$

Proof: Let $\Omega = \{\omega_1, \omega_2, \dots\}$. Fix $n \geq 1$ and $\lambda \in \mathbb{R}$ and observe that

$$0 \leq \sum_{i=1}^{n} \left( |X(\omega_i)| - \lambda |Y(\omega_i)| \right)^2 P(\{\omega_i\})$$

for a fixed $n \in \mathbb{N}$

$$= \lambda^2 \underbrace{\sum_{i=1}^{n} (Y(\omega_i))^2 P(\{\omega_i\})}_{a} - 2\lambda \underbrace{\sum_{i=1}^{n} |X(\omega_i) Y(\omega_i)| P(\{\omega_i\})}_{b/2} + \underbrace{\sum_{i=1}^{n} (X(\omega_i))^2 P(\{\omega_i\})}_{c}$$

Recall that if for some $a, b, c \in \mathbb{R}$

$a\lambda^2 + b\lambda + c \geq 0$ for all $\lambda \in \mathbb{R}$ then $b^2 - 4ac \leq 0$, i.e.

$$4 \left( \sum_{i=1}^{n} |X(\omega_i) Y(\omega_i)| P(\{\omega_i\}) \right)^2 \leq 4 \left( \sum_{i=1}^{n} (X(\omega_i))^2 P(\{\omega_i\}) \right) \left( \sum_{i=1}^{n} (Y(\omega_i))^2 P(\{\omega_i\}) \right)$$

$$\Rightarrow \left( \sum_{i=1}^{n} |X(w_i) Y(w_i)| P(\{w_i\}) \right)^2 \leq \left( \sum_{i=1}^{\infty} (X(w_i))^2 P(\{w_i\}) \right) \left( \sum_{i=1}^{\infty} (Y(w_i))^2 P(\{w_i\}) \right)$$

$$= E(X^2)\, E(Y^2)$$

We have prove that for every $n \geq 1$,

$$\sum_{i=1}^{n} |X(w_i) Y(w_i)| P(\{w_i\}) \leq \sqrt{E(X^2)\, E(Y^2)}$$

and hence, letting $n \to \infty$

$$\sum_{i=1}^{\infty} |X(w_i) Y(w_i)| P(\{w_i\}) \leq \sqrt{E(X^2)\, E(Y^2)} \Rightarrow \underline{E(|XY|) \leq \sqrt{E(X^2)\, E(Y^2)}}.$$

Equation ② is proved.

Thus $E(XY)$ is defined.

Furthermore, for every $\lambda \in \mathbb{R}$

$$0 \leq E\left[(X-\lambda Y)^2\right] = E(X^2) - 2\lambda E(XY) + \lambda^2 E(Y^2)$$

$$0 \leq E(X^2) - 2\lambda E(XY) + \lambda^2 E(Y^2)$$

Therefore, $$4\,(E(XY))^2 \leq 4\,E(X^2)\,E(Y^2)$$

$$\Rightarrow (E(XY))^2 \leq E(X^2)\,E(Y^2)$$

$$\Rightarrow |E(XY)| \leq \sqrt{E(X^2)\,E(Y^2)} \qquad \square$$

**Corollary:** If $E(X^2)$ is defined, then so is $E(X)$ and

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

**Pf:** Choose $Y \equiv 1$ on the Cauchy-Schwartz inequality.

## Covariance:

**Definition:** For two random variables $X$ and $Y$ with expectations $\mu_X$ and $\mu_Y$ respectively, then covariance between $X$ and $Y$ is

$$\mathrm{Cov}(X,Y) := E\Big((X-\mu_X)(Y-\mu_Y)\Big)$$

which is defined whenever the expectation on the right hand side is finite.

**Theorem:** Let $X$ and $Y$ be random variables whose covariance is defined.

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

2. If $X$ has finite variance, then $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$.

3. For $a, b, c, d \in \mathbb{R}$, $\mathrm{Cov}(aX+b, cY+d) = ac \, \mathrm{cov}(X, Y)$.

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

4. If variances of $X$ and $Y$ are defined, then

   $$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{cov}(X, Y).$$

5. If $X$ and $Y$ are independent, then,

   $\mathrm{Cov}(X, Y) = 0$. (also, $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$)

Proofs!

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

$$= E(XY) - \mu_Y\mu_X - \mu_X\mu_Y + \mu_X\mu_Y = E(XY) - E(X)E(Y)$$

② $\mathrm{Cov}(X,X) = E(X^2) - E(X)E(X) = \mathrm{Var}(X)$ (we used ①).

③ $\mathrm{Cov}(aX+b, cY+d) = E\Big(\big(aX+b - E(aX+b)\big)\big(cY+d - E(cY+d)\big)\Big)$

$$= E\Big(\big(aX - aE(X)\big)\big(cY - cE(Y)\big)\Big)$$

$$= E\big(acXY - acXE(Y) - acY\,E(X) + ac\,E(X)E(Y)\big)$$

$$= ac\big(E(XY) - E(X)E(Y) - E(Y)E(X) + E(X)E(Y)\big)$$

$$= ac\big(E(XY) - E(X)E(Y)\big) = ac\,\mathrm{Cov}(X,Y).$$

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

④ $\mathrm{Var}(X+Y) = \mathrm{Cov}(X+Y, X+Y) = E\left((X+Y)^2\right) - \left(E(X+Y)\right)^2$

$$= E(X^2) + 2E(XY) + E(Y^2)$$

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

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

⑤ $\mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y) = E(X)E(Y) - E(X)E(Y) = 0$

if $X$ and $Y$ are independent.

Q: $E(X) < \infty$, $E(Y) < \infty \Rightarrow E(XY) < \infty$? True whenever $E((X-\mu_X)(Y-\mu_Y)) < \infty$.

Remark: If $E(X) < \infty$ then it does not imply that $X < \infty$.

Example: Choose $X = n$ w.p. $\frac{c}{n^r}$ for any $r > 2$.

$$E(X) = \sum_{n=1}^{\infty} \frac{c}{n^{r-1}} < \infty \quad \text{for any } r > 2.$$

But $X$ is not bounded.