Probability theory 1

5/11/21

- We have been introduced with random variables, PMFs, CDFs, variances, expectations.
- Today we proceed with a particulad kind of random variables which can be writtean as a sum of indicator random variables and compute its expectation.

$$P(A) = \sum_{i=1}^{\infty} P(A \mid A_i) P(A_i)$$

- Definition (Indicator function):

The indicator random variable of a subset A of a sample space

Ω is a function

$$\mathbf{1}_A \equiv \mathbb{1}_A(\cdot) \equiv \mathbb{1}(A)(\cdot) : \Omega \to \{0,1\}$$

$$\mathbb{1}(A)(\omega) \in \{0,1\}$$
$$\omega \in \Omega$$

definded as $\quad \mathbb{1}(A)(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A \end{cases}$.

Observe that for an indicator random variable ($\mathbb{1}_A$, say) its expectation is

$$E(\mathbb{1}_A) = 1 \times P(\mathbb{1}_A = 1) + 0 \times \mathbb{P}(\mathbb{1}_A = 0)$$
$$= P(\omega \in A) = P(A).$$

Hence, $\boxed{E(\mathbb{1}_A) = P(A)}$

Example: Let $X \sim$ the number of empty boxes when n distinct balls are thrown at random in k boxes. What is $E(X)$?

$\mathbb{1}(A) = \mathbb{1}_A$

Ans: $$X = \sum_{i=1}^{k} \mathbb{1}(\text{the ith box is empty})$$

where the boxes are numbered from $1, 2, \ldots, k$.

So, $$E(X) = \sum_{i=1}^{k} E\left(\mathbb{1}(\text{the ith box is empty})\right) = \sum_{i=1}^{k} P(\text{the ith box is empty})$$

$$= k\, P(\text{the 1st box is empty}) = k \frac{(k-1)^n}{k^n} = k^{1-n}(k-1)^n.$$

$$\binom{n+k-1}{k-1}$$

Example: Let $X \sim$ the number of letters that come between A and B when the four letters A, B, C, D are randomly permuted. What is $E(X)$?

Ans: Observe that,

$$X = \sum_{\alpha \in \{A,B,C,D\} \setminus \{A,B\}} \mathbb{1}\Big(\alpha \text{ (is a letter) comes in between A and B)} \Big)$$

So, $$E(X) = \sum_{\alpha \in \{C,D\}} P(\alpha \text{ comes in between A and B})$$

$$= \frac{2}{4} \times P(\alpha \overset{(=C)}{} \text{ comes in between A and B})$$

$\alpha \in Z \setminus \{A, B\}$.

$E(X) = (k-2) \times P(\alpha \cdots)$

$$= 2\,\frac{\#(\alpha \text{ comes in between A and B})}{\#(A, B, C, D \text{ can be permuted})} = 2\,\frac{?}{4!} = 2\cdot\frac{\frac{4!}{3}}{4!} = \frac{2}{3}. \quad \square \qquad (8)$$

$\#(\alpha$ comes in between A and B when $n (\geq 3)$ letters are permuted$) = \dfrac{n!}{3}$

For $n = 3$ the statement is verified.

Suppose the statement is true for $n = k$.

Then, $\#(\alpha \text{ comes in between A \& B}) = \dfrac{k!}{3}$.

$\underline{A, B, C}$ $\}$ $\underline{n = 3}$

ABC, $\underline{ACB}$, CAB

BAC, $\underline{BCA}$, CBA

A B

$$\frac{3!}{3} = \frac{6}{3} = 2$$

A B C

D

When $n = k+1$

$x_1 \quad x_2 \; \cdot \; A \cdots - \cdot \; B \; - \cdots \; x_{k-2}$

x_{k-1}

$$\#(\alpha \text{ comes in between } A \& B) = (k+1)\frac{k!}{3}$$

$$= \frac{(k+1)!}{3}$$

ABC, ACB, CAB

BAC, BCA, CBA

$\frac{3!}{3}$

$\Rightarrow$ D $\xrightarrow{\text{1st place}}$

DABC, DACB, DCAB

DBAC, DBCA, DCBA

$\xrightarrow{\text{2nd place}}$ ADBC, - , ..

$$4.2 = \frac{4.3.2}{3} = \frac{4!}{3}$$

$$E(X) = (k-2)\cdot\frac{1}{3}$$ when $X \sim$ the number of letters that comes in between A & B when k many letters (including A, B) are permutted.

$$= (k-2)\,\frac{\frac{k!}{3}}{k!}$$

Joint distributions:

Definition: For two random variables X and Y taking values $x_1, x_2, \cdots$ and $y_1, y_2, \cdots$, respectively, their joint PMF is defined by

$$p(x_i, y_j) := \mathbb{P}\left(X = x_i, Y = y_j\right) \text{ for any } i, j \in \mathbb{N}.$$

$$= \mathbb{P}\left(\{X = x_i\} \cap \{Y = y_j\}\right)$$ (wherever X & Y defined on same Ω).

Theorem: If $p(\cdot,\cdot)$ is the joint PMF of X and Y, show that

$$P(X=x_i) = \sum_j p(x_i, y_j) \quad \text{for any } x_i,$$

$(\mathbb{R}^2, +)$

$\sum (x, y_j) = (x, \sum y_j)$

and for all y_j $\quad P(Y=y_j) = \sum_i p(x_i, y_j)$

Proof:

$$\sum_j p(x_i,y_j) = \sum_j P(X=x_i, Y=y_j) = P\Big(\bigcup_j \{X=x_i, Y=y_j\}\Big)$$

$$= P\Big(\{X=x_i\}, \bigcup_j \{Y=y_j\}\Big) = P\big(\{X=x_i\}, \Omega\big)$$

$$= P(X=x_i). \quad \square$$

Example: An urn contains 10 red balls, 5 black balls, and 5 white balls. Three balls are drawn from the urn with replacement. Let X and Y denote the number of black and white balls drawn, respectively. Then,

P. Billingsley

$$P(X=x, Y=y) = \frac{3!}{x!\,y!\,(3-x-y)!}\left(\frac{5}{20}\right)^{x}\left(\frac{5}{20}\right)^{y}\left(\frac{10}{20}\right)^{3-x-y}.$$

for $x, y = 0, 1, 2, 3$ such that $x + y \leq 3$ $\square$

Definition: Random variables X and Y are independent if

$$P(X=x_i, Y=y_j) = P(X=x_i)\, P(Y=y_j) \quad \text{for all } i,j \in \mathbb{N}.$$

Example: Suppose 10 distinct letters $=\{A, B, C, \ldots, J\}$ are given and asked to form a word of 10 letters. Let $X \sim$ 2nd placed letter and $Y \sim$ 5th placed letter. Then show that X and Y are independent.

HW: $P(X=x, Y=y) = P(X=x)\, P(Y=y)$ for any $x, y \in \{A, B, \ldots, J\}$.

Theorem 1: Let X and Y be two independent random variables taking values in $\mathbb{R}$, such that $E(XY)$ exists. Then

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

[Chose $f(x,y) = xy$ in the next theorem].

Theorem 2: Under the setup of Theorem 1. Let $f: \mathbb{R}^2 \to \mathbb{R}$ be such that $f(X,Y)$ has an expectation. Then,

$$E[f(X,Y)] = \sum_i \sum_j f(x_i, y_j)\, \mathbb{P}(X = x_i)\, \mathbb{P}(Y = y_j).$$

Pf: Let X takes the values $x_1, x_2, \ldots$

Y takes the values $y_1, y_2, \ldots$

$$E\left(f(X,Y)\right) = \sum_{\omega\in\Omega} \left(f(X,Y)\right)(\omega)\ \mathbb{P}\left(\{\omega\}\right) = \sum_{\omega\in\Omega} f\left(X(\omega), Y(\omega)\right)\ \mathbb{P}\left(\{\omega\}\right)$$

$$= \sum_i \sum_j \sum_{\omega\in\Omega:\ X(\omega)=x_i,\ Y(\omega)=y_j} f(x_i, y_j)\ \mathbb{P}\left(\{\omega\}\right)$$

$$\sum_{\omega\in A} \mathbb{P}(\omega) = \mathbb{P}(A)$$

$$= \sum_i \sum_j f(x_i, y_j)\ \mathbb{P}\left(\omega : X(\omega)=x_i,\ Y(\omega)=y_j\right)$$

$$= \sum_i \sum_j f(x_i, y_j)\ \mathbb{P}\left(X=x_i, Y=y_j\right)$$

(If X & Y are not independent then we arrive upto this expression).

using independence

$$= \sum_i \sum_j f(x_i, y_j)\ \mathbb{P}\left(X=x_i\right)\ \mathbb{P}\left(Y=y_j\right).$$

HW: Find X & Y (random variables) s.t. $E(X) < \infty$, $E(Y) < \infty$ but $E(XY)$ does not exists.

Ex1: Chose $X = Y$. &

we already got examples when

$E(X^2) = \infty$, $E(X) = E(Y) < \infty$.

~~Ex2:~~ $\boxed{X \perp Y}$

$$X = n \quad \text{w.p.} \quad \frac{c}{n^3}$$

$$Y = n + \frac{1}{n} \quad \text{w.p.} \quad \frac{c}{n^3}$$

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

$$E(XY) = \sum_n \sum_m n\left(m + \frac{1}{m}\right) \frac{c}{n^3} \frac{c}{n^3}$$

$X \perp Y$

$E f(X, Y))$

$= f(E(X), E(Y))$

Chose Y s.t.

$E(Y) = 0$

& $f(x, y) = \frac{x}{y}$

$f: \mathbb{R}^+ \to \mathbb{R}^+$

$$= \sum_n n \frac{c}{n^3} \sum_m \left(m + \frac{1}{m}\right) \frac{c}{n^3}$$