Random Variables (rv)

Functions of a Random Variable [Ross S4.4]

Say we have a random variable X. Let $Y = g(X)$ for some function $g(.)$. Then:

- X is a function of the outcome $s \in S$
- Y is a function of X

$\Rightarrow$ Y is a function of the outcome $s \in S$

$\Rightarrow$ Y is a random variable.

Y has a PMF $p_Y(y)$. We can find it from $p_X(x)$.

Example 10.1: Let X be a random variable such that

$$P[X=-1]=0.1, \qquad P[X=0]=0.3, \qquad P[X=1]=0.6.$$

Let $Y = X^2$. What are $E[X]$ and $E[Y]$?

Solution:

$$\begin{aligned} E[X] &= -1 \times 0.1 + 0 \times 0.3 + 1 \times 0.6 \\ &= 0.5 \end{aligned}$$

$$\begin{aligned} P[Y=0] &= P[X^2=0] \\ &= P[X=0] \\ &= 0.3 \end{aligned}$$

$$\begin{aligned} P[Y=1] &= P[X^2=1] \\ &= P[\{X=1\} \cup \{X=-1\}] \\ &= 0.1 + 0.6 \end{aligned}$$

So

$$\begin{aligned} E[X^2] = E[Y] &= 0 \times 0.3 + 1 \times 0.7 \\ &= 0.7 \end{aligned}$$

Note: $(E[X])^2 = (0.5)^2 \neq 0.7 = E[X^2]$.

So $E[g(X)] \neq g(E[X])$ in general.

Proposition 10.1 *If X is a rv with possible values $\mathcal{X} = \{x_1, x_2, \ldots\}$ then*

$$E[g(X)] = \sum_{i \geq 1} g(x_i) p_X(x_i)$$

Why is this true? Let $Y = g(X)$.

Let $\mathcal{Y} = \{y_1, y_2, \ldots\}$ be all possible values of Y.

$$\begin{aligned} \sum_{i\geq 1} g(x_i)p_X(x_i) &= \sum_{j\geq 1} \sum_{i:g(x_i)=y_j} g(x_i)p_X(x_i) \\ &= \sum_{j\geq 1} \sum_{i:g(x_i)=y_j} y_j p_X(x_i) \\ &= \sum_{j\geq 1} y_j \sum_{i:g(x_i)=y_j} p_X(x_i) \\ &= \sum_{j\geq 1} y_j P[g(X)=y_j] \\ &= \sum_{j\geq 1} y_j P[Y=y_j] \\ &= E[Y] \\ &= E[g(X)] \end{aligned}$$

Example 10.2: In Example 10.1,

$$\begin{aligned} E[X^2] &= \sum_i x_i^2 p_X(x_i) \\ &= (-1)^2 \times p_X(-1) + 0^2 \times p_X(0) + 1^2 \times p_X(1) \\ &= 1 \times 0.1 + 0 \times 0.3 + 1 \times 0.6 \\ &= 0.7 \end{aligned}$$

Corollary 10.1 *If a and b are constants, then $E[aX+b] = aE[X]+b$.*

Why?

$$\begin{aligned} E[aX+b] &= \sum_{x\in\mathcal{X}} (ax+b)p_X(x) \\ &= a\sum_{x\in\mathcal{X}} xp_X(x) + b\sum_{x\in\mathcal{X}} p_X(x) \\ &= aE[X]+b \end{aligned}$$

Example 10.3: Say $E[X]=3$. Then $E[10X+4] = 10 \times 3 + 4 = 34$.

Note: $E[X]$ is called **mean** of X. $E[X^n]$ is called the n-th **moment** of X.

Often write $\mu_X = E[X]$.

Variance [Ross S4.5]

Given X, it is useful to summarize some essential properties of X.

$E[X]$ tells us about the "center" of how X is distributed.

Example 10.4: Let

$$\begin{aligned} P[W=0] &= 1 \\ P[Y=1] = P[Y=-1] &= \frac{1}{2} \\ P[Z=100] = P[Z=-100] &= \frac{1}{2} \end{aligned}$$

Then $E[W] = 0 = E[Y] = E[Z]$, but these are not equally spread...

Definition 10.1: The **variance** of X is

$$\begin{aligned} Var[X] &= E[\,(X-E[X])^2\,] \\ &= E[\,(X-\mu_X)^2\,] \end{aligned}$$

We often write $\sigma_X^2 = Var[X]$.

Note: Since $(X-\mu_X)^2 \geq 0$, then $Var[X] \geq 0$. $\qquad (*)$

Also

$$\begin{aligned} Var[X] &= E[\,(X-\mu_X)^2\,] \\ &= \sum_{x\in\mathcal{X}} (x-\mu_X)^2 p(x) \\ &= \sum_{x\in\mathcal{X}} \left(x^2 - 2\mu_X x + \mu_X^2\right) p(x) \\ &= \sum_{x\in\mathcal{X}} x^2 p_X(x) - 2\mu_X \underbrace{\sum_{x\in\mathcal{X}} x p_X(x)}_{\mu_X} + \mu_X^2 \sum_{x\in\mathcal{X}} p_X(x) \\ &= E[X^2] - 2\mu_X^2 + \mu_X^2 \qquad (\,\mu_X = E[X]\,) \\ &= E[X^2] - (E[X])^2 \end{aligned} \tag{10.1}$$

Also, combinining $(*)$ with (10.1), we get

$$E[X^2] \geq (E[X])^2 \tag{10.2}$$

and, if $E[X] > 0$, then

$$\frac{E[X^2]}{E[X]} \geq E[X] \tag{10.3}$$

Example 10.5: Let X be the the outcome of a dice roll. What is $Var[X]$?

Solution:

$$\begin{aligned} E[X] &= 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + \cdots + 6 \times \frac{1}{6} = \frac{7}{2} \\ E[X^2] &= 1 \times \frac{1}{6} + 2^2 \times \frac{1}{6} + \cdots + 6^2 \times \frac{1}{6} = \frac{91}{6} \end{aligned}$$

So, $\qquad Var[X] = E[X^2] - (E[X])^2 = \dfrac{91}{6} - \left(\dfrac{7}{2}\right)^2 = \dfrac{35}{12}$

Also:

$$\begin{aligned} E[\,(X-E[X])^2\,] &= \left(1-\frac{7}{2}\right)^2 \times \frac{1}{6} + \left(2-\frac{7}{2}\right)^2 \times \frac{1}{6} + \cdots + \left(6-\frac{7}{2}\right)^2 \times \frac{1}{6} \\ &= \frac{35}{12} \end{aligned}$$

Example 10.6: The distance from Vancouver to Boston is 4200km. If the wind is good (with probability 0.7), the speed of a plane is $V = 700$ km/h. If the wind is not good (probability 0.3), the speed is $V = 600$ km/h.

What is the average flight time?

Solution: If the wind is good, the flight time $T = 4200/700 = 6$ hours.

If the wind is not good, then $T = 4200/600 = 7$ hours.

So,

$$P[T=6] = 0.7, \qquad P[T=7] = 0.3$$

and $E[T] = 6 \times 0.7 + 7 \times 0.3 = 6.3$ hours

Note, that this is not the same as computing the average speed

$$E[V] = 700 \times 0.7 + 600 \times 0.3 = 670 \text{ km/h},$$

and then computing $4200/670 \approx 6.27$ hours.

In other words, even though $T = \frac{4200}{V}$, $E[T] \neq \frac{4200}{E[V]}$.