Random Variables (rv)

Examples [Ross S4.5]

Example 11.1: [Friendship Paradox]

There are n people named $1, 2, \dots, n$.

Person i has $f(i)$ friends. Let $m = \sum_{i=1}^{n} f(i)$.

Let X be a random person, equally likely to be any of the n people.

Let $Z = f(X)$, i.e., Z is # of friends of random person X.

Then

$$E[Z] = \sum_{i=1}^{n} f(i) \underbrace{P[X = i]}_{1/n} = \frac{m}{n} \qquad \text{[by Prop. 10.1]}$$

$$E[Z^2] = \sum_{i=1}^{n} (f(i))^2 \, P[X = i] = \frac{1}{n} \sum_{i=1}^{n} (f(i))^2$$

Now, each person writes the names of their friends on a sheet of paper (one sheet per friend).

There are m sheets, and one sheet is drawn at random, each sheet being equally likely to be chosen.

Let

$Y =$ name of friend on drawn sheet
$W = f(Y)$

Now

$$P[Y = i] = \frac{f(i)}{m} \qquad \left[\text{as opposed to } \frac{1}{n}\right]$$

$$\begin{aligned} E[W] &= E[f(Y)] \\ &= \sum_i f(i) \, P[Y = i] \\ &= \sum_i f(i) \times \frac{f(i)}{m} \\ &= \frac{n}{m} \times \frac{1}{n} \sum_i (f(i))^2 \\ &= \frac{E[Z^2]}{E[Z]} \\ &\geq E[Z] \qquad \text{[since } E[Z^2] \geq (E[Z])^2\text{]} \end{aligned}$$

So:

$$\begin{aligned} &(\text{expected \# of friends of random person } = E[Z]) \\ &\leq (\text{expected \# of friends of random friend } = E[W]) \end{aligned}$$

Example 11.2: There are n days in a year.

Persons $1, 2$ and 3 are independently born on day r with probability p_r, for $r = 1, 2, \dots, n$.

Let $A_{i,j} = \{\text{persons } i \text{ and } j \text{ born on same day}\}$

a) Find $P[A_{1,3}]$

b) Find $P[A_{1,3} \mid A_{1,2}]$

Solution:

a)

$$\begin{aligned} P[A_{1,3}] &= P[\cup_r \{1 \text{ and } 3 \text{ both born on day } r\}] \\ &= \sum_r P[\{1 \text{ and } 3 \text{ both born on day } r\}] \\ &= \sum_r P[\{1 \text{ born on day } r\}] P[\{3 \text{ born on day } r\}] \\ &= \sum_r p_r^2 \end{aligned}$$

b)

$$\begin{aligned} P[A_{1,3} \mid A_{1,2}] &= \frac{P[A_{1,3} A_{1,2}]}{P[A_{1,2}]} \\ &= \frac{P[\{1, 2 \text{ and } 3 \text{ born on same day}\}]}{P[\{1 \text{ and } 2 \text{ born on same day}\}]} \\ &= \frac{\sum_r p_r^3}{\sum_r p_r^2} \end{aligned}$$

Remark 11.1: We had $E[aX + b] = aE[X] + b$. What about $Var[aX + b]$?

$$\begin{aligned} Var[aX + b] &= E\left[(aX + b - E[aX + b])^2\right] \\ &= E\left[(aX + b - aE[X] - b)^2\right] \\ &= E\left[(aX - aE[X])^2\right] \\ &= E\left[a^2 (X - E[X])^2\right] \\ &= E\left[a^2 Y\right] \qquad \text{where } Y = (X - E[X])^2 \\ &= a^2 E[Y] \\ &= a^2 E\left[(X - E[X])^2\right] \\ &= a^2 \, Var[X] \end{aligned}$$

Remark 11.2: If X has units of, say, kg, then:

- $E[X] = \mu_X$ has units of kg,
- $Var[X] = \sigma_X^2$ has units of kg^2.

We also define $SD[X] = \sqrt{Var[X]} = \sigma_X$, called **standard deviation**.

$SD[X]$ has units of kg again.