M362K: March 29th, 2024.

Section 3.2.

Google.
Anscombe Quartet.
Dinosaurus Data.

Review.

Def'n. Let X be a discrete r.v. w/ pmf $p_X(\cdot)$. We define its $\boxed{\text{expectation}}$ as

$$\boxed{\mathbb{E}[X] = \sum_{\text{all } x} x \cdot p_X(x)}$$

Example.

$$\boxed{\mathbb{E}[X \cdot Y] = \sum_{\text{all}(x,y)} x \cdot y \cdot p_{X,Y}(x,y)}$$

Def'n. We say that r.v.s X and Y are independent if

$$\boxed{\mathbb{P}[X \leq x, Y \leq y] = \mathbb{P}[X \leq x] \cdot \mathbb{P}[Y \leq y] = F_X(x) \cdot F_Y(y)} \quad \text{for all } x, y$$

Proposition. Let X and Y be discrete r.v. w/

joint pmf $p_{X,Y}(x,y)$

and

marginal pmfs $p_X(x)$ and $p_Y(y)$.

Then, X and Y are independent

$$\iff$$

$$p_{X,Y}(x,y) = p_X(x) \cdot p_Y(y)$$

If X and Y are independent,

then, $\boxed{\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]}$

Review. **Def'n.** Let X be a random variable w/ a finite mean μ_X. Then, the variance of X is defined as

$$\boxed{\mathrm{Var}[X] := \mathbb{E}\left[(X-\mu_X)^2\right] = \mathbb{E}[X^2] - \mu_X^2}$$

Def'n. Let X be a r.v. w/ a finite variance. Its standard deviation is defined as $\boxed{SD[X] = \sqrt{\mathrm{Var}[X]}}$.

We usually write $\boxed{\sigma_X = SD[X]}$.

Problem. Consider a box w/ four tickets in it numbered 0, 1, 1, 2. The tickets are extracted from the box @ random and the result is denoted by the r.v. X.
Find $\mathbb{E}[X]$ and $SD[X]$.

→ : $\mathrm{Support}(X) = \{0, 1, 2\}$

pmf of X.

x	0	1	2
$p_X(x)$	1/4	1/2	1/4

- $\mu_X = \mathbb{E}[X] = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = 1$
- $\mathbb{E}[X^2] = 0^2 \cdot \frac{1}{4} + 1^2 \cdot \frac{1}{2} + 2^2 \cdot \frac{1}{4} = \frac{3}{2}$

$$\Rightarrow \mathrm{Var}[X] = \mathbb{E}[X^2] - \mu_X^2 = \frac{3}{2} - 1 = \frac{1}{2}$$

- $\sigma_X = \sqrt{\frac{1}{2}} = \frac{\sqrt{2}}{2}$

□

Linear Transforms of Random Variables.

Let X have a finite variance and let a and b be constants. Define

$$\boxed{Y = a \cdot X + b}$$

$$\mathrm{Var}[Y] = \mathrm{Var}[aX + \underbrace{b}_{\text{shift}}]$$

$$= \mathbb{E}\left[\left((aX+b) - \mathbb{E}[aX+b]\right)^2\right]$$ linearity of $\mathbb{E}$

$$= \mathbb{E}\left[\left(aX + \cancel{b} - a\cdot\mathbb{E}[X] - \cancel{b}\right)^2\right]$$

$$= \mathbb{E}\left[a^2(X-\mathbb{E}[X])^2\right] = a^2\,\boxed{\mathbb{E}\left[(X-\mathbb{E}[X])^2\right]} = a^2\cdot\boxed{\mathrm{Var}[X]}$$

linearity of $\mathbb{E}$

Also,

$$SD[aX+b] = \sqrt{\mathrm{Var}[aX+b]} = \sqrt{a^2\cdot \mathrm{Var}[X]}$$

$$= |a|\cdot SD[X]$$

Standard Units.

Let X be a r.v. w/ mean μ_X and standard deviation σ_X.

In standard units $\boxed{X^* = \dfrac{X-\mu_X}{\sigma_X}}$

- $\mathbb{E}[X^*] = \mathbb{E}\left[\dfrac{X-\mu_X}{\sigma_X}\right] = \dfrac{1}{\sigma_X}\left(\underbrace{\mathbb{E}[X]-\mu_X}_{=0}\right) = 0$ ⌣

  linearity of $\mathbb{E}$

- $\mathrm{Var}[X^*] = \mathrm{Var}\left[\dfrac{1}{\sigma_X}\cdot X - \dfrac{\mu_X}{\sigma_X}\right] = \mathrm{Var}\left[\dfrac{1}{\sigma_X}\cdot X\right] = \dfrac{1}{\sigma_X^2}\mathrm{Var}[X]$

  shift

  $= 1$

- $SD[X^*] = 1$

Law of Averages.

Law of Large Numbers.

Let $\{X_1, X_2, \ldots, X_n, \ldots\}$ be a sequence of independent r.v.s.

Assume they are all identically distributed and that their common mean is μ_X.

$$\boxed{\frac{X_1+X_2+\cdots+X_n}{n} \longrightarrow \mu_X}$$

<u>Def'n.</u> Say that X and Y are r.v.s w/ finite variances. Then, the covariance between X and Y is

$$\boxed{\mathrm{Cov}[X,Y] = \mathbb{E}\left[(X-\mu_X)(Y-\mu_Y)\right]}$$

- $\mathrm{Cov}[X,X] = \mathrm{Var}[X]$
- $\mathrm{Cov}[X,Y] = \mathbb{E}[(X-\mu_X)(Y-\mu_Y)]$

$$= \mathbb{E}[XY - \mu_X \cdot Y - X \cdot \mu_Y + \mu_X \mu_Y]$$

linearity of $\mathbb{E}$

$$= \mathbb{E}[XY] - \mu_X \cdot \underbrace{\mathbb{E}[Y]}_{=\mu_Y} - \underbrace{\mathbb{E}[X]}_{=\mu_X} \mu_Y + \mu_X \mu_Y$$

$$= \mathbb{E}[XY] - \mu_X \cdot \mu_Y$$

- If X and Y are independent, then $\boxed{\mathrm{Cov}[X,Y] = 0}$