CS 70, July 26, 2022 Jingjia Chen

Variance & correlation.

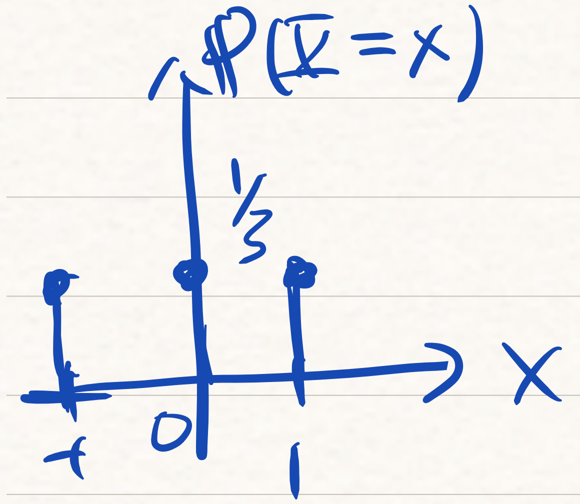
Function of Random Variables

Ex ① $Y = cX$ $c \in \mathbb{R}$ $(c \neq 0)$. find $\mathbb{E}(Y)$

$$\mathbb{E}(f(X)) = \sum_x f(x) \, \mathbb{P}_X(X = x)$$

$$\mathbb{E}(Y) = \sum_x cx \, \mathbb{P}_X(X = x) = c\,\mathbb{E}(X)$$

Ex ②. $X \sim U[-1, 1]$, $Y = X^2$ find $\underline{\mathbb{P}_Y}$ & $\mathbb{E}[Y]$



$y = x^2$	1	0	1
x	-1	0	1
$\mathbb{P}(X=x)$	$1/3$	$1/3$	$1/3$

$\mathbb{P}(Y=0) = 1/3 \quad \mathbb{P}(Y=1) = 2/3$

$\mathbb{E}(Y) = 0 \cdot 1/3 + 1 \cdot 2/3 = 2/3$

Ex③ $X \sim U[-N, N]$, $Y = X^2$ find $\mathbb{P}_Y$ and $\mathbb{E}(Y)$

$$\boxed{\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}}$$

$$\mathbb{P}_X(X = x) = \begin{cases} \frac{1}{2N+1} & x \in [-N, N] \\ 0 & \text{o.w.} \end{cases}$$

$$\mathbb{P}_Y(Y = N^2 + N) = 0$$

$$\mathbb{P}_Y(Y = y = x^2) = \begin{cases} \frac{2}{2N+1} & y \in [1, N^2] \leftarrow \\ \frac{1}{2N+1} & y = 0 \leftarrow \\ 0 & \text{o.w.} \leftarrow \end{cases}$$

$$\mathbb{E}(Y) = 0 \cdot \frac{1}{2N+1} + \sum_{k=1}^{N} k^2 \frac{2}{2N+1} + 0 = \frac{2}{2N+1} \sum_{k=1}^{N} k^2 = \frac{N(N+1)}{3}$$

Variance of Random Variable (r.v) $X : \Omega \to \mathbb{R}$

Def. a. r.v. X w/ $\mathbb{E}(X) = \mu$. the variance of X is

$$\mathrm{Var}(X) = \mathbb{E}((X-\mu)^2)$$

$\sigma(X) := \sqrt{\mathrm{Var}(X)}$ is standard deviation of X

$$\mathrm{Var}(X) = \mathbb{E}((X-\mu)^2) = \mathbb{E}(X^2+\mu^2-2\mu X) = \mathbb{E}(X^2)+\mu^2-2\mu \overset{\swarrow}{\mathbb{E}(X)}$$

$$= \mathbb{E}(X^2) - \mu^2 = \mathbb{E}(X^2) - \mathbb{E}(X)^2 \quad \text{(long − short)}$$

Ex① $Y = cX$: $\mathbb{E}(Y) = c\,\mathbb{E}(X)$ if X : $\mathrm{var}(X)$. what is $\mathrm{var}(Y)$?

$$\mathrm{var}(Y) = \mathbb{E}(Y^2) - \mathbb{E}(Y)^2 = c^2\,\mathbb{E}(X^2) - c^2\,\mathbb{E}(X)^2 = \underset{\triangle}{c^2}\,\mathrm{Var}(X)$$

1. 6- face dice X, find var(X)

$$\mathbb{E}(X) = \frac{1}{6}(1+2+\cdots+6) = \frac{(1+6)\cdot 6}{6\cdot 2} = \frac{7}{2}$$

$$\mathbb{E}(X^2) = \frac{1}{6}\cdot 1^2 + \frac{1}{6}\cdot 2^2 + \cdots + \frac{1}{6}\cdot 6^2 = \frac{91}{6}$$

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{91}{6} - \frac{49}{4} = \frac{35}{12}$$

2: Uniform dist. $X \sim U\{1, 2, \cdots, n\}$. $\mathbb{P}(X = x) = \frac{1}{n}$ $x \in [1, n]$

$$\underline{\mathbb{E}(X)} = \frac{n+1}{2} = \frac{1}{n}(1+2+\cdots+n) = \frac{1}{n}\cdot\frac{n(1+n)}{2}$$

$$\mathrm{Var}(X) = \underline{\mathbb{E}(X^2)} - \underline{\mathbb{E}(X)^2} = \underline{\frac{1}{n}(1^2+2^2+\cdots n^2)} - \frac{(n+1)^2}{\underline{4}}$$

$$= \frac{n^2-1}{12}$$

$$\mathbb{E}(X^2) = \frac{1}{n}(1^2+2^2+\cdots+n^2)$$

$$= \frac{1}{n}\frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6}$$

$$\boxed{\sum_{n=1}^{N} n^2 = \frac{N(N+1)(2N+1)}{6}}$$

$$\mathbb{E}(X)^2 = \frac{(n+1)^2}{4}$$

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{2(n+1)(2n+1)-3(n+1)^2}{12}$$

$$= \frac{4n^2+6n+2-3n^2-3-6n}{12}$$

$$= \frac{n^2-1}{12}$$

two r.v. X, Y, $\mathbb{E}(X)$, $\mathbb{E}(Y)$, $var(X)$, $var(Y)$

new r.v. $Z = X+Y$. find $var(Z)$.

$$Var(Z) = Var(X+Y) = \mathbb{E}((X+Y)^2) - (\mathbb{E}(X+Y))^2$$

$$= \mathbb{E}(X^2+Y^2+2XY) - (\mathbb{E}(X)+\mathbb{E}(Y))^2$$

$$= \underline{\mathbb{E}(X^2)} + \underline{\mathbb{E}(Y^2)} + \underline{2\mathbb{E}(XY)} - \underline{\mathbb{E}(X)^2} - \underline{\mathbb{E}(Y)^2} - 2\mathbb{E}(X)\mathbb{E}(Y)$$

$$= var(X) + var(Y) + 2(\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y))$$

if $X \perp Y$ then $\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = 0$

↓

meaning X independet of Y

→ (the other way is not true !)

$\mathbb{E}(Y\mathbb{E}(X)) = \mathbb{E}(X)\mathbb{E}(Y)$

Covariance

Def: $\operatorname{cov}(X,Y) = \mathbb{E}((X-\mathbb{E}(X))(Y-\mathbb{E}(Y)))$

$= \mathbb{E}(XY - Y\mathbb{E}(X) - X\mathbb{E}(Y) + \mathbb{E}(X)\mathbb{E}(Y))$

$= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$

facts:

1) $\operatorname{var}(X+Y) = \operatorname{Var}(X) + \operatorname{var}(Y) + 2\operatorname{cov}(X,Y)$
2) $\operatorname{cov}(X,X) = \operatorname{Var}(X)$
3) $\operatorname{cov}(aX, bY) = ab\operatorname{cov}(X,Y)$
4) $X \perp Y \Rightarrow \operatorname{cov}(X,Y) = 0$

correlation (coefficient)

Def. r.v. X, Y. $\sigma(X) > 0$. $\sigma(Y) > 0$

$$\operatorname{corr}(X,Y) = \frac{\operatorname{cov}(X,Y)}{\sigma(X)\sigma(Y)} \in [-1,1]$$

→ normalized cov.

Ex. show $\operatorname{corr}(\cdot,\cdot) \in [-1,1]$

proof: let $\tilde{X} = (X - \mathbb{E}(X))/\sigma(X)$

$\mathbb{E}(\tilde{X}) = 0$

$\operatorname{var}(\tilde{X}) = 1$

$\mathbb{E}(\tilde{X}^2) = 1$

$$\mathbb{E}(\tilde{X}^2) = \mathbb{E}\left(\frac{X^2 + \mathbb{E}^2(X) - 2X\mathbb{E}(X)}{\sigma^2(X)}\right)$$

$$= \frac{\mathbb{E}(X^2) - \mathbb{E}(X)^2}{\operatorname{var}(X)} = 1$$

$$\tilde{X} = (X - \mathbb{E}(X))/\sigma(X) \qquad \mathbb{E}(\tilde{X}) = 0, \ \mathrm{Var}(\tilde{X}) = 1$$

$$\hat{Y} = (Y - \mathbb{E}(Y))/\sigma(Y) \qquad \mathbb{E}(\tilde{Y}) = 0, \ \mathrm{var}(\tilde{Y}) = 1$$

for $\underline{\underline{\tilde{X}+\tilde{Y}}}$ $\quad$ $\underline{\underline{\tilde{X}-\tilde{Y}}}$

$$\underline{\underline{0}} \le \mathbb{E}((\hat{X}+\tilde{Y})^2) = \overset{\overset{1}{\swarrow}}{\mathbb{E}(\tilde{X}^2)} + \overset{\overset{1}{\swarrow}}{\mathbb{E}(\tilde{Y}^2)} + \overset{\swarrow}{2\mathbb{E}(\tilde{X}\tilde{Y})} \le \underline{2 + \boxed{2\mathbb{E}(\tilde{X}\tilde{Y})}}$$

$$\underline{\underline{0}} \le \mathbb{E}((\hat{X}-\tilde{Y})^2) = \mathbb{E}(\tilde{X}^2) + \mathbb{E}(\tilde{Y}^2) - 2\mathbb{E}(\tilde{X}\tilde{Y}) \le \underline{2 - 2\mathbb{E}(\tilde{X}\tilde{Y})}$$

$\underline{\underline{\mathbb{E}(\tilde{X}\tilde{Y}) \in [-1, 1]}}$ (see next page)

$$-1 \le \underbrace{\mathrm{corr}(X, Y)}_{} \overset{}{\underset{*}{=}} \underbrace{\mathrm{cov}(\hat{X}, \tilde{Y})}_{} = \mathbb{E}(\tilde{X}\tilde{Y}) \le 1$$

next page

$$\left.\begin{array}{l} 0 \le 2+2\mathbb{E}(\tilde{X}\tilde{Y}) \Rightarrow \mathbb{E}(\tilde{X}\tilde{Y}) \ge -1 \\ 0 \le 2-2\mathbb{E}(\tilde{X}\tilde{Y}) \Rightarrow \mathbb{E}(\tilde{X}\tilde{Y}) \le 1 \end{array}\right\} \Rightarrow -1 \le \mathbb{E}(\tilde{X}\tilde{Y}) \le 1$$

when $X = Y$ $\operatorname{corr}(X,Y) = 1$ when $X = -Y$, $\operatorname{corr}(X,Y) = -1$

$*$ $\underline{\operatorname{corr}(X,Y)} = \dfrac{\operatorname{cov}(X,Y)}{\sigma(X)\sigma(Y)} = \operatorname{cov}\left(\dfrac{X}{\sigma(X)}, \dfrac{Y}{\sigma(Y)}\right)$

$$= \mathbb{E}\left(\left(\frac{X}{\sigma(X)} - \mathbb{E}\left(\frac{X}{\sigma(X)}\right)\right)\left(\frac{Y}{\sigma(Y)} - \mathbb{E}\left(\frac{Y}{\sigma(Y)}\right)\right)\right)$$

$$= \mathbb{E}\left(\left(\frac{X-\mathbb{E}(X)}{\sigma(X)}\right)\left(\frac{Y-\mathbb{E}(Y)}{\sigma(Y)}\right)\right) = \mathbb{E}(\tilde{X}\tilde{Y}) = \mathbb{E}(\tilde{X}\tilde{Y}) - \overset{\overset{0}{\downarrow}}{\mathbb{E}(\tilde{X})}\,\overset{\overset{0}{\downarrow}}{\mathbb{E}(\tilde{Y})}$$

$$= \underline{\operatorname{cov}(\tilde{X}\tilde{Y})}$$

$\uparrow$ $\tilde{X} = \dfrac{X - \mathbb{E}(X)}{\sigma(X)}$, $\tilde{Y} = \dfrac{Y - \mathbb{E}(Y)}{\sigma(Y)}$

if $\overline{X = Y}$ then $\operatorname{corr}(X,Y) = 1$

↑ there's a glitch here

I can't remove it :'(

Independence $\Rightarrow$ uncorrelated

$\not\Leftarrow$

<u>indep.</u> $X \perp Y \iff \mathbb{P}_{XY}(x,y) = \mathbb{P}_X(x)\,\mathbb{P}_Y(y)$

<u>uncorr.</u> $\operatorname{cov}(X,Y) = 0 \iff \mathbb{E}(XY) = \mathbb{E}(X)\,\mathbb{E}(Y)$

two r.v. can be uncorrelated but dependent,
but if they are independent, they must are uncorrelated.

Ex ④ $X \sim U[-1,1]$, $Y = X^2$

$\boxed{\text{uncorrelated} \not\Rightarrow \text{independence}}$

↳ they are dependent

$\mathbb{E}(X)=0$, $\mathbb{E}(Y)=2/3$ $\underline{\mathbb{E}(XY)} = 0$

$\underline{cov(X,Y)=0}$

X	-1	0	1
Y	1	0	1
XY	-1	0	1
$\mathbb{P}$	$1/3$	$1/3$	$1/3$

$$\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$
$$= 0 - 0 \cdot 2/3$$
$$= 0$$

$$\boxed{0 = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)}$$

↓ def of uncorrelated

$$\mathbb{E}(XY) = -1 \cdot 1/3 + 0 \cdot 1/3 + 1 \cdot 1/3 = 0$$