Midterm:

Time : Next Friday, normal lecture time

Alt: Time : TBD.

Scope: Everything before Sec 6.2, except Sec 5.4.

Where : On Gradescope.

Length: 100 minutes to solve questions and

10 minutes to scan & upload your work.

note: start before 6:10 p.m to enjoy 110 minutes

Open book, open notes, calculators and Python notebook allows.

No communication with anyone else except me and the GSI.

Last time.

Sec. 5.3 method of indicators

Sec. 5.5 conditional expectation

5.6 expectation by conditioning

Today.

Sec 6.1 & 6.2 Variance & Standard deviation

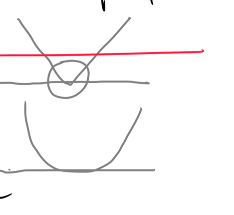
Variation

Let X be a R.V., $\mu = \mathbb{E}X$ is the mean / expectation

The deviation of X (from the mean) is said to be

$D := X - \mu$.

Problem is $\mathbb{E}D = \mathbb{E}X - \mu = \mu - \mu = 0$.

Two solutions: $\begin{cases} \cancel{\mathbb{E}|D|} \\ \mathbb{E}D^2 \end{cases}$  "Variance"

Denote as $\mathrm{Var}(X) = \mathbb{E}D^2$

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

Example. $X \sim \mathrm{Binom}(3, \frac{1}{2})$ $\mu = \frac{3}{2}$.

$g(x) = (x - \frac{3}{2})^2$

X	0	1	2	3
prob.	1/8	3/8	3/8	1/8

$$\begin{aligned}\mathrm{Var}(X) &= \mathbb{E}(X-\mu)^2 \\ &= \mathbb{E}\, g(X) \\ &= \sum_x g(x)\, \mathbb{P}(X=x) \\ &= (0-\tfrac{3}{2})^2 \cdot \tfrac{1}{8} + (1-\tfrac{3}{2})^2 \cdot \tfrac{3}{8} + (2-\tfrac{3}{2})^2 \cdot \tfrac{3}{8} + (3-\tfrac{3}{2})^2 \cdot \tfrac{1}{8} \\ &= \tfrac{9}{4} \cdot \tfrac{1}{8} + \tfrac{1}{4} \cdot \tfrac{3}{8} + \tfrac{1}{4} \cdot \tfrac{3}{8} + \tfrac{9}{4} \cdot \tfrac{1}{8} = \tfrac{24}{32} = \tfrac{3}{4}.\end{aligned}$$

$\Rightarrow SD(X) = \sqrt{\frac{3}{4}} = \frac{\sqrt{3}}{2}$

Alt. sol

$$\begin{aligned}\mathrm{Var}(X) &= \mathbb{E}X^2 - \mu^2 \\ &= 0^2 \cdot \tfrac{1}{8} + 1^2 \cdot \tfrac{3}{8} + 2^2 \cdot \tfrac{3}{8} + 3^2 \cdot \tfrac{1}{8} - (\tfrac{3}{2})^2 \\ &= \tfrac{3}{4}.\end{aligned}$$

Property.

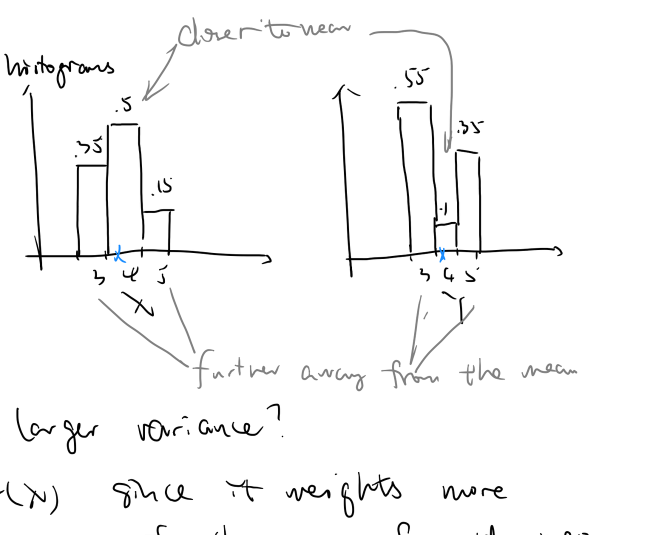
① non-negativity: $\mathrm{Var}(X) \geq 0$, and $\mathrm{Var}(X) = 0$ only when X is a constant.

② shift-invariance: $\mathrm{Var}(X) = \mathrm{Var}(X+c)$, $\forall c \in \mathbb{R}$

Since $D_c := X + c - \mathbb{E}(X+c) = X - \mu = D$.

Example 2:

	3	4	5
$\mathbb{P}(X=\cdot)$	35%	50%	15%
$\mathbb{P}(Y=\cdot)$	55%	10%	35%



$\mathbb{E}X = \mathbb{E}Y = 3.8$

Which of X or Y has a larger variance?

Intuitively. $\mathrm{Var}(Y) > \mathrm{Var}(X)$ since it weights more on the numbers that are further away from the mean.

$\mathrm{Var}(X) = (3-3.8)^2 \cdot 35\% + (4-3.8)^2 \cdot 50\% + (5-3.8)^2 \cdot 15\% \approx 0.68$

$\mathrm{Var}(Y) \approx 0.93$.

Linear transform of Variance

Example X is the room temperature measured in °F

Y is the - - - - - - - °C

$X = \frac{9}{5}Y + 32$

Question : what is the relationship between $\mathrm{Var}(X)$ & $\mathrm{Var}(Y)$

First thing (easy): $\mathrm{Var}(X) = \mathrm{Var}(\frac{9}{5}Y + 32)$

$= \mathrm{Var}(aY)$. $a = \frac{9}{5}$

$$\begin{aligned}\mathrm{Var}(aY) &= \mathbb{E}\left[(aY - \mathbb{E}(aY))^2\right] \\ &= \mathbb{E}(aY - a\mathbb{E}Y)^2 \\ &= a^2\, \mathbb{E}(Y - \mathbb{E}Y)^2 = a^2\, \mathrm{Var}(Y)\end{aligned}$$

It makes sense since the unit of $\mathrm{Var}(Y)$ is square of the unit of Y.

Generally. $\mathrm{Var}(aX + b) = a^2\, \mathrm{Var}(X)$, $\forall a, b \in \mathbb{R}$.

Problem : unit of $\mathrm{Var}(Y)$ is not the same as Y.

Solution: Consider $SD(Y) = \sqrt{\mathrm{Var}(Y)}$ Standard deviation.

"root mean square of deviations from average."

⑤	④	③	②	①
$\sqrt{\mathrm{Var}(X)}$	$\mathbb{E}D^2$	D^2	$D = X - \mu$	$\mu = \mathbb{E}X$
$SD(X)$	$\mathrm{Var}(X)$			

Another way to get $\mathrm{Var}(X)$.

$$\begin{aligned}\mathrm{Var}(X) &= \mathbb{E}(X - \mathbb{E}X)^2 \\ &= \mathbb{E}(X^2 - 2X \cdot \mu + \mu^2) \\ &= \mathbb{E}X^2 - \mathbb{E}\,2\mu X + \mathbb{E}\mu^2 \\ &= \mathbb{E}X^2 - 2\mu \underbrace{\mathbb{E}X}_{\mu} + \mu^2 \\ &= \mathbb{E}X^2 - \mu^2.\end{aligned}$$

Sec 6.3 (Not in midterm).

Markov's inequality

It bounds the right tail prob. of X using only $\mu = \mathbb{E}X$.

$\mathbb{P}(X \geq c)$

X has to be nonnegative.

$$\begin{aligned}\mathbb{E}X &= \Big(\sum_{x<c} + \sum_{x \geq c}\Big) x\, \mathbb{P}(X = x) \\ &\geq \sum_{x \geq c} x\, \mathbb{P}(X = x) \quad \text{(since } X \text{ is non-negative)} \\ &\geq \sum_{x \geq c} c\, \mathbb{P}(X = x) \\ &= c\, \mathbb{P}(X \geq c)\end{aligned}$$

$\Rightarrow \mathbb{P}(X \geq c) \leq \dfrac{\mu}{c}$ Markov inequality.

① The smaller μ is, the smaller the bound is.

② The larger c is, - —— - - --

③ We may say instead

$\mathbb{P}(X \geq k\mu) \leq \frac{1}{k}$.

$\mathbb{P}(X \geq 5\mu) \leq \frac{1}{5}$

This is just an upper bound, which may contains little to none information.

e.g. $\mathbb{P}(X \geq \mu) \leq 1$

e.g. $X \sim \mathrm{Binom}(100, \frac{1}{2})$ $\mu = 50$

$\mathbb{P}(X \geq 200) \leq \frac{1}{4}$. which is useless.

since we know $\mathbb{P}(X \geq 200) = 0$.

$\mathbb{P}(X \geq 75) \leq \frac{1}{1.5}$