⓪ Comment on yesterday's lecture

Pace was too fast at the end of lecture
I went quickly to try to finish everything on time

Today: Answer fewer questions from chat
No recap

I know some people like these things, but time is finite & going at an appropriate pace is more important

Start of lecture today: review last part of lecture 5B

① Combining random variables

X, Y random variables

New random variables by combining X and Y: $X+Y$, $3X+Y$, X^2, etc

Example Flip a fair coin 3 times

$X = \#$ of heads

$$Y = \begin{cases} 1 & \text{if } 1^{st} \text{ flip is heads} \\ 0 & \text{if } 1^{st} \text{ flip is tails} \end{cases}$$

$Z = X^2 + 3Y$

Question What is $Z(HHT)$?

Answer $X(HHT) = 2$
$Y(HHT) = 1$ $\Rightarrow$ $Z(HHT) = 2^2 + 3 \cdot 1 = 4+3 = 7$

② Expected value

$X : \Omega \to \mathbb{R}$ random variable

Definition: $E(X) = \sum_{a \in \text{range}(X)} a \cdot P(X = a)$ → sum over possible values of X

Equivalent Formula: $E(X) = \sum_{\omega \in \Omega} X(\omega) P(\omega)$ → sum over outcomes in Ω

Example Flip a coin 2 times

$X = \#$ of heads

sum over possible values: $E(X) = 0 \cdot P(X=0) + 1 \cdot P(X=1) + 2 \cdot P(X=2)$

$= 0 \cdot (1/4) + 1 \cdot (1/2) + 2 \cdot (1/4)$

$= 0 + 1/2 + 1/2 = 1$

sum over outcomes: $E(X) = X(HH)P(HH) + X(HT)P(HT) + X(TH)P(TH)$

$+ X(TT)P(TT)$

$= 2 \cdot (1/4) + 1 \cdot (1/4) + 1 \cdot (1/4) + 0 \cdot (1/4)$

$= 1/2 + 1/4 + 1/4 + 0 = 1$

(2.1) Useful Facts

X, Y random variables
$a \in \mathbb{R}$

Prop $E(aX) = aE(X)$

Prop $E(a) = a$

↳ constant random variable equal to a no matter what the outcome is

"law of unconscious statisticians"

Prop For any function $f: \mathbb{R} \times \mathbb{R} \to \mathbb{R}$

$$E(f(X,Y)) = \sum_{a \in \text{range}(X)} \sum_{b \in \text{range}(Y)} f(a,b)\, P(X=a \text{ and } Y=b)$$

Example $f(X,Y) = X \cdot Y$ X and Y both have range $\{0,1,2,\ldots,n\}$

$$E(XY) = \sum_{i=0}^{n} \sum_{j=0}^{n} i \cdot j \cdot P(X=i \text{ and } Y=j)$$

(2.2) Linearity of Expectation

Requires **no** assumptions abt $X_1, \dots, X_n$!

Thm $E(X_1 + X_2 + \dots + X_n) = E(X_1) + E(X_2) + \dots + E(X_n)$

Example n people, each wearing a hat
collect all hats & reassign them randomly

$X = 1$

X = # of students who get their own hat back

What is $E(X)$?

Attempt 1 Possible values of X: $0, 1, 2, \dots, n$

$$E(X) = \sum_{k=0}^{n} k \cdot \underbrace{P(X=k)}_{} = ??$$

Can calculate using inclusion-exclusion

Thm $E(X_1+X_2+\dots+X_n) = E(X_1)+E(X_2)+\dots+E(X_n)$

Example n people, each wearing a hat
collect all hats & reassign them randomly
X = # of students who get their own hat back
What is $E(X)$?

Attempt 2 Define random variables

$$X_1 = \begin{cases} 1 & \text{if person 1 gets their hat back} \\ 0 & \text{if person 1 does not get their hat back} \end{cases}$$

$$\vdots$$

$$X_n = \begin{cases} 1 & \text{if person } n \text{ gets their hat back} \\ 0 & \text{if person } n \text{ does not get their hat back} \end{cases}$$

indicator variables

$X = X_1+X_2+\dots+X_n$

$$E(X_i) = \frac{(n-1)!}{n!} = \frac{\cancel{(n-1)}\cancel{(n-2)}\cdots\cancel{1}}{n\cdot\cancel{(n-1)}\cdot\cancel{(n-2)}\cdots\cancel{1}} = \frac{1}{n}$$

$(n-1)!$ → permutations where i gets their own hat

$n!$ → all permutations

$$E(X) = E(X_1+X_2+\dots+X_n) = E(X_1)+E(X_2)+\dots+E(X_n)$$
$$= \frac{1}{n}+\frac{1}{n}+\dots+\frac{1}{n} = n\cdot\left(\frac{1}{n}\right) = 1$$

③ Variance

X is a random variable

How far is X from its expected value?

Def If X is a random variable with expected value $E(X) = a$ then the variance of X, written $\mathrm{Var}(X)$ is

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

Comment Standard deviation of X: $\sigma(X) = \sqrt{\mathrm{Var}(X)}$

Def If X is a random variable with expected value $E(X) = a$ then the variance of X, written $Var(X)$ is

$$Var(X) = E((X-a)^2)$$

Example Roll a fair 6-sided die. X = value of roll

$$E(X) = 1\cdot(1/6) + 2\cdot(1/6) + 3\cdot(1/6) + 4\cdot(1/6) + 5\cdot(1/6) + 6\cdot(1/6) = 21/6 = 3.5$$

outcome	X	$(X-3.5)^2$
1	1	$(1-3.5)^2 = 6.25$
2	2	$(2-3.5)^2 = 2.25$
3	3	$(3-3.5)^2 = 0.25$
4	4	$(4-3.5)^2 = 0.25$
5	5	$(5-3.5)^2 = 2.25$
6	6	$(6-3.5)^2 = 6.25$

$$Var(X) = E((X-3.5)^2) = 6.25\cdot(1/6) + 2.25(1/6) + 0.25(1/6) + 0.25\cdot(1/6) + 2.25(1/6) + 6.25(1/6) = 17.5/6 \approx 2.91$$

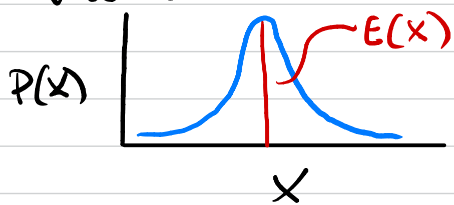
sum over outcomes

But what does it mean?

3.1 Meaning of variance

$a = E(X)$
$Var(X) = E((X-a)^2)$
Intuitively: measures how much X deviates from its expected value

low variance:



high variance:

Question Why not $E(X-a)$?
Answer Positive and negative deviations cancel out

Question Why not $E(|X-a|)$? mean absolute deviation
Answer Harder to calculate.
Actually this is sometimes used

Question Why not $E(X-a)$?

Answer Positive and negative deviations cancel out

Example Roll a fair 6-sided die

X = value of roll

$E(X) = 3.5$

outcome	X	$X - 3.5$
1	1	-2.5
2	2	-1.5
3	3	-0.5
4	4	0.5
5	5	1.5
6	6	2.5

$$E(X-3.5) = -2.5(1/6) - 1.5(1/6) - 0.5(1/6) + 0.5(1/6) + 1.5(1/6) + 2.5(1/6) = 0$$

3.2 Alternative formula for variance

Thm For any random variable X,

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

pf Let $a = E(X)$

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

$$= E(X^2 - 2aX + a^2)$$

linearity of expectation

$$= E(X^2) + E(-2aX) + E(a^2)$$

$$= E(X^2) - 2a\,E(X) + a^2$$

definition of a

$$= E(X^2) - 2E(X)E(X) + E(X)^2 = E(X^2) - E(X)^2$$

Thm For any random variable X, $Var(X) = E(X^2) - (E(X))^2$

Example (A bit tricky) $X \sim Bin(n,p)$ → flip n biased coins, p = heads prob. X = # of heads

What is $Var(X)$?

$Var(X) = E(X^2) - E(X)^2$, $E(X)^2 = (np)^2$ Need to find $E(X^2)$

Define $X_i = \begin{cases} 1 & \text{if coin } i \text{ is heads} \\ 0 & \text{if coin } i \text{ is tails} \end{cases}$ $X = X_1 + X_2 + \dots + X_n$

$$E(X^2) = E((X_1 + \dots + X_n)^2) = E\left(\sum_{i=1}^{n} \sum_{j=1}^{n} X_i X_j\right)$$

$$= \sum_{i,j} E(X_i X_j) = n \cdot p + \underbrace{(n^2 - n)}_{\hookrightarrow \text{ \# of pairs } i \neq j} \cdot p^2$$

$X_i X_j = \begin{cases} 1 & \text{if coins } i \ \& \ j \text{ are both heads} \\ 0 & \text{else} \end{cases}$

$E(X_i X_j) = \begin{cases} p^2 & \text{if } i \neq j \\ p & \text{if } i = j \end{cases}$

Final answer

$$Var(X) = E(X^2) - E(X)^2$$
$$= np + (n^2 - n)p^2 - (np)^2$$
$$= np + \cancel{n^2 p^2} - np^2 - \cancel{n^2 p^2} = np(1-p)$$

④ Covariance

Random variables X and Y

When X is high is Y usually high, low or "it depends"

relative to $E(X)$

Def If X is a random variable with $E(X) = a$ and Y is a random variable with $E(Y) = b$ then the covariance of X and Y, written $\mathrm{Cov}(X, Y)$ is

$$\mathrm{Cov}(X, Y) = E((X-a)(Y-b))$$

Question If you just know the distribution of X and the distribution of Y, can you calculate $\mathrm{Cov}(X, Y)$?

Answer No! You need the joint distribution

$\mathrm{Cov}(X, Y)$ measures something about the relationship between X and Y

Def $E(X) = a$ $E(Y) = b$ $\Rightarrow$ $Cov(X, Y) = E((X-a)(Y-b))$

Example Flip a fair coin 3 times

$X = \#$ of heads $\quad E(X) = 1.5$

$Y = \#$ of tails $\quad E(Y) = 1.5$

$$Cov(X, Y) = E((X-1.5)(Y-1.5)) = -(1.5)^2\left(\tfrac{2}{8}\right) - (0.5)^2\left(\tfrac{6}{8}\right)$$
$$= -0.75$$

Negatively correlated: when X is large, Y is small (& vice-versa)

Outcome	X	Y	$X-1.5$	$Y-1.5$	$(X-1.5)(Y-1.5)$
HHH	3	0	1.5	−1.5	$-(1.5)^2$
HHT	2	1	0.5	−0.5	$-(0.5)^2$
HTH	2	1	0.5	−0.5	$-(0.5)^2$
THH	2	1	0.5	−0.5	$-(0.5)^2$
HTT	1	2	−0.5	0.5	$-(0.5)^2$
THT	1	2	−0.5	0.5	$-(0.5)^2$
TTH	1	2	−0.5	0.5	$-(0.5)^2$
TTT	0	3	−1.5	1.5	$-(1.5)^2$

4.1 Meaning of covariance

Sign of $Cov(X,Y)$

$Cov(X,Y) < 0$	$Cov(X,Y) = 0$	$Cov(X,Y) > 0$
X and Y usually move in opposite directions	Just knowing X is large doesn't tell you whether Y is large or small	X and Y usually move in the same direction
X high ⇒ Y low X low ⇒ Y high	X high ⇒ Y can be low or high	X high ⇒ Y high X low ⇒ Y low

Magnitude of $Cov(X,Y)$

Hard to interpret because it depends on the variance of X and Y

More meaningfull measure:

$$Corr(X,Y) = \frac{Cov(X,Y)}{\sigma(X)\sigma(Y)}$$

→ Correlation of X and Y

↳ Same sign as $Cov(X,Y)$ but always between -1 and 1

Examples What is the sign of $\mathrm{Cov}(X,Y)$?

① Pick a random day from 2020
X = max temperature in Berkeley that day
Y = max temperature in Berkeley the next day

$\mathrm{Cov}(X,Y) > 0$
If it's hot one day, more likely to be hot the next day

② Pick a random day from 2020
X = max temperature in Berkeley that day
Y = number of people in Berkeley who used their heater that day

$\mathrm{Cov}(X,Y) < 0$
If it's hot, people usually use their heaters less

③ Pick a random day from 2020
X = max temperature in Berkeley that day
Y = what day of the month it was

$\mathrm{Cov}(X,Y) \approx 0$
(probably not exactly 0)

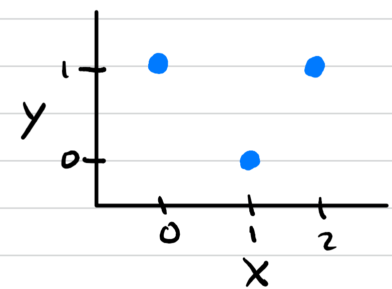
Warning: $\mathrm{Cov}(X,Y)=0$ does not mean X, Y independent!

Example Flip a fair coin 2 times

$X = \#$ of heads $\qquad E(X) = 1$

$$Y = \begin{cases} 1 & \text{if the two flips agree} \\ 0 & \text{if two flips are different} \end{cases} \qquad E(Y) = 1/2$$

Outcome	X	Y	$X-1$	$Y-1/2$	$(X-1)(Y-1/2)$
HH	2	1	1	1/2	1/2
HT	1	0	0	−1/2	0
TH	1	0	0	−1/2	0
TT	0	1	−1	1/2	−1/2



$$\mathrm{Cov}(X,Y) = (1/2)(1/4) + 0\cdot(1/4) + 0\cdot(1/4) + (-1/2)\cdot(1/4) = 1/8 - 1/8 = 0$$

X, Y not independent, but Y is large both when X small & when X large

⑤ Facts about $E(X)$, $Var(X)$, $Cov(X,Y)$

Properties are useful, the proofs are not that important

① Covariance is symmetric

$Cov(Y, X) = Cov(X, Y)$ (order of X and Y doesn't matter)

② Covariance is bilinear

$Cov(X+Y, Z) = Cov(X, Z) + Cov(Y, Z)$ $\quad a = E(X) \quad b = E(Y) \quad c = E(Z)$

$\Rightarrow E(X+Y) = a+b$

$$
\begin{aligned}
Cov(X+Y, Z) &= E((X+Y-(a+b))(Z-c)) \\
&= E(((X-a)+(Y-b))(Z-c)) \\
&= E((X-a)(Z-c) + (Y-b)(Z-c)) \\
&= E((X-a)(Z-c)) + E((Y-b)(Z-c)) \\
&= Cov(X, Z) + Cov(Y, Z)
\end{aligned}
$$

linearity of expectation

$Cov(a \cdot X, Y) = a \cdot Cov(X, Y)$

$Cov(a_1 X_1 + a_2 X_2 + \ldots + a_n X_n, Y) = a_1 Cov(X_1, Y) + \ldots + a_n Cov(X_n, Y)$

③ Variance in terms of covariance

$$Var(X) = Cov(X,X) \qquad a = E(X)$$

$$Cov(X,X) = E((X-a)(X-a)) = E((X-a)^2) = Var(X)$$

④ Covariance in terms of expected value

$$Cov(X,Y) = E(XY) - E(X)E(Y) \qquad a = E(X) \quad b = E(Y)$$

$$\begin{aligned} Cov(X,Y) &= E((X-a)(Y-b)) \\ &= E(XY - aY - bX + ab) \\ &= E(XY) - aE(Y) - bE(X) + ab \\ &= E(XY) - E(X)E(Y) - E(Y)E(X) + E(X)E(Y) \\ &= E(XY) - E(X)E(Y) \end{aligned}$$

⑤ Variance of a sum

$$Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$$

$$\begin{aligned} Var(X+Y) &= Cov(X+Y, X+Y) \\ &= Cov(X, X+Y) + Cov(Y, X+Y) \\ &= Cov(X,X) + Cov(X,Y) + Cov(Y,X) + Cov(Y,Y) \\ &= Var(X) + Var(Y) + 2Cov(X,Y) \end{aligned}$$

5.1 When X, Y are independent

Suppose X and Y are independent

① $E(XY) = E(X)E(Y)$

$$E(XY) = \sum_{a \in range(X)} \sum_{b \in range(Y)} a \cdot b \cdot P(X=a \text{ and } Y=b)$$

"law of unconscious statisticians"

$$= \sum_a \sum_b a \cdot b \cdot P(X=a)P(Y=b)$$

independence

$$= \left(\sum_a a \cdot P(X=a)\right)\left(\sum_b b \cdot P(Y=b)\right) = E(X)E(Y)$$

② $Cov(X,Y) = 0$

$$Cov(X,Y) = E(XY) - E(X)E(Y)$$
$$= E(X)E(Y) - E(X)E(Y) = 0$$

③ $Var(X+Y) = Var(X) + Var(Y)$

$$Var(X+Y) = Var(X) + Var(Y) + 2\,\underbrace{Cov(X,Y)}_{=0}$$

=0 by independence

$$= Var(X) + Var(Y)$$