Problem 3 on pset 3:
parts d) & e) extra credit.

3.2 More on Expectation
3.3 Variance and Standard Deviation

Anna Karlin
Most Slides by Alex Tsun

Agenda

- Linearity of Expectation (LoE) recap
- Law of the Unconscious Statistician (LOTUS)
- Variance
- Independence of random variables
- Properties of variance

Linearity of Expectation (LoE)

Linearity of Expectation: Let Ω be the sample space of an experiment, $X, Y: \Omega \to \mathbb{R}$ be (possibly "dependent") random variables both defined on Ω, and $a, b, c \in \mathbb{R}$ be scalars. Then,

$$E[X+Y] = E[X] + E[Y]$$

and

$$E[\underbrace{aX+b}_{Y}] = aE[X] + b$$

Combining them gives,

$$E[aX + bY + c] = aE[X] + bE[Y] + c$$

Corollary: linearity for sum of lots of r.v.s

$$E(X_1 + X_2 + \ldots + X_n) = E(X_1) + E(X_2) + \ldots + E(X_n)$$

Proof by induction!

Indicator random variable

- For any event A, can define the indicator random variable for A

$$X_A = \begin{cases} 1 & \omega \in A \\ 0 & \text{o.w.} \end{cases}$$

Computing complicated expectations

- Often boils down to finding the right way to decompose the random variable into simple random variables (often indicator random variables) and then applying linearity of expectation.

Linearity is special!

- In general $\boxed{E(g(X)) \neq g(E(X))}$

$Y = X^2$

$\boxed{Y = 1 \text{ with prob } 1}$

~~$E[X \cdot Y] = E[X] \cdot E[Y]$~~
~~$E[X^2] = E[X]^2$~~
~~$E[X/Y] = E[X] / E[Y]$~~
~~$E[\operatorname{asinh}(X)] = \operatorname{asinh}(E[X])$~~
~~.~~
~~.~~
~~.~~

$$X = \begin{cases} 1 & \text{with prob } 1/2 \\ -1 & \text{with prob } 1/2 \end{cases}$$

$E(X) = 1 \cdot \frac{1}{2} + (-1) \cdot \frac{1}{2} = 0$

$E(X^2) = 1 \cdot 1 = 1$

$\boxed{E(X^2) \neq [E(X)]^2}$

	$E(X)$	$E(X^2)$
a)	0	0
b)	**0**	**1**
c)	1	1
d)	-1	1

$E(X^2) = (1)^2 \cdot \frac{1}{2} + (-1)^2 \frac{1}{2}$

$g(x) = x^2$

LINEARITY IS SPECIAL!

- In general $E(g(X)) \neq g(E(X))$

- How DO we compute $\boxed{E(\underbrace{g(X)}_{Y})}$?

pmf of Y

$p_Y(y)$

$E(Y) = \sum_{y \in \Omega_Y} y \, p_Y(y)$

Homeworks of 3 students returned randomly

$Y = X^2 \bmod 2.$

- Each permutation equally likely
- X: # people who get their own homework
- What is $E(X^2 \bmod 2)$?

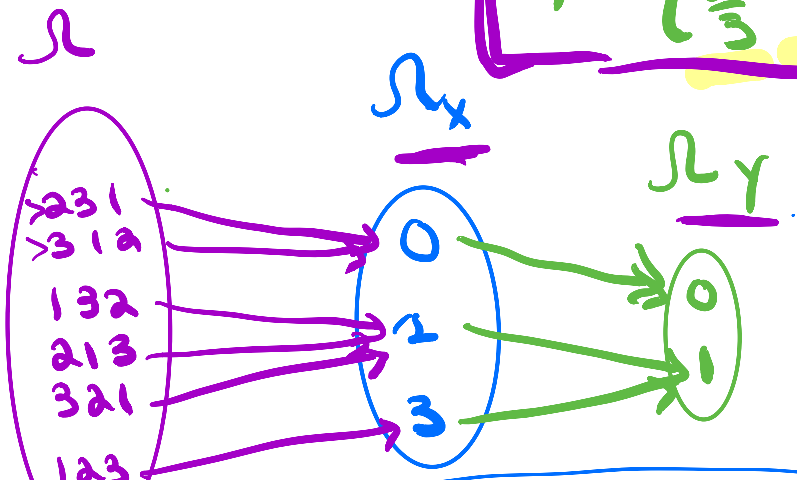
$g(x) = x^2 \bmod 2.$

$$P_X(x) = \begin{cases} \frac{1}{3} & x = 0 \\ \frac{1}{2} & x = 1 \\ \frac{1}{6} & x = 3 \end{cases}$$

$$P_Y(y) = \begin{cases} \frac{1}{3} & y = 0 \\ \frac{2}{3} & y = 1 \end{cases}$$

Prob	Outcome w	X(w)	Y(w)
1/6	1 2 3	3	$3^2 \bmod 2 = 1$
1/6	1 3 2	1	1
1/6	2 1 3	1	1
1/6	2 3 1	0	0
1/6	3 1 2	0	0
1/6	3 2 1	1	1

$Y = X^2 \bmod 2.$



$(0^2 \bmod 2) P_X(0) + (1^2 \bmod 2) P_X(1) + 3^2 \bmod 2 \cdot$

$$E(Y) = \sum_{\omega \in \Omega} g(X(\omega)) P(\omega) = \boxed{\sum_{x \in \Omega_X} g(x) p_X(x)} = \sum_{y \in \Omega_Y} y\, p_Y(y)$$

$Y = g(X)$

Law of the Unconscious Statistician (LOTUS)

Law of the Unconscious Statistician (LOTUS): Let X be a discrete random variable with range Ω_X and $g: D \to \mathbb{R}$ be a function defined at least over Ω_X ($\Omega_X \subseteq D$). Then,

$$\boxed{E[g(X)] = \sum_{b \in \Omega_X} g(b) p_X(b)} =$$

Note that in general, $E[g(X)] \neq g(E[X])$. For example, $E[X^2] \neq (E[X])^2$, or $E[\log(X)] \neq \log(E[X])$.

$$\boxed{E[X^2] = \sum_{k \in \Omega_X} k^2 p_X(k)}$$

Variance (Intuition)

Which game would you rather play? We flip a fair coin.

Game 1:

- If heads, You pay me \$1.
- If Tails, I pay you \$1.

$1 \cdot \frac{1}{2} + (-1)\frac{1}{2}$

Game 2:

- If Heads, you pay me \$1000.
- If Tails, I pay you \$1000.

$1000\frac{1}{2} + (-1000)\frac{1}{2}$

E(my payoff)

	Game 1	Game 2
a)	0	0
b)	0	1000
c)	1	0
d)	1	1000

Variance (Intuition)

How far is a random variable from its mean, on average?

$$X - E[X]$$

$$E(X - E(X)) = E(X) - \overbrace{E(\underbrace{E(X)}_{\text{number.}})}^{= E(X)}$$

$$= 0$$

Variance (Intuition)

How far is a random variable from its mean, on average?

$$|X - E[X]|$$

Variance (Intuition)

How far is a random variable from its mean, on average?

$$E[\,|X - E[X]|\,] \neq \sigma_X$$

$$E[\,(X - E[X])^2\,]$$

Variance and Standard Deviation (SD)

Variance: The variance of a random variable X is

$E(aX) = aE(X)$

$$Var(X) = E[(X - E[X])^2]$$

The variance is always nonnegative since we take an expectation of a nonnegative random variable $(X - E[X])^2$. We can also show that for any scalars $a, b \in \mathbb{R}$,

$$Var(aX + b) = a^2\, Var(X)$$

$$E[(X - \underbrace{E(X)}_{\mu})^2] = E[X^2 - 2\mu X + \mu^2]$$

$$= E[X^2] + \underbrace{E[-2\mu X]}_{\substack{-2\mu E(X) \\ -2\mu^2}} + \underbrace{E[\mu^2]}_{\mu^2}$$

$$\boxed{= E[X^2] - \mu^2}$$

Variance and Standard Deviation (SD)

Variance: The variance of a random variable X is

More Useful

$$Var(X) = E[(X - E[X])^2] = \boxed{E[X^2] - (E[X])^2}$$

The variance is always nonnegative since we take an expectation of a nonnegative random variable $(X - E[X])^2$. We can also show that for any scalars $a, b \in \mathbb{R}$,

$$Var(aX + b) = a^2\, Var(X)$$

Variance and Standard Deviation (SD)

Variance: The variance of a random variable X is

MORE USEFUL

$$Var(X) = E[(X - E[X])^2] = \boxed{E[X^2] - (E[X])^2}$$

The variance is always nonnegative since we take an expectation of a nonnegative random variable $(X - E[X])^2$. We can also show that for any scalars $a, b \in \mathbb{R}$,

$$Var(aX + b) = a^2\, Var(X)$$

Standard Deviation (SD): The standard deviation of a random variable X is

$$\sigma_X = \sqrt{Var(X)}$$

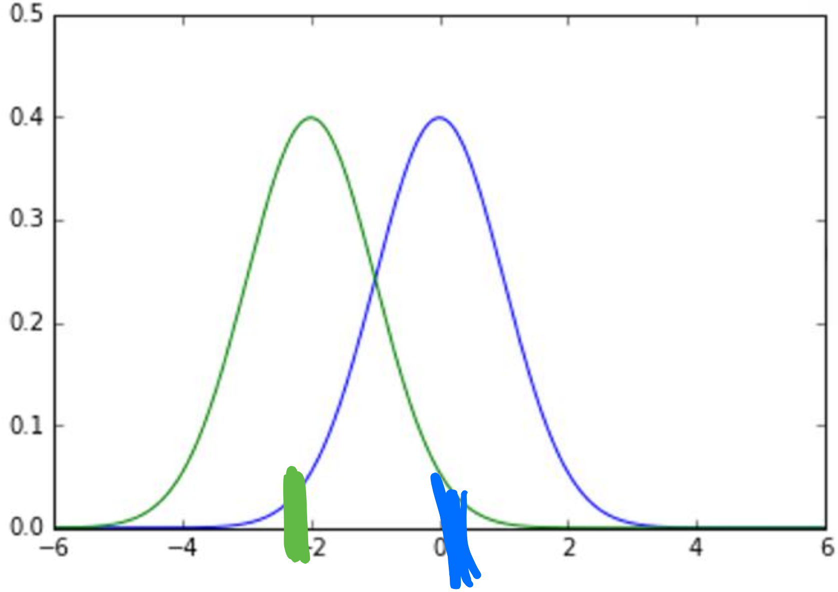
We want this because the units of variance are squared in terms of the original variable X, and this "undo's" our squaring, returning the units to the same as X.

Variance (Property)

Var(aX)

$$Var(aX+b) = a^2 Var(X)$$

$$Var(X+b) = Var(X)$$



Variance (Property)

$$Var(aX + b) = a^2Var(X)$$

$$\boxed{Var(aX)} = E[(aX)^2] - (E[aX])^2 = E[a^2X^2] - (aE[X])^2$$

$$= a^2E[X^2] - a^2(E[X])^2 = a^2(E[X^2] - E[X]^2) = \boxed{a^2Var(X)}$$

Variance (Example)

Let X be the outcome of a fair 6-sided die roll. What is $Var(X)$?

$$Var(X) = E[X^2] - E[X]^2$$

$$E[X] = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + \cdots + 6 \cdot \frac{1}{6} = 3.5$$

$$E[X^2] = 1^2 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6} + 3^2 \cdot \frac{1}{6} + \cdots + 6^2 \cdot \frac{1}{6} = \frac{91}{6}$$

LOTUS

$$Var(X) = E[X^2] - E[X]^2 = \frac{91}{6} - (3.5)^2 = \frac{35}{12}$$

Variance

Which game would you rather play? We flip a fair coin.

Game 1:

- If heads, You pay me \$1.
- If Tails, I pay you \$1.

Game 2:

- If Heads, you pay me \$1000.
- If Tails, I pay you \$1000.

X = payoff in game

$\mathrm{Var}(X)$

	Game 1	Game 2
a)	0	0
b)	1	1000
c)	1	1,000,000

$$\mathrm{Var}(X) = E(X^2) - [\underbrace{E(X)}_{=0}]^2$$

$$= \underline{E(X^2)}$$

$$\boxed{1^2\cdot\tfrac{1}{2} + (-1)^2\cdot\tfrac{1}{2} = 1}$$

$$\boxed{(1000)^2\,\tfrac{1}{2} + (-1000)^2\,\tfrac{1}{2}}$$

Game 1 Game 2

$\sigma_X = 1$ $\sigma_X = \sqrt{Var(X)}$ $\sigma_X = 1000$

IN GENERAL, VAR(X+Y) $\neq$ VAR(X) + VAR(Y)

Example 1:

- $X = \boxed{\pm 1 \text{ each with prob } ½}$
- E(X)= 0 Var(X) = 1
- How about $\boxed{Y = -X}$

$E(Y) = 0$ $Var(Y) = 1$

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

$Var(X + \underset{-X}{Y}) = Var(0) = 0$

Example 2: What is Var(X+X)?

$\neq 2Var(X)$

$Var(2X) = 2^2 Var(X)$

$Var(aX) = a^2 Var(X)$

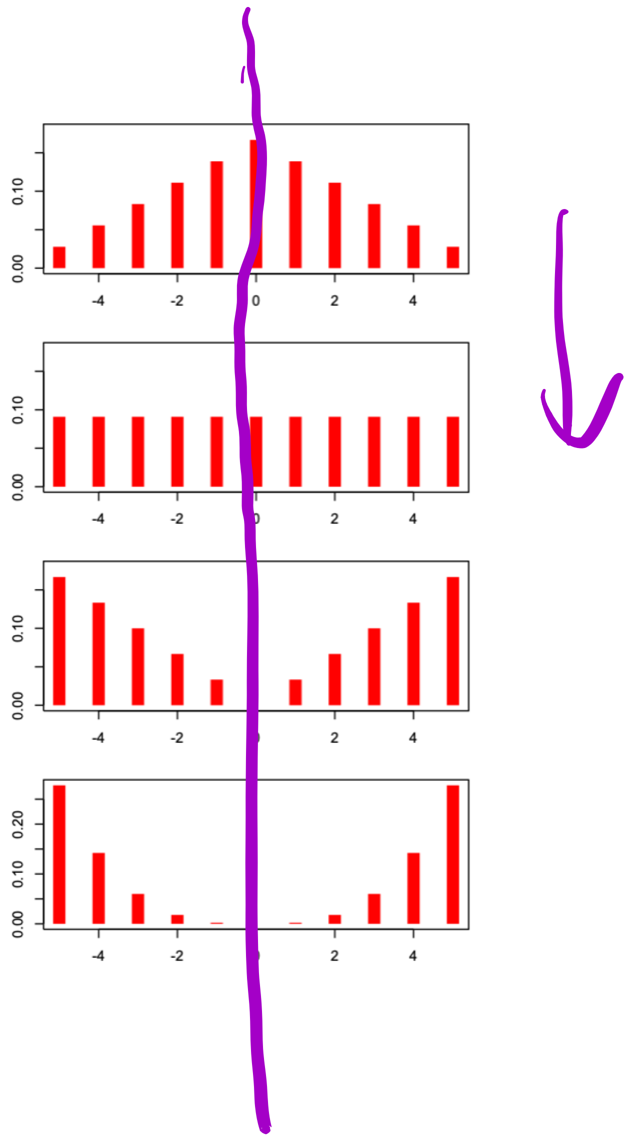
Variance in pictures

$\sigma^2 = 5.83$

$\sigma^2 = 10$

$\sigma^2 = 15$

$\sigma^2 = 19.7$



Random Picture

Random variables and independence

Random variable X and event E are independent if the event E is independent of the event {X=x} (for any fixed x), i.e.

$$\forall x \; P(X = x \text{ and } E) = P(X=x) \cdot P(E)$$

Two random variables X and Y are independent if the events {X=x} and {Y=y} are independent for any fixed x, y, i.e.

$$\forall x, y \; P(X = x \text{ and } Y=y) = P(X=x) \cdot P(Y=y)$$

$$P(X=x \mid Y=y) = P(X=x)$$

Intuition as before: knowing X doesn't help you guess Y or E and vice versa.

Example

Random variable X and event E are independent if the event E is independent of the event {X=x} (for any fixed x), i.e.

$$\forall x\ P(X = x \text{ and } E) = P(X=x) \cdot P(E)$$

$\frac{1}{2}$.

Example: Let X be number of heads in n independent coin flips. Let E be the event that the number of heads is even.

$$Pr(X=k) = \binom{n}{k} \frac{1}{2^n}$$

X & E indep?

E happen

a) yes

$$Pr(X=5 \mid E) \overset{?}{=} P(X=k)$$

$\boxed{\text{b) no}}$

"
0

"

Example

Two random variables X and Y are independent if the events {X=x} and {Y=y} are independent (for any fixed x, y), i.e.

$$\forall x, y \; P(X = x \text{ and } Y=y) = P(X=x) \cdot P(Y=y)$$

Example: Let X be number of heads in first n of 2n independent coin flips, Y be number in the last n flips, and let Z be the total.

$n>2$

$$\Pr(X=n \mid Z=2) = 0$$

$$\neq P(X=n)$$

	X & Y indep.	X & Z indep.
a)	no	no
b)	no	yes
c)	yes	no
d)	yes	yes

Example continued

Example: Let X be number of heads in first n of 2n independent coin flips, Y be number in the last n flips, and let Z be the total.

$$P(X = j) = \binom{n}{j} 2^{-n}$$

$$P(Y = k) = \binom{n}{k} 2^{-n}$$

$$P(X = j \wedge Y = k) = \binom{n}{j}\binom{n}{k} 2^{-2n} = P(X = j)P(Y = k)$$

Important facts about independent random variables

Theorem: If X & Y are independent, then $E[X \cdot Y] = E[X] \cdot E[Y]$

Theorem: If X and Y are independent, then
$Var[X + Y] = Var[X] + Var[Y]$

Corollary: If $X_1 + X_2 + \dots + X_n$ are mutually independent then
$Var[X_1 + X_2 + \dots + X_n] = Var[X_1] + Var\ [X_2] + \dots + Var[X_n]$

E[XY] for independent random variables

- Theorem: If X & Y are independent, then E[X•Y] = E[X]•E[Y]
- Proof:

Let $x_i, y_i, i = 1, 2, \ldots$ be the possible values of X, Y.

$$\begin{aligned} E[X \cdot Y] &= \sum_i \sum_j x_i \cdot y_j \cdot P(X = x_i \wedge Y = y_j) \\ &= \sum_i \sum_j x_i \cdot y_j \cdot P(X = x_i) \cdot P(Y = y_j) \\ &= \sum_i x_i \cdot P(X = x_i) \cdot \left(\sum_j y_j \cdot P(Y = y_j) \right) \\ &= E[X] \cdot E[Y] \end{aligned}$$

independence

Note: *NOT* true in general; see earlier example $E[X^2] \neq E[X]^2$

x

Variance of a sum of independent r.v.s

Theorem: If X and Y are independent, then

Var[X + Y] = Var[X] + Var[Y]

Proof:

$$
\begin{aligned}
&Var[X+Y] \\
&= E[(X+Y)^2] - (E[X+Y])^2 \\
&= E[X^2 + 2XY + Y^2] - (E[X] + E[Y])^2 \\
&= E[X^2] + 2E[XY] + E[Y^2] - ((E[X])^2 + 2E[X]E[Y] + (E[Y])^2) \\
&= E[X^2] - (E[X])^2 + E[Y^2] - (E[Y])^2 + 2(E[XY] - E[X]E[Y]) \\
&= Var[X] + Var[Y] + 2(E[X]E[Y] - E[X]E[Y]) \\
&= Var[X] + Var[Y]
\end{aligned}
$$

using independence of x & y

x

Independent vs dependent r.v.s

- Dependent r.v.s can reinforce/cancel/correlate in arbitrary ways.
- Independent r.v.s are, well, independent.

Example:

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

X_i is indicator r.v. with probability 1/2 of being 1.

versus

$W = n\, X_1$

	$E(Z)$	$E(W)$
a)	$\frac{n}{2}$	n
b)	n	$\frac{n}{2}$
c)	0	0
d)	$\frac{n}{2}$	$\frac{n}{2}$