Lec 14 10/26/17 Math 289

Custom r.v.'s

Roulette in America. Bet \$1 on black. payout is 1:1.

$$X \sim \begin{cases} \$1 & \text{w.p. } \frac{18}{38} \\ -\$1 & \text{w.p. } \frac{20}{38} \end{cases} \qquad \mu = \$1\,\frac{18}{38} + -\$1\,\frac{20}{38} = -\$0.053$$

If I play many many times... my avg. winnings will be ...

$$\bar{X} \to \mu \quad \text{Law of Large \#'s.}$$

$\lim_{n \to \infty} T = -\infty$ If I play forever... I lose all my money!

Bet on lucky #7. Payout: 35:1

$$X \sim \begin{cases} \$35 & \text{wp } \frac{1}{38} \\ -\$1 & \text{wp } \frac{37}{38} \end{cases}$$

Bet on Dozen 1-12. Payout 2:1

$$X \sim \begin{cases} \$2 & \text{wp } \frac{12}{38} \\ -\$1 & \text{wp } \frac{26}{38} \end{cases}$$

All bets the same expectation. Story...

Europe bet on black. Same payout but...

$$X \sim \begin{cases} \$1 \\ -\$1 \end{cases}$$ $E(X) = -\$0.027$ "much fairer"

Def: "Fair Game" $E[X] = 0$

Basic r.v. Transformations

Uber example

$P(\text{traffic}) = 0.3$ If traffic → street, else Van Wyck

$$W \sim \begin{cases} 7\text{ min} & \text{wp } 0.7 \\ 12\text{ min} & \text{wp } 0.3 \end{cases}$$ (not a Bernoulli)

$E[W] = 0.7 \cdot 7\text{min} + 0.3 \cdot 12\text{ min} = \cancel{7.65}\ 8.5\text{ min}$ Interpret...

Over many trips, my average time in the taxi is $\approx$ ~~7.65~~ 8.5 min. but any given trip...

3

TOceil this further... $W \sim \begin{cases} 7\text{ min} & \text{wp } 0.7 \\ 12\text{ min} & \text{wp } 0.3 \end{cases}$ $E[W] = 8.5$ min

Uber Charges \$0.40/min

$B = \$0.40/\text{min} \cdot W$ } a transformation of a r.v.

$B \sim \begin{cases} \$2.80 & \text{wp } 0.7 \\ \$4.80 & \text{wp } 0.3 \end{cases}$ $E[B] = \$2.80 \cdot 0.7 + \$4.80 \cdot 0.3 = \$3.12$

figured out this PMF through common sense... Math 621 covers the theory to do this axiomatically

~~Is there an easier way? B is a function of another variable.~~

~~Note that $B = g(W) = 0.4W$ the function is simple: it's a "scale". def: $g(x) = a f(x)$ s.t. $a \in \mathbb{R}$, g is a scaling function.~~

Is $E(B)$ related to $E(W)$ through g?

~~Beyond scope of course... "FAKE PROOF"~~

~~$E(X) := \int_\Omega X(\omega) P(d\omega)$ A "Lebesgue" integral on a measurable function~~

~~If X is discrete, $Supp(X) = \{x_1, x_2, x_3, \ldots\}$ Recall $\int_1^3 f(x)dx = \int_1^2 f(x)dx + \int_2^3 f(x)dx$~~

~~$X(\omega) \in Supp(X)$ $\forall \omega \in \Omega$ by def.~~

~~$= \int_{\{\omega : X(\omega) = x_1\}} X(\omega) P(d\omega) + \int_{\{\omega : X(\omega) = x_2\}} X(\omega) P(d\omega) + \ldots$ Matches the def $= x_1 \int P(d\omega) + x_2 \int P(d\omega) + \ldots$ $x_1 P(X = x_1) + x_2 P(X = x_2) + \ldots$~~

$$E[g(X)] = \int_{\Omega} g(X(\omega))\, P(d\omega)$$

$$= \int_{\{\omega : X(\omega) = x_1\}} g(X(\omega))\, P(d\omega) + \int_{\{\omega : X(\omega) = x_2\}} g(X(\omega))\, P(d\omega) + \dots$$

$$= g(x_1)\, P(X = x_1) + g(x_2)\, P(X = x_2) + \dots$$

$$= \sum_{x \in \mathrm{Supp}[X]} g(x)\, p(x)$$

Rule:

$$E[g(X)] = \sum_{x \in \mathrm{Supp}(X)} g(x)\, p(x)$$

Proof needs Ω and ω's... don't need to do this...

What if $\underbrace{Y = aX}_{g(X)}$ scaling just like $B = 0.4W$

$$\underset{\text{by def p.t.}}{E(Y)} = E(aX) = \sum_{x \in \mathrm{Supp}(X)} ax\, p(x) = a \underbrace{\sum_{x \in \mathrm{Supp}(X)} x\, p(x)} = a\,E(X)$$

So.. $E(B) = 0.4\, E(W) = \$3.12$

there is a base fee... \$3

$T = B + \$3$

$E(T)$?

$Y = X + c$

$$E(Y) = E(X + c) = \sum_{x \in \mathrm{Supp}(X)} (x + c)\, p(x) = \sum x\, p(x) + \sum c\, p(x) = E(X) + c \sum p(x) = E(X) + c$$

RULE

$$\boxed{Y = aX + c \Rightarrow E(Y) = aE(X) + c}$$

$$E(Y) = E(X) + \$3 = \$3.12 + 3 = \boxed{\$6.12}$$

$X \sim \text{Bin}\left(6, \frac{1}{2}\right)$, $Y = X^2$

$E(X) = 6 \cdot \frac{1}{2} = 3$ $\quad$ $E(Y) = ?$

? No!

$$E(Y) = \sum x^2 p(x) = \sum x^2 \binom{6}{x} \frac{1}{2^6} = 12.5$$

gotta do it...

Note: $E(g(X)) \neq g(E(X))$ generally...

~~However... $\text{Quantile}[g(X), p] = g(\text{Quantile}[X, p])$~~

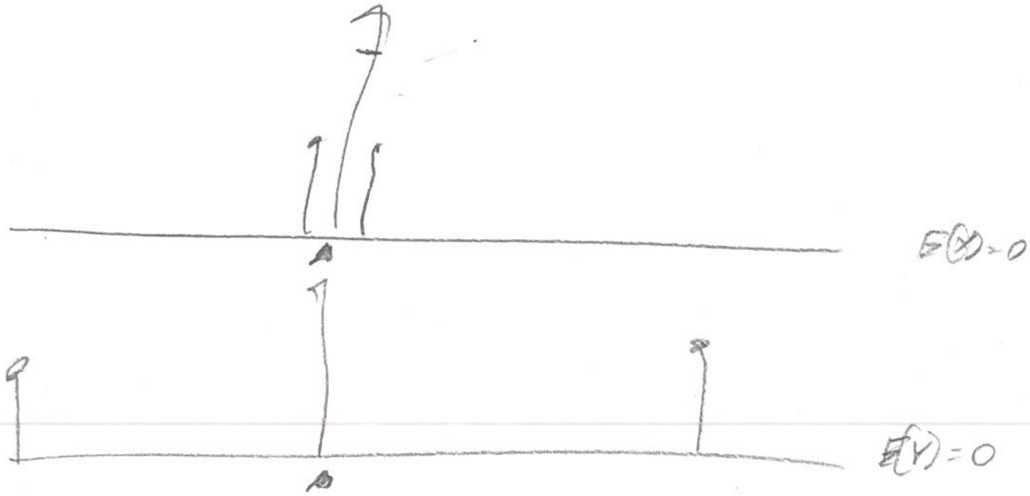
~~Why? $Q[X, p] = \min\{x : F(x) \geq p\}$~~

~~$Q[g(X), p] = \min\{g(x) : P(X \leq x) \geq p\}$~~

~~$P(g(X) \leq g(x)) \geq p$~~

$X \sim \text{Rademacher} := \begin{cases} 1 & \text{w.p. } \frac{1}{2} \\ -1 & \text{w.p. } \frac{1}{2} \end{cases}$ $\quad E(X) = (1)\cdot\frac{1}{2} + (-1)\cdot\frac{1}{2} = 0$

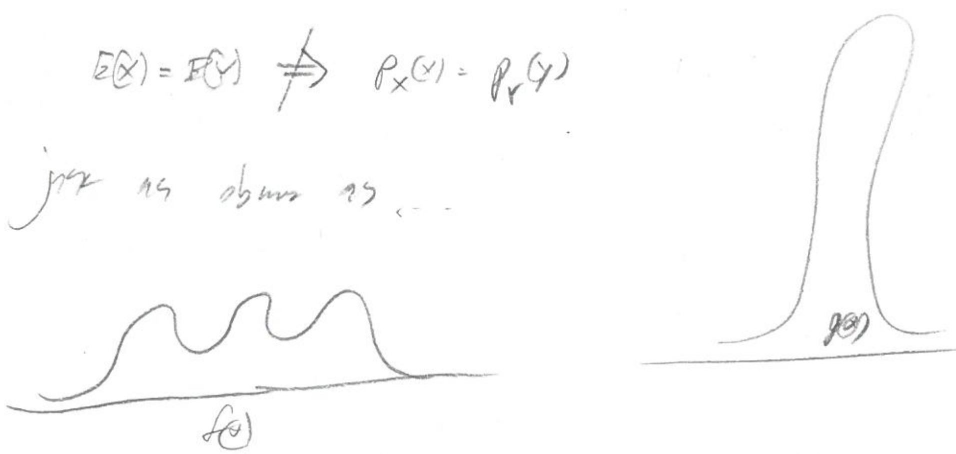
$Y = 10X \Rightarrow E(Y) = 10E(X) = 10 \cdot 0 = 0$



PMF's

$E(X) = E(Y) \not\Rightarrow p_X(x) = p_Y(y)$

just as obvious as ...



$\int_{\mathbb{R}} f(x)\,dx = 17$ $\qquad \int_{\mathbb{R}} g(x)\,dx = 17$ $\quad$ but $f(x) \neq g(x)$

Of course many different shapes can yield the same area!
f is more "diffuse" or "spread" or "varied" than g

We need a metric to capture part of this difference btwn r.v.'s X, Y.

How about something to do with how far the support is away from pivot

In X, close to pivot; in Y, far from pivot. Y more dispersal.

"error function", "loss function"

$e(X, \mu) = X - \mu$ but it doesn't care if it's $\oplus$ or $\ominus$

$e(X, \mu) = |X - \mu|$ bad for taking derivative! Hard to prove things...

$e(X, \mu) = (X - \mu)^2$ "squared error loss" or "L2 error".

let $L := (X - \mu)^2$ (r.v.'s)

$E[L]$ is what? The expected sqd error distance from μ_X.

It means "how far away on avg is a realization of X from its own mean"?

$$E[L] = E[g(X)] = \sum_{x \in Supp(X)} g(x)\, p(x) = \sum_{x \in Supp(X)} (x - \mu)^2 p(x)$$

e.g.

$X \sim$ Rademacher $\Rightarrow \mu = E(X) = 0$

$$E[L] = ((-1) - (0))^2 p(-1) + ((1) - (0))^2 p(1) = 1 \cdot 0.5 + 1 \cdot 0.5 = \boxed{1}$$

$Y = 10X \Rightarrow \mu = 0$

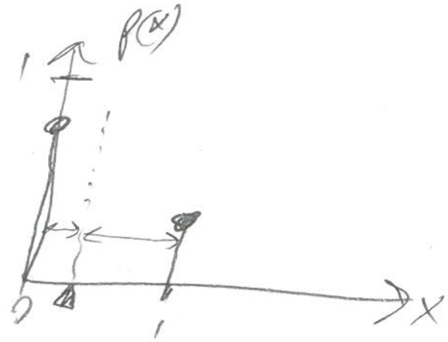
$$E[L] = (-10 - 0)^2 p(-1) + (10 - 0)^2 p(1) = 100 \cdot 0.5 + 100 \cdot 0.5 = \boxed{100}$$

We did our job.. It's thing so useful has its own name

Variance

$$\sigma^2 := Var(X) := E[(X - \mu)^2]$$

$X \sim \text{Bern}\left(\frac{1}{3}\right)$



$E(X) = \frac{1}{3}$

$Var(X) = \left(0 - \frac{1}{3}\right)^2 \frac{2}{3} + \left(1 - \frac{1}{3}\right)^2 \cdot \frac{1}{3}$

$= \frac{1}{9} \cdot \frac{2}{3} + \frac{4}{9} \cdot \frac{1}{3} = \frac{6}{27} = \frac{2}{9} = .259$

genl formula?

$X \sim \text{Bern}(p)$

$Var(X) = ((0) - (p))^2 (1-p) + ((1) - (p))^2 p$

$= p^2(1-p) + (1-p)^2 p$

$= (1-p)\left(\cancel{p^2} + \overbrace{(1-p)p}^{p - \cancel{p^2}}\right)$

$= \boxed{p(1-p)}$

Roulette: Bet on lucky #7

$X_7 \sim \begin{cases} \$35 & \text{wp } \frac{1}{38} \\ -\$1 & \text{wp } \frac{37}{38} \end{cases}$ $\mu = -\$0.053$

$Var(X_7) = (\$35 - -\$0.053)^2 \frac{1}{38} + (-\$1 - -\$.053)^2 \cdot \frac{37}{38}$

$= 33.207 \ \$^2$

" " Bet on Black

$X_B \sim \begin{cases} \$1 & \text{wp } \frac{18}{38} \\ -\$1 & \text{wp } \frac{20}{38} \end{cases}$ $\mu = -\$0.053$

$Var(X_B) = (\$1 - -\$0.053)^2 \frac{18}{38} + (-\$1 - -\$0.053)^2 \frac{20}{38}$

$= 0.997 \ \$^2$

$\bar{X}_A \to \mu$

$\bar{X}_B \to \mu$

which goes faster?
the one with lowest variance...
(we will prove why later)

Ouch! $\$^2$ ... has no meaning!

Easiest way to solve $\sqrt{\$^2} = \$$ ← nice and interpretable

let $\sigma := SE(X) := \sqrt{Var(X)} = \sqrt{\sigma^2}$

"Standard error" or "standard deviation"

$\sigma_A = \$5.79$, $\sigma_B = \$1.00$

σ^2 interpretation

σ interpretation ... not so clear!

It is not an expectation... it is just a practical strategy to report spread.

But... it will be useful later...

$T_2 := X_1 + X_2$ $\quad E(T) = \sum_{t \in Supp(T)} t\, p(t)$

$Supp(X_1)$	$Supp(X_2)$	T	$p(t)$
x_1		⋮	⋮
x_2			
x_3			
⋮			

aggregate

Hard to get

Is there a better way?

X_1, X_2 indep $p(x_1, x_2) = p(x_1)\,p(x_2)$