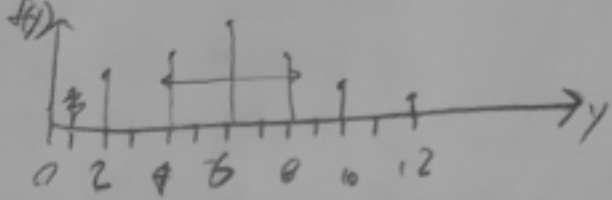
Lecture 15 11/4/14

$X \sim \text{Binom}(6, \frac{1}{2})$ $\qquad$ $Y = 2X$



$E[X] = np = 3$

$Var[X] = np(1-p) = 6 \cdot \frac{1}{2} \cdot \frac{1}{2} = 1.5$

$SD[X] = \sqrt{1.5} \approx 1.22$

$E[Y] = E[2X] = 2E(X) = 6$

$Var[Y] = ?$ we discussed previously $Var(X+c) = Var(X)$

It appears that $SD[Y] = 2\,SD[X]$ since everything is scaled up by some factor. Is this true?

Def: $Var(g(X)) := E(g(X) - E[g(X)])^2 := \sum_{x \in Supp(X)} (g(x) - E[g(X)])^2 f(x)$

$Var(aX) = \sum_{x \in Supp(X)} (ax - a\mu)^2 f(x) = a^2 \underbrace{\sum_{x \in Supp X} (x-\mu)^2 f(x)}_{Var(X)} = a^2 \sigma^2$

$E[aX] = a\mu$

Rule:

$\Rightarrow \boxed{Var[aX] = a^2 Var(X)}$, $\boxed{SD(aX) = \sqrt{a^2 Var(X)} = |a|\,SD(X)}$ since $\sqrt{(-3)^2} = 3 \neq -3$

$\boxed{Var(aX) = a^2 \sigma^2}$ $\quad$ or $\boxed{SD(aX) = a|\sigma|}$

$Z = \frac{X - \mu}{\sigma}$

Z is X "standardized"

$E[Z] = E\left[\frac{X-\mu}{\sigma}\right] = \frac{1}{\sigma} E[X - \mu] = \frac{1}{\sigma}(E(X) - \mu) = 0$ ✓

$Var[Z] = Var\left[\frac{X-\mu}{\sigma}\right] = Var\left[\frac{X}{\sigma} - \frac{\mu}{\sigma}\right] = Var\left[\frac{X}{\sigma}\right] = \frac{1}{\sigma^2} Var(X) = 1$ ✓

A "standardized r.v." has mean 0 and Var/SD = 1

$T_2 = X_1 + X_2$

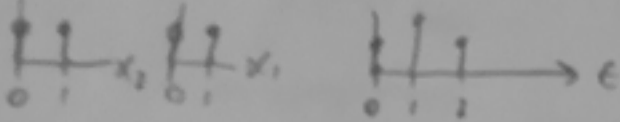
Looking for $E[T]$

By def $E(T) = \sum_{t \in Supp(T)} t\, f_T(t)$ → What is this?

$Supp[T] = \{x + y : x \in Supp(X), y \in Supp(Y)\}$

$X_1 \sim \text{Bernoulli}(\frac{1}{2})$

$X_2 \sim \text{Bernoulli}(\frac{1}{2})$



Shape is not the same!

No easy way to figure this one!

$P(T=0) = P(X_1=0)\,P(X_2=0) = (\frac{1}{2})^2 = \frac{1}{4}$

$P(T=2) = P(X_1=1)\,P(X_2=1) = (\frac{1}{2})^2 = \frac{1}{4}$

$P(T=1) = P(X_1=0)\,P(X_2=1) + P(X_1=1)\,P(X_2=0) = (\frac{1}{2})^2 + (\frac{1}{2})^2 = \frac{1}{2}$

$$f_T(t) = \sum_{\substack{x_1, x_2 \\ \text{s.t. } t = x_1 + x_2}} f(x_1, x_2)$$

joint mass function (JMF)

$P(X_1 = x_1, X_2 = x_2)$

both things happen "joint" = intersection

If X_1, X_2 ind,

$$f(x_1, x_2) = f_{X_1}(x_1)\, f_{X_2}(x_2) \qquad P(A,B) = P(A)P(B)$$

$$f_T(t) = \sum_{\substack{x_1, x_2 \\ \text{s.t. } t = x_1 + x_2}} f_{X_1}(x_1)\, f_{X_2}(x_2) = \sum_{x_1 \in \text{Supp}(X_1)} f_{X_1}(x_1)\, f_{X_2}(t - x_1) := f_{X_1}(x_1) * f_{X_2}(x_2)$$

"Convolution"

If t known, x_1 free, x_2 is pinned!

$$\Rightarrow E(T) = \sum_{t \in \text{Supp}(T)} t \sum_{x_1 \in \text{Supp}(X_1)} f_{X_1}(x_1)\, f_{X_2}(t - x_1)$$

not pretty! But at least we learned convolution!

This strategy fails to give an elegant answer... but it is a very elegant if you use mgf's (after midterm).

Need another way...

$$E[T] = E[\underbrace{X+Y}_{g(X,Y)}]$$

an arbitrary function of two r.v.'s

JMF again!

$$E[g(X,Y)] = \sum_{(x,y)\in Supp(X,Y)} g(x,y)\, f(x,y) = \sum_{x\in Supp(X)}\sum_{y\in Supp(Y)} g(x,y)\, f(x,y)$$

as long as this returns 0 for invalid (x,y) in $Supp(X,Y)$, okay!

$$E[X+Y] = \sum_{x}\sum_{y} (x+y)\, f(x,y)$$

shorthand

$$= \sum_x\sum_y x\, f(x,y) + \sum_x\sum_y y\, f(x,y) \quad \text{(algebra)}$$

$$= \sum_x x \sum_y f(x,y) + \sum_y y \sum_x f(x,y) \quad \text{(switch order of sums, factor out constants)}$$

What is $\sum_y f(x,y)$?

Imagine... $Supp(X) = Supp(Y) = \{1,2,3\}$

$f(x,y) = \frac{1}{36}xy$ Valid JMF?

$\sum_{x=1}^{3}\sum_{y=1}^{3} \frac{1}{36}xy = 1$ ✓

$f(x,y) > 0 \ \forall (x,y) \in Supp(X,Y)$ ✓

x \ y	1	2	3	row totals
1	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{6}{36}$
2	$\frac{2}{36}$	$\frac{4}{36}$	$\frac{6}{36}$	$\frac{12}{36}$
3	$\frac{3}{36}$	$\frac{6}{36}$	$\frac{9}{36}$	$\frac{18}{36}$
col totals	$\frac{6}{36}$	$\frac{12}{36}$	$\frac{18}{36}$	

a function of x

$$\sum_y f(x,y) = f(x)$$

marginal mass function!

$= P(X=x)$

this is "marginalizing out" Y

likewise $\sum_x f(x,y) = f(y)$

"marginalizing out" X

$$= \sum_x x\, f(x) + \sum_y y\, f(y)$$

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

Works <u>all</u> the time!

Makes sense? For X,Y indep... sure!

For X,Y dependent... not so clear to me...

Using induction $E[X_1 + \dots + X_n] = \sum_{i=1}^{n} E[X_i]$ *

If $X_1, \dots, X_n$ equal in distr (do not need iid) all w/ $E[X]$

$\Rightarrow E[X_1 + \dots + X_n] = n\mu$ * n times expectation of one of them...

$X_1, \dots, X_n$ iid Bern(p) $T_n = \sum X_i \sim$ Binomial(n,p) $E[T_n] = E[X_1 + \dots + X_n] = \boxed{np}$ Cool! Also seen with general/any binomial.

$Var[X+Y] = ?$ $Var[g(X,Y)] := E\left[\left(g(X,Y) - E[g(X,Y)]\right)^2\right] = \sum_x \sum_y (g(x,y) - E[g(X,Y)])^2 f(x,y)$ ← PMF

$$= \sum_x \sum_y \Big(\overbrace{x+y-(\mu_X+\mu_Y)}^{x+y-\mu_X-\mu_Y}\Big)^2 f(x,y)$$

$$= \sum_x \sum_y \left(x^2 + y^2 + \mu_X^2 + \mu_Y^2 + 2xy - 2x\mu_X - 2x\mu_Y - 2y\mu_X - 2y\mu_Y + 2\mu_X\mu_Y\right) f(x,y)$$

Start grouping things...

$$= \underbrace{\sum_x \underbrace{(x^2 - 2x\mu_X + \mu_X^2)}_{(x-\mu_X)^2} \underbrace{\sum_y f(x,y)}_{f(x)}}_{\underbrace{Var[X]}_{\sigma_X^2}} + \underbrace{\sum_y \underbrace{(y^2 - 2y\mu_Y + \mu_Y^2)}_{(y-\mu_Y)^2} \underbrace{\sum_x f(x,y)}_{f(y)}}_{\underbrace{Var[Y]}_{\sigma_Y^2}} + \underbrace{2\sum_x \sum_y \underbrace{(xy - x\mu_Y - y\mu_X + \mu_X\mu_Y)}_{(x-\mu_X)(y-\mu_Y)} f(x,y)}_{2\underbrace{E[(X-\mu_X)(Y-\mu_Y)]}_{2\,Cov[X,Y]}}$$

$2\,Cov[X,Y]$ (σ_{XY}) by definition

$Var[X+Y] = \sigma_X^2 + \sigma_Y^2 + 2\sigma_{XY}$ * need to know $\sigma_{XY} \neq 0$ for X, Y dependent!

Examine this again... $E[(X-\mu_X)(Y-\mu_Y)] = E(XY) - E(\mu_X Y) - E(\mu_Y X) + E[\mu_X\mu_Y]$

$= E(XY) - \mu_X\mu_Y - \mu_X\mu_Y + \mu_X\mu_Y$

$= E(XY) - \mu_X\mu_Y$

an equivalent def of $Cov(X,Y)$

just like $Var(X) = E(X^2) - \mu^2$

What is $E[XY]$? Again $g(X,Y) = XY$ $\quad E[g(X,Y)] = \sum_{(x,y) \in \operatorname{Supp}(X,Y)} g(x,y) f(x,y)$

$$E[XY] = \sum_x \sum_y xy\, f(x,y) \quad \text{Can't budge!}$$

But what if X ind Y? $f(x,y) = f(x) f(y)$ did this before...

$$E[XY] = \sum_x \sum_y xy\, f(x) f(y) = \sum_x x f(x) \underbrace{\sum_y y f(y)}_{\mu_Y} = \mu_Y \underbrace{\sum_x x f(x)}_{\mu_X} = \mu_X \mu_Y$$

$\Rightarrow \operatorname{Cov}[X,Y] = E[XY] - \mu_X \mu_Y = \mu_X \mu_Y - \mu_X \mu_Y = 0$ if X, Y indep!

$\Rightarrow \operatorname{Var}[X_1 + X_2] = \operatorname{Var}[X_1] + \operatorname{Var}[X_2]$ * if X_1, X_2 are indep.

$\Rightarrow \operatorname{Var}[X_1 + \ldots + X_n] = \sum_{i=1}^{n} \operatorname{Var}[X_i]$ *

if $X_1, \ldots X_n$ iid need iid now.

$\operatorname{Var}[X_1 + \ldots + X_n] = n\sigma^2$ *

Point Estimation Theory

"Estimator", $\bar{x}$ "estimate"

$E[\bar{X}_n] = E\left[\frac{X_1 + \ldots + X_n}{n}\right] = \frac{1}{n} \sum E[X_i]$ *, if identically drawn

$= \frac{1}{n} n\mu = \boxed{\mu}$ * A property called "unbiasedness" the estimator on average yields what you want

$\operatorname{Var}[\bar{X}_n] = \operatorname{Var}\left[\frac{X_1 + \ldots + X_n}{n}\right] = \frac{1}{n^2} \operatorname{Var}[X_1 + \ldots + X_n]$ *

if iid

$= \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}$ *

As $n \to \infty$ $\operatorname{Var}[\bar{X}_n]$ drops to zero

MIDTERM II ↑

FINAL ↓