Lecture 16 11/3/15 Math 241

What if not independent?

$$E[X_1+X_2] = \sum_{x_1}\sum_{x_2}(x_1+x_2)\,p(x_1,x_2) = \sum_{x_1}\sum_{x_2} x_1\,p(x_1,x_2) + \sum_{x_1}\sum_{x_2} x_2\,p(x_1,x_2)$$

$$= \sum_{x_1} x_1 \underbrace{\sum_{x_2} p(x_1,x_2)}_{?} + \sum_{x_2} x_2 \underbrace{\sum_{x_1} p(x_1,x_2)}_{?}$$

Imagine $Supp[X_1] = \{1, 7, 19\}$

$Supp[X_2] = \{5, 23, 88\}$

$2+10+4 = \frac{16}{30} = \frac{8}{15}$

$X_2 \backslash X_1$	1	7	19	
5	$\frac{1}{15}$	$\frac{1}{3}$	$\frac{2}{15}$	$\frac{16}{30}$
23	$\frac{1}{30}$	$\frac{1}{10}$	$\frac{1}{30}$	$\frac{5}{30}$
88	$\frac{1}{30}$	$\frac{1}{5}$	$\frac{1}{15}$	$\frac{9}{30}$
	$\frac{4}{30}$	$\frac{19}{30}$	$\frac{2}{30}$	1

Valid JMF $\sum_{x_1}\sum_{x_2} p(x_1,x_2) = 1$

Margin

"Marginal PMF"

$$P(X_2=5) = P(X_2=5, X_1=1) + P(\quad) + P(\quad)$$
$$\frac{2}{5}$$

X_1 ind of X_2?

No...

$\frac{1}{5} = \frac{\frac{1}{30}}{\frac{1}{6}} = P(X_1=1 \mid X_2=88) \neq P(X_1=1) = \frac{4}{30} = \frac{0.66}{5}$

$\left.\begin{array}{l} \sum_{x_2} p(x_1,x_2) = p(x_1) \\ \sum_{x_1} p(x_1,x_2) = p(x_2) \end{array}\right\}$ Marginalizing over

$$\int_{\mathbb{R}} f(x,y)\,dy = f(x)$$

$$= \sum_{x_1} x_1 \, p(x_1) + \sum_{x_2} x_2 \, p(x_2) = E(X_1) + E(X_2)$$

$$\Rightarrow E[X_1 + X_2] = E(X_1) + E(X_2) \text{ for any } X_1, X_2$$

$$\Rightarrow E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i]$$

$$E[\bar{X}] = \ldots\ldots = \mu$$

$$E[T] = \ldots\ldots = n\mu$$

Now... Recall $X_1, \ldots, X_r \overset{iid}{\sim} \text{Geom}(p)$

$$T = X_1 + \ldots + X_r \sim \text{NegBin}(r, p)$$

$$E[T] = \sum_{i=1}^{r} E[X_i] = \sum_{i=1}^{r} \frac{1}{p} = \boxed{\frac{r}{p}}$$

Recall $X \sim \text{Hyper}(n, K, N)$

$$X = X_1 + X_2 + \ldots + X_n \qquad E(X) = \sum_{i=1}^{n} E(X_i) = \boxed{n \frac{K}{N}}$$

where $X_1, X_2, \ldots X_n \sim \text{Bern}\left(\frac{K}{N}\right)$

but not independent

Back to Variance...

$\mathrm{Var}(X) ::$

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

$$\boxed{\sigma^2 = \mathrm{Var}(X) = E[X^2] - \mu^2} \quad \Rightarrow \boxed{E(X^2) = \sigma^2 + \mu^2}$$

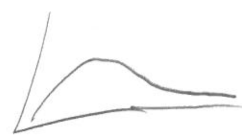
Note

$E[X]$ first moment — $E[X-\mu]$ first central moment

$E(X^2)$ second moment — $\sigma^2 = E(X-\mu)^2$ "

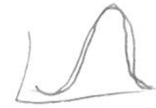
$E(X^3)$ third moment — $E[(X-\mu)^3]$ "

etc



Dimensionless

$\frac{E[X]}{\sigma}$ first standardized moment — $\frac{E(X-\mu)^3}{\sigma^3}$ skewness $+, -$



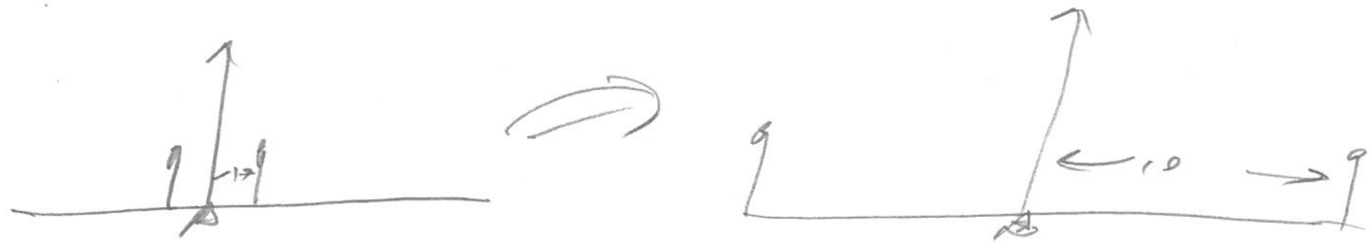
$\frac{E[(X-\mu)]}{\sigma^2}$ second — $\frac{E(X-\mu)^4}{\sigma^4}$ Kurtosis (tailedness)

$$\mathrm{Var}[aX] = E(aX - E(aX))^2 = \ldots \; a^2\,\mathrm{Var}(X)$$

$$SE(aX) = \sqrt{\mathrm{Var}(aX)} = \sqrt{a^2 \mathrm{Var}(X)} = |a|\sqrt{\mathrm{Var}(X)} = |a|\,SE(X) = |a|\sigma$$

$$\mathrm{Var}[X+c] = E[((X+c) - E[X+c])^2] = \ldots \; \mathrm{Var}(X)$$

$$\mathrm{Var}(aX+c) = a^2\,\mathrm{Var}(X) \qquad SE(aX+c) = |a|\sigma$$



$X \sim$ Rademacher

distances get multiplied by 10 and then squared hence $Var(9X) = 9^2 Var(X)$

Same distances!

$$Var[X_1 + X_2] = E\left[\left((X_1 + X_2) - \overbrace{E[X_1 + X_2]}^{\mu_1 + \mu_2}\right)^2\right]$$

$$= E\left[(X_1 + X_2 - \mu_1 - \mu_2)^2\right]$$

$$= E\left[X_1^2 + X_2^2 + \mu_1^2 + \mu_2^2 + 2X_1X_2 - 2X_1\mu_1 - 2X_2\mu_1 - 2X_1\mu_2 - 2X_2\mu_2 + 2\mu_1\mu_2\right]$$

$$= E[X_1^2] + E[X_2^2] + \mu_1^2 + \mu_2^2 + 2E[X_1X_2] - 2\mu_1 E[X_1] - 2\mu_1 E[X_2] - 2\mu_2 E[X_1] - 2\mu_2 E[X_2] + 2\mu_1\mu_2$$

$$= \sigma_1^2 + \cancel{\mu_1^2} + \sigma_2^2 + \cancel{\mu_2^2} + \cancel{\mu_1^2} + \cancel{\mu_2^2} + 2E[X_1X_2] - \cancel{2\mu_1^2} - 2\mu_1\mu_2 - 2\mu_2\mu_1 - \cancel{2\mu_2^2} + \cancel{2\mu_1\mu_2}$$

$$= Var[X_1] + Var[X_2] + 2\underbrace{\left(E[X_1X_2] - \mu_1\mu_2\right)}_{Cov[X_1, X_2] :=}$$

$\sum_{x_1}\sum_{x_2} x_1 x_2\, p(x_1, x_2)$ No dice!

bc of ind...

$\sum_{x_1}\sum_{x_2} x_1 x_2\, p(x_1) p(x_2)$

$\underbrace{\sum_{x_1} x_1 p(x_1)}_{\mu_1} \underbrace{\sum_{x_2} x_2 p(x_2)}_{\mu_2}$

$\Rightarrow Var[X_1 + X_2] = Var[X_1] + Var[X_2]$
if X_1, X_2 ind.

$Var[\sum X_i] = \sum Var(X_i)$ if $X_1, \dots, X_n$ ind.

$SE[X_1 + X_2] = \sqrt{\sigma_1^2 + \sigma_2^2}$ No real pattern...

$$Var(\bar{X}_n) = Var\left(\frac{\sum X_i}{n}\right) = \frac{1}{n^2}\sum Var(X_i) = \frac{\sigma^2 + \dots + \sigma^2}{n^2} = \frac{n\sigma^2}{n^2} = \boxed{\frac{\sigma^2}{n}}$$

$$\boxed{SE(\bar{X}) = \frac{\sigma}{\sqrt{n}}}$$

Key formulas!!

$X_1, \dots X_n \overset{iid}{\sim} Bern(p)$

$T = X_1 + \dots + X_n \sim Bin(n, p)$

$X \sim Hyper(n, K, N)$

$Var(X)$? Hard...

$$Var(T) = \sum Var(X_i) = \boxed{np(1-p)} \Rightarrow SE(T) = \sqrt{np(1-p)}$$

$$Var(T) = \overset{E(T-\mu)^2}{\sum_{x=0}^{n}} (x-np)^2 \binom{n}{x} p^x (1-p)^{n-x}$$

Hmm

$$= E(T^2) - \mu^2 = \sum_{x=0}^{n} x^2 \binom{n}{x} p^x (1-p)^{n-x} - n^2p^2$$

$$= \sum_{x=1}^{n} x^2 \binom{n}{x} p^x (1-p)^{n-x} = np \sum_{x=1}^{n} x \binom{n-1}{x-1} p^{x-1} (1-p)^{n-x}$$

$$np((n-1)p+1) - n^2p^2$$

$$np(np-p+1) - n^2p^2$$

$$\cancel{n^2p^2} - np^2 + np - \cancel{n^2p^2}$$

$$\Rightarrow n(p-p^2) = \boxed{np(1-p)}$$

let $y = x-1$

$$\sum_{y=0}^{n-1} (y+1)\binom{n-1}{y} p^y (1-p)^{n-y-1}$$

let $m = n-1$

$$\sum_{y=0}^{m} y \binom{m}{y} p^y (1-p)^{m-y} + \sum_{y=0}^{m} \binom{m}{y} p^y (1-p)^{m-y}$$

$Y \sim Bin(m, p) \longrightarrow$ mp $(n-1)p$

1

$X \sim \text{Geom}(p)$ $\mu = E(X) = \frac{1}{p}$

On HW...

$\text{Var}(X) = E[(X-\mu)^2] = \sum_{x=1}^{\infty} (x - \frac{1}{p})^2 (1-p)^{x-1} p$ ... algebra ... $= \boxed{\frac{1-p}{p^2}}$

conditional prob...

What is $P(X = 17 \mid X > 10)$

X>10

X<10

X=17

$= \frac{P(X=17 \;\&\; X>10)}{P(X>10)} = \frac{P(X=17)}{P(X>10)} = \frac{(1-p)^{16} p}{(1-p)^{10}} = (1-p)^6 p = P(X=7)$

$P(X = b + x \mid X > x) = \frac{P(X = b+x)}{P(X>x)} = \frac{(1-p)^{b+x-1}}{(1-p)^x} p = (1-p)^{x-1} p$

$= P(X = x)$

$b \in \mathbb{N} \leftarrow$ some position

"memorylessness" property

↑ midterm 2

↓ Final