Lec 7 MATH 340/640

Bag of apples and bananas. Draw n fruits one-at-a-time with replacement.

$p_1 := P(\text{drawing an apple})$ $\qquad X_1 :=$ # of apples drawn

$p_2 := P(\text{drawing a banana}) = 1 - p_1$ $\qquad X_2 :=$ # of bananas drawn

These p_1, p_2 do not change throughout the draws since we are replacing the fruit each time after the draw

$$X_1 \sim \text{Bin}(n, p_1)$$

$$X_2 \sim \text{Bin}(n, p_2) = \text{Bin}(n, 1-p_1)$$

not identically distr. Not independent! Note $X_1 + X_2 = n$

$$P(X_1 = x) \neq \underbrace{P(X_1 = x \mid X_2 = 1)}_{\text{this will be Deg}(n-1)\text{ so either } 0 \text{ or } 1}$$

$$\binom{n}{x} p_1^x (1-p_1)^{n-x} \in (0,1)$$

$$\vec{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim p_{X_1,X_2}(x_1,x_2) = \frac{n!}{x_1!\,x_2!} p_1^{x_1} p_2^{x_2} \mathbb{1}_{x_1+x_2=n} \mathbb{1}_{x_1 \in \{0,\dots,n\}} \underbrace{\mathbb{1}_{x_2 \in \{0,\dots,n\}} \mathbb{1}_{n \in \mathbb{N}_0}}$$

Define $\binom{n}{x_1, x_2} := \frac{n!}{x_1!\,x_2!}$

$\Rightarrow \vec{X} \sim \binom{n}{x_1,x_2} p_1^{x_1} p_2^{x_2} = \text{Multi}\left(n, \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}\right)$, the Multinomial distr.

Looks like Binomial, but 2-dimensional! $\Rightarrow$ not Binomial!!

Now add cantelopes to basket

$p_1 := P(\text{picking apple})$, $X_1 :=$ # apples drawn

$p_2 := P(\text{picking banana})$, $X_2 :=$ # bananas drawn

$p_3 := P(\text{picking cantelope})$, $X_3 :=$ # cantelopes drawn

$\vec{p} := \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}$

$$\vec{X} \sim \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \sim \text{Multi}(n, \vec{p}) = \binom{n}{x_1, x_2, x_3} p_1^{x_1} p_2^{x_2} p_3^{x_3}$$

let $n = 7$, $x_1 = 2$, $x_2 = 2$, $x_3 = 3$ $\quad$ A A B B C C C

How many ways? $\dfrac{7!}{2!\,2!\,3!}$

Generally, K types of objects, $p_1, \ldots, p_K$ are prob's of picking each object, $\vec{p} = \begin{bmatrix} p_1 \\ \vdots \\ p_K \end{bmatrix}$

$$\vec{X} \sim \text{Multi}(n, \vec{p}) := \binom{n}{x_1, \ldots, x_K} \prod_{j=1}^{K} p_j^{x_j}$$

$$\binom{n}{x_1 \ldots x_K} := \frac{n!}{\prod_{j=1}^{K} x_j!} \prod_{j=1}^{K} \mathbb{1}_{x_j \in \{0, \ldots, n\}} \mathbb{1}_{\sum x_j = n} \mathbb{1}_{n \in \mathbb{N}_0}$$

The multichoose notation "takes care of" the indicator function in the PMF.

$$S_{\vec{X}} = \{\vec{x} : \vec{x} \in \{0, 1, \ldots, n\}^K \text{ and } \vec{x} \cdot \vec{1}_K = n\}$$

Param Space: $n \in \mathbb{N}$ if $n = 0 \Rightarrow \vec{X} \sim \text{Deg}(\vec{0}_K)$

$$\vec{p} \in \{\vec{v} : \vec{v} \in (0,1)^K \text{ and } \vec{v} \cdot \vec{1}_K = 1\}$$

$$\sum_{\vec{x} \in \mathbb{R}^n} p_{\vec{X}}(\vec{x}) = 1$$

$$\Downarrow$$

$$\sum_{x_1 \in \mathbb{R}, x_2 \in \mathbb{R}, \ldots x_K \in \mathbb{R}} \binom{n}{x_1 \ldots x_K} p_1^{x_1} \ldots p_K^{x_K} = 1$$

multinomial thm (HW)

$$\vec{X} \sim \text{Multi}(n, \vec{p}) \sim \binom{n}{x_1 \ldots x_K} p_1^{x_1} p_2^{x_2} \cdots p_K^{x_K}$$

Intuitively, $X_i \sim p_{X_i}(x_i) = \text{Bin}(n, p_i)$. How to prove it?

$p_{X_i}(x)$ is a marginal distr. To obtain, we need to use Law of Total Prob to margin out all other dimensions. WLOG, let $i=1$. Thus,

$$p_{X_1}(x_1) = P(X_1 = x_1) = \sum_{\substack{x_2 \in \mathbb{R},\\ x_3 \in \mathbb{R},\\ \vdots \\ x_K \in \mathbb{R}}} p_{X_1,\ldots,X_K}(x_1,\ldots,x_K) = \sum_{\vec{x}_{-1} \in \mathbb{R}^{K-1}} \frac{n!}{x_1! \cdots x_K!} p_1^{x_1} \cdots p_K^{x_K} \mathbb{1}_{x_1+\ldots+x_K=n} \mathbb{1}_{n \in \mathbb{N}_0} \prod_{j=1}^{K} \mathbb{1}_{x_j \in \{0,\ldots,n\}}$$

$$= \frac{n!}{x_1!} p_1^{x_1} \mathbb{1}_{x_1 \in \{0,\ldots,n\}} \mathbb{1}_{n \in \mathbb{N}_0} \sum_{\vec{x}_{-1} \in \mathbb{R}^{K-1}} \frac{1}{x_2! \cdots x_K!} p_2^{x_2} \cdots p_K^{x_K} \prod_{j=2}^{K} \mathbb{1}_{x_j \in \{0,\ldots,n\}} \cdot \frac{(n-x_1)!}{(n-x_1)!}$$

Note: $\mathbb{1}_{x_1 \in \{0,\ldots,n\}} = \mathbb{1}_{x_1 \in \{0,\ldots,n\}} \mathbb{1}_{n-x_1 \in \mathbb{N}_0}$

$$= \frac{n!}{x_1!(n-x_1)!} p_1^{x_1} \mathbb{1}_{x_1 \in \{0,\ldots,n\}} \mathbb{1}_{n \in \mathbb{N}_0} \sum_{\vec{x}_{-1} \in \mathbb{R}^{K-1}} \frac{(n-x_1)!}{x_2! \cdots x_K!} p_2^{x_2} \cdots p_K^{x_K} \mathbb{1}_{n-x_1 \in \mathbb{N}_0} \prod_{j=2}^{K} \mathbb{1}_{x_j \in \{0,\ldots,n\}}$$

Note: $p_2 + \ldots + p_K = 1 - p_1 \neq 1 \Rightarrow$ mult top and bottom by $\frac{(1-p_1)^{x_2}}{(1-p_1)^{x_2}} \cdot \ldots \cdot \frac{(1-p_1)^{x_K}}{(1-p_1)^{x_K}} = \frac{(1-p_1)^{n-x_1}}{(1-p_1)^{x_2} \cdots (1-p_1)^{x_K}}$

$$= \binom{n}{x_1} p_1^{x_1} (1-p_1)^{n-x_1} \underbrace{\sum_{\vec{x}_{-1} \in \mathbb{R}^{K-1}} \binom{n-x_1}{x_2 \ldots x_K} \left(\frac{p_2}{1-p_1}\right)^{x_2} \cdots \left(\frac{p_K}{1-p_1}\right)^{x_K}}_{\text{PMF of Multi}\left(n-x_1, \begin{bmatrix} \frac{p_2}{1-p_1} \\ \vdots \\ \frac{p_K}{1-p_1} \end{bmatrix}\right)} = \text{Bin}(n, p_1) \checkmark$$

Happy Dumpty $= 1$

We will see after midterm that using ch.f. allows this to be done in one line!!!

Is X_j, X_k independent? No if you know X_k is large $\Rightarrow X_j$ must be smaller since many samples were accounted for. $X_k \uparrow \Leftrightarrow X_j \downarrow$
They "inversely covary".
How do we measure dependence? One metric is "covariance".

Let X_1, X_2 be two random variables with means $\mu_1, \mu_2 < \infty$ and variances $\sigma_1^2, \sigma_2^2 < \infty$.

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

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

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

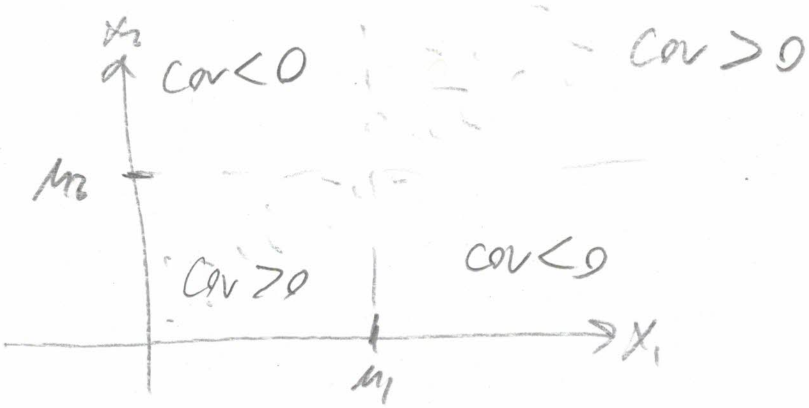
$$= \underbrace{E[X_1^2] - \mu_1^2}_{\sigma_1^2} + \underbrace{E[X_2^2] - \mu_2^2}_{\sigma_2^2} + 2(\underbrace{E[X_1X_2] - \mu_1\mu_2}_{\sigma_{12} = \sigma_{21} := Cov[X_1, X_2]})$$

Covariance

usually...

$$\sigma_{12} := Cov[X_1, X_2] := E[(X_1 - \mu_1)(X_2 - \mu_2)] = E[X_1X_2 - \mu_1 X_2 - \mu_2 X_1 + \mu_1\mu_2]$$

$$= E[X_1X_2] - \mu_1\mu_2 - \cancel{\mu_1\mu_2 + \mu_1\mu_2}$$



Average of all prob-weighted mean-centered support values is the covariance.

$\vec{X} \sim Multi(n, \vec{p})$, $Cov[X_i, X_j] < 0$? Yes as $X_i \uparrow \Leftrightarrow X_j \downarrow$

Keep this in mind...

Rules for Covariance

(1) If X_i, X_j independent

$$E[X_i X_j] = \sum_{x_i \in \mathbb{R}} \sum_{x_j \in \mathbb{R}} x_i x_j \, p_{X_i X_j}(x_i, x_j) = \sum_{x_i \in \mathbb{R}} \sum_{x_j \in \mathbb{R}} x_i x_j \, p_{X_i}(x_i) \, p_{X_j}(x_j)$$

$$= \sum_{x_i \in \mathbb{R}} x_i \, p_{X_i}(x_i) \sum_{x_j \in \mathbb{R}} x_j \, p_{X_j}(x_j) = E(X_i) E(X_j) = \mu_i \mu_j$$

if X_i, X_j ind.

$$\Rightarrow Cov[X_i X_j] = E[X_i X_j] - \mu_i \mu_j = 0 \checkmark \Rightarrow Var(X_i + X_j) = Var(X_i) + Var(X_j)$$

~~Is the covariance $Cov[X_i, X_j]$ related to the variances $Var(X_i) Var(X_j)$?~~

~~Yes... through a very famous inequality.~~

~~let $W =$~~

(2) $Cov[X_i, X_i] = E(X^2) - \mu^2 = Var(X) = \sigma^2$

(3) $Cov[X_i, X_j] = Cov[X_j, X_i]$ (Her) $\quad \sigma_{ij} = \sigma_{ji}$

(4) $Cov[X_i + X_j, X_k] = Cov[X_i, X_k] + Cov[X_j, X_k]$

(5) $Cov[aX_i, bX_j] = ab\,\sigma_{ij}$

(6) $Var[X_1 + \dots + X_n] = \sum_{i=1}^{n} \sum_{j=1}^{n} Cov[X_i, X_j]$

Is the $Cov[X_i, X_j]$ related to $Var(X_i)$ and $Var(X_j)$?

Yes, through a very famous inequality:

Consider rv's X, Y with $\mu_X < \infty$, $\mu_Y < \infty$, $\sigma_X^2 < \infty$, $\sigma_Y^2 < \infty$.

Let $W = (X - cY)^2 \quad \forall c \in \mathbb{R}$

Define: non-neg rv means:

$S_W \geq 0 \Rightarrow E[W] \geq 0$ why? $\int_{w \in S_W} \underbrace{w}_{\geq 0} \underbrace{f_W(w)}_{\geq 0} \, dw \geq 0$ Same for discrete

$\Rightarrow E[(X - cY)^2] \geq 0$

$\Rightarrow E[X^2 - 2cXY + c^2Y^2] \geq 0$

$\Rightarrow E[X^2] - 2c\,E(XY) + c^2 E(Y^2) \geq 0$

let $c = \dfrac{E(XY)}{E(Y^2)} \in \mathbb{R}$

$$\Rightarrow E(X^2) - 2\frac{E(XY)}{E(Y^2)}E(XY) + \frac{E(XY)^2}{E(Y^2)^{\cancel{2}}}\cancel{E(Y^2)} \geq 0$$

mult by $E(Y^2)$

$\Rightarrow E(X^2)E(Y^2) - 2E(XY)^2 + E(XY)^2 \geq 0$

$\Rightarrow E(X^2)E(Y^2) - E(XY)^2 \geq 0$

$\Rightarrow E(XY)^2 \leq E(X^2)E(Y^2)$

$\Rightarrow |E(XY)| \leq \sqrt{E(X^2)E(Y^2)} \Rightarrow E(XY) \leq \sqrt{E(X^2)E(Y^2)}$

Cauchy Schwarz Inequality   if X, Y non-neg

We use this to prove the cov. ineq.:

$$Cov(X,Y)^2 = E[(X-\mu_X)(Y-\mu_Y)]^2 \leq E[(X-\mu_X)^2]\,E[(Y-\mu_Y)^2] = Var(X)\,Var(Y)$$

$$\Rightarrow Var(X) \geq \frac{Cov(X,Y)^2}{Var(Y)}$$