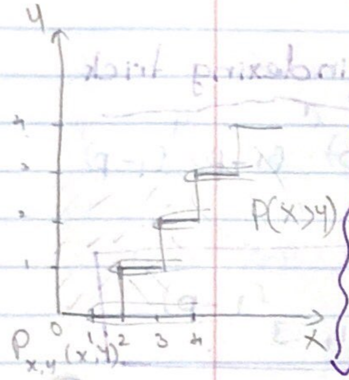
Lecture - 04 09/09/2020

Let $x, y \overset{iid}{\sim} \text{Geom}(p)$

$P(x > y) = ?$



$$P(x>y) = P(y>x)$$

$$P(x>y) + P(y>x) + P(x=y) = 1$$

$$\Rightarrow 2P(x>y) + P(x=y) = 1$$

$$\Rightarrow P(x>y) = \frac{1 - P(x=y)}{2} < 1/2.$$

since $P(X=Y) > 0$

$$P(x>y) = \sum_{x \in \mathbb{R}} \sum_{y \in \mathbb{R}} \underbrace{P_{x,y}(x,y)}_{P_x(x)\,P_y(y)} \mathbb{1}_{x>y}$$

$$= \sum_{x \in \mathbb{R}} \sum_{y \in \mathbb{R}} P_x(x)\, P_y(y)\, \mathbb{1}_{x>y}$$

$$= \sum_{y \in \mathbb{R}} \sum_{x \in \mathbb{R}} P_x(x)\, P_y(y)\, \mathbb{1}_{x>y}$$

$$= \sum_{y \in \mathbb{R}} P_y(y) \sum_{x \in \mathbb{R}} P_x(x)\, \mathbb{1}_{x>y}$$

$$= \sum_{y \in \{0,1,\dots\}} \boxed{p(1-p)^y} \sum_{x \in \{0,1,\dots\}} p(1-p)^x\, \mathbb{1}_{\underbrace{x>y}_{x \geq y+1}}$$

$$= p^2 \sum_{y \in \{0,1,\dots\}} (1-p)^y \sum_{x \in \{y+1, y+2, \dots\}} (1-p)^x$$

let $x' = x - (y+1) \Rightarrow x' \in \{0,1,\dots\}$

$\Rightarrow x = x' + y + 1$

reindexing trick

$$= p^2 \sum_{y \in \{0,1,\dots\}} (1-p)^y \sum_{x' \in \{0,1,\dots\}} (1-p)^{x'} (1-p)^y (1-p)^1$$

$$= p^2(1-p) \sum_{y \in \{0,1,\dots\}} (1-p)^{2y} \boxed{\sum_{x' \in \{0,1,\dots\}} (1-p)^{x'}}$$

$$= \frac{1}{1-(1-p)} = 1/p$$

Geometric series, for $a \in (-1,1) \setminus \{0\}$

$$\sum_{i=0}^{\infty} a^i = 1/_{1-a}$$

$$= p^2(1-p) \sum_{y \in \{0,1,\dots\}} (1-p)^{2y} \cdot (1/p)$$

$$= p(1-p) \boxed{\sum_{y \in \{0,1,\dots\}} \left((1-p)^2\right)^y} = \frac{1}{1-(1-p)^2} = \frac{1}{1-(1-2p+p^2)}$$

$$= \frac{1}{2p-p^2} = \frac{1}{p(2-p)}$$

$$= \frac{p(1-p)}{p(2-p)} = \boxed{\frac{1-p}{2-p}} < \frac{1}{2}$$

Bag of fruits; apples and bananas.

Draw with replacement n times
Let X_1 = # apples, $P_1 = P(\text{apple})$

$\Rightarrow X_1 \sim Bin(n, p_1)$; Draw n with replacement.

X_1 = # apples, X_2 = # bananas.

$X_1 \sim Bin(n, p_1)$ $\quad X_2 \sim Bin(n, p_2)$

Are X_1 and X_2 independent?

Since $X_1 + X_2 = n \Rightarrow X_1, X_2$ dependent.

$$\vec{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$

$$\vec{X} \sim P_{\vec{X}}(\vec{x}) = P_{\vec{X}}(x_1, x_2) =$$

$$= \underbrace{\frac{n!}{x_1! \, x_2!}}\; p_1^{x_1} p_2^{x_2} \, \mathbb{1}_{\underbrace{x_1 + x_2 = n}} \, \mathbb{1}_{\underbrace{x_1 \in \{0,1,\dots,n\}}} \, \mathbb{1}_{\underbrace{x_2 \in \{0,1,\dots,n\}}}$$

$\binom{n}{x_1, x_2}$ multichoose notation

$$\Rightarrow \vec{X} \sim Multi(n, \vec{p}) = \binom{n}{x_1, x_2} p_1^{x_1} p_2^{x_2}, \quad \vec{p} = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}$$

multinomial rv of dim $= 2$

Since X_1, X_2 are dependent, we cannot factor this JMF.

Bag of fruit now has cantaloupes. You draw cantaloupes with probability p_3 and X_3 is the count of cantaloupes.

$$\vec{X} \sim \text{Multi}(n, \vec{p}) = \binom{n}{x_1, x_2, x_3} p_1^{x_1} p_2^{x_2} p_3^{x_3} \; \mathbb{1}_{x_1 \in \{0,1,\dots,n\}} \, \mathbb{1}_{x_2 \in \{0,1,\dots,n\}} \, \mathbb{1}_{x_3 \in \{0,1,\dots,n\}}$$

$$= \frac{n!}{x_1!\, x_2!\, x_3!} p_1^{x_1} p_2^{x_2} p_3^{x_3} \; \mathbb{1}_{x_1 + x_2 + x_3 = n}$$

In general, If there are K types of fruit (# categories) then the general multinomial rv of dim K is:

$$\vec{X} \sim \text{multi}(n, \underset{\begin{bmatrix} p_1 \\ \vdots \\ p_k \end{bmatrix}}{\vec{p}}) = \binom{n}{x_1, x_2, \dots, x_k} \prod_{k=1}^{K} p_k^{x_k}$$

Parameter Space: $n \in \mathbb{N}$, $\vec{p} \in \{ \vec{v} : \vec{v} \cdot \vec{1} = 1, v_1 \in (0,1), \dots, v_k \in (0,1) \}$

Support: $\text{Supp}[\vec{X}] = \{ \vec{x} : \vec{x} \cdot \vec{1} = n, x_1 \in \{0,1,\dots,n\}, \dots, x_k \in \{0,1,\dots,n\} \}$

$k = 2$

$\vec{X} \sim \text{Multi}\left(n, \begin{bmatrix} p \\ 1-p \end{bmatrix}\right) = \binom{n}{x_1, x_2} p^{x_1} (1-p)^{x_2}$. Dependent?

$P(X_1 = x_1 \mid X_2 = x_2) \overset{?}{=} P(X_1 = x_1) = \text{Bin}(n, p_1)$

$\quad \parallel$

$\text{Deg}(n - x_2) \implies$ Dependent!

Conditional PMF

$$P_{X_1|X_2}(x_1, x_2) := \frac{P_{X_1,X_2}(x_1, x_2)}{P_{X_2}(x_2)}$$

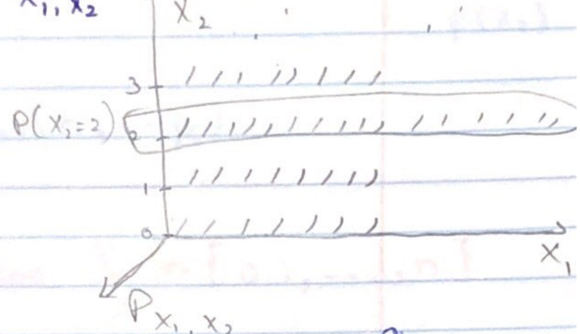
JMF

Marginal PMF of X_2

want to show: $X_2 \sim \text{Bin}(n, p_2)$

$$P_{X_2}(x_2) = P(X_2 = x_2) = \sum_{x_1 \in \mathbb{R}} P_{X_1,X_2}(x_1, x_2)$$

"Margining out x_1"



$$= \sum_{x_1 \in \mathbb{R}} \binom{n}{x_1, x_2} p^{x_1} (1-p)^{x_2}$$

$$= \sum_{x_1 \in \mathbb{R}} \frac{n!}{x_1!\, x_2!} p^{x_1} (1-p)^{x_2} \mathbb{1}_{x_1 + x_2 = n} \mathbb{1}_{x_1 \in \{0,\ldots,n\}} \mathbb{1}_{x_2 \in \{0,1,\ldots,n\}}$$

$$= \frac{n!}{x_2!} (1-p)^{x_2} \mathbb{1}_{x_2 \in \{0,1,\ldots,n\}} \sum_{x_1 \in \{0,1,\ldots,n\}} \frac{p^{x_1}}{x_1!} \mathbb{1}_{x_1 = n - x_2}$$

$$= \frac{n!}{x_2!} (1-p)^{x_2} \mathbb{1}_{x_2 \in \{0,1,\ldots,n\}} \frac{p^{n-x_2}}{(n-x_2)!}$$

$$= \binom{n}{x_2} p^{n-x_2} (1-p)^{x_2} = \text{Bin}(n, 1-p)$$

Margining a multinomial to yield one dimension is binomial