$X, Y \overset{iid}{\sim} Geom(p)$

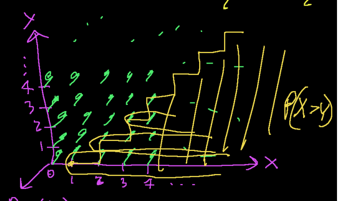
$P(X>Y) \overset{?}{=} \frac{1}{2}$ good guess

$P(X>Y) = P(Y>X)$ > 0

$\underbrace{P(X>Y) + P(Y>X)} + P(X=Y) = 1$

$2P(X>Y) = 1 - P(X=Y)$

$\Rightarrow P(X>Y) = \frac{1-P(X=Y)}{2} < \frac{1}{2}$



$$P(X>Y) = \sum_{x\in\mathbb{R}} \sum_{y\in\mathbb{R}} p_{X,Y}(x,y) \mathbb{1}_{x>y}$$

X,Y indep

$$= \sum_{x\in\mathbb{R}} \sum_{y\in\mathbb{R}} p_X(x) p_Y(y) \mathbb{1}_{x>y}$$

$$= \sum_{x\in\{0,1,\ldots\}} \sum_{y\in\{0,1,\ldots\}} (1-p)^x p (1-p)^y p \, \mathbb{1}_{x>y}$$

$x \geq y+1 \Rightarrow x \in \{y+1, y+2, \ldots\}$

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

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

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

"Reindexing Trick" kind of like a u subst.

$x' = x-(y+1) \in \{0,1,2,\ldots\}$

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

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

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

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

Geometric Series Formula

$\sum_{i=0}^{\infty} a^i = \frac{1}{1-a}$ for $a\in(-1,1)\setminus\{0\}$

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

precalc: $(1-p)^{2y} = ((1-p)^2)^y$

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

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

Consider a bag of fruit that has apples and bananas. You now draw with replacement n samples from this bag and you count how many are apples and how many are bananas. Let X_1 be the rv that counts the number of apples and let X_2 be the rv that counts the number of bananas. Let p_1 be the probability of picking an apple and p_2 be the probability of picking a banana.

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

Are X_1 and X_2 independent? NO $\Rightarrow$ dependent

$P(X_1 = x_1 \mid X_2 = x_2) \overset{?}{=} P(X_1 = x_1)$

$n = 10$ $P(X_1 = 6 \mid X_2 = 4) = 1 \neq \binom{10}{6} p_1^6 (1-p_1)^4$

$P(X_1 = 3 \mid X_2 = 4) = 0$

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

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

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

Multinomial rv of dim = 2

Let's add cantaloupes to the bag. Let X_3 count the number of cantaloupes and p_3 be the probability of drawing a cantaloupe.

$$\vec{X} \sim Multin(n, \vec{p}) := \binom{n}{x_1,x_2,x_3} p_1^{x_1} p_2^{x_2} p_3^{x_3} = \frac{n!}{x_1!\,x_2!\,x_3!} p_1^{x_1} p_2^{x_2} p_3^{x_3} \mathbb{1}_{x_1+x_2+x_3=n} \mathbb{1}\mathbb{1}\mathbb{1}$$

$Multin(n,\vec{p}) = p_{\vec{X}}(x_1,x_2,x_3)$, $\vec{p} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}$

The general multinomial rv of dim = K has PMF:

$$\vec{X} \sim Multin(n, \vec{p}) := \binom{n}{x_1,x_2,\ldots,x_K} \prod_{k=1}^{K} p_k^{x_k}$$

$\vec{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_K \end{bmatrix}$, $\vec{p} = \begin{bmatrix} p_1 \\ \vdots \\ p_K \end{bmatrix}$

Param Space

$n \in \mathbb{N}$

$\vec{p} \in \{\vec{v} : \vec{v}\cdot\vec{1} = 1, v_1 \in (0,1), \ldots, v_K \in (0,1)\}$

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

I want to derive the condition PMF and the marginal PMF's in the case of K = 2 (apples and bananas).

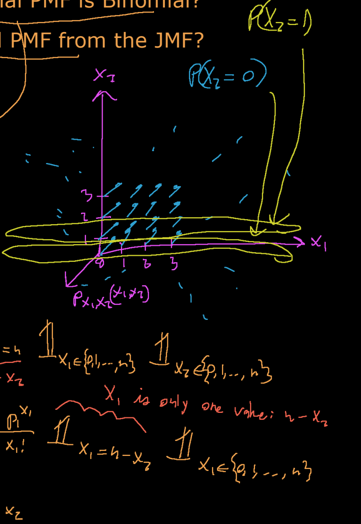
$$Deg(n-x_2) = p_{X_1|X_2}(x_1,x_2) = \frac{p_{X_1,X_2}(x_1,x_2)}{p_{X_2}(x_2)}$$

JMF

marginal PMF

How do we prove that the marginal PMF is Binomial?

How do we compute the marginal PMF from the JMF?



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

$$= \sum_{x_1\in\mathbb{R}} \binom{n}{x_1,x_2} p_1^{x_1} p_2^{x_2}$$

$$= p_2^{x_2} \sum_{x_1\in\mathbb{R}} \frac{n!}{x_1!\,x_2!} p_1^{x_1} \mathbb{1}_{x_1+x_2=n} \mathbb{1}_{x_1\in\{0,1,\ldots,n\}} \mathbb{1}_{x_2\in\{0,1,\ldots,n\}}$$

$x_1 = n - x_2$

$$= p_2^{x_2} \frac{n!}{x_2!} \mathbb{1}_{x_2\in\{0,1,\ldots,n\}} \sum_{x_1\in\mathbb{R}} \frac{p_1^{x_1}}{x_1!} \mathbb{1}_{x_1=n-x_2} \mathbb{1}_{x_1\in\{0,1,\ldots,n\}}$$

x_1 is only one value: $n - x_2$

$$= p_2^{x_2} \frac{n!}{x_2!} \mathbb{1}_{x_2\in\{0,1,\ldots,n\}} \frac{p_1^{n-x_2}}{(n-x_2)!} = \binom{n}{x_2} p_2^{x_2} p_1^{n-x_2} = Bin(n, p_2)$$

$\binom{n}{x_2}$

Since $p_1 + p_2 = 1 \Rightarrow p_1 = 1 - p_2$