$X, Y \overset{iid}{\sim} Geom(p)$ $\quad P(X > Y) \overset{?}{=} \frac{1}{2}$

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

$$P(X > Y) + P(Y > X) + P(Y = X) = 1$$

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

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

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

$$\overset{X,Y \text{ indep}}{=} \sum_{x \in \mathbb{R}} \sum_{y \in \mathbb{R}} P_X(x) P_Y(y)\, \mathbb{1}_{x > y} = \sum_{x \in \{0,1,..\}} \sum_{y \in \{0,1,..\}} (1-p)^x p (1-p)^y p\, \mathbb{1}_{x > y}$$

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

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

"Reindexing trick":
$x' = x - (y+1) \in \{0,1,2,...\}$
$\Rightarrow x = x' + y + 1$

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

$$= p^2 (1-p) \sum_{y \in \{0,1,..\}} (1-p)^{2y} \left[ \sum_{x' \in \{0,1,..\}} (1-p)^{x'} \right]$$

Geometric formula
$$\sum_{i=0} a^i = \frac{1}{1-a}$$

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

Now consider a bag of fruit that has apples and bananas. You draw w/ replacement $n$ samples and count how many are apples and how many are bananas.

Let $X_1$ be an r.v that counts num of apples

Let $X_2$ be an r.v that counts num of bananas

Let $p_1$ be proba for picking apple

Let $p_2$ be proba for picking bananas

$X_1 \sim Bin(n, p)$ and $X_2 \sim Bin(n, p_2)$

Are $X_1$ and $X_2$ indep? NO!

Let $\vec{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$, $\vec{X} \sim P_{\vec{X}}(x_1, x_2) = p_1^{x_1} p_2^{x_2} \underbrace{\frac{n!}{x_1!\, x_2!} \mathbb{1}_{x_1+x_2=n} \mathbb{1}_{x_1 \in \mathbb{N}_0} \mathbb{1}_{x_2 \in \mathbb{N}_0}}_{\binom{n}{x_1, x_2} \text{ multichoose notation}}$

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

multinomial r.v of dim=2

Add cantaloupes to the bag. Let $X_3$ count the num of cantaloupes and $p_3$ the proba.

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

The general multinomial r.v of dim $k$ has PMF:

$\vec{X} \sim Multi(n, \vec{p}) = \binom{n}{x_1, \ldots, x_k} \prod_{k=1}^{k} p_k^{x_k}$

Param space: $n \in \mathbb{N}$ $\vec{p} \in \{\vec{v} : \vec{v} \cdot \vec{1}, v_1 \in (0,1), \ldots, v_k \in (0,1)\}$

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

I want to derive the condition PMF and the marginal PMF's

ie the marginal PMF in the $k=2$ (apples and bananas)

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

How do prove that the marginal PMF is Binomial?

How do we compute marginal PMF from the JMF?

$$P_{X_2}(x_2) = \sum_{x_1 \in \mathbb{R}} P_{X_1 X_2}(x_1, x_2) = \sum_{x_1 \in \mathbb{R}} \binom{n}{x_1, x_2} p_1^{x_1} p_2^{x_2}$$

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

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

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

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

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