Let $X, Y \overset{iid}{\sim} \text{Geom}(p) = (1-p)^x p \, \mathbb{1}_{x \in \mathbb{N}_0}$

$P(X > Y)$ it is not equal to 1/2 since X = Y sometimes but it should be slightly less than 1/2.

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

$X, Y \overset{iid}{\sim}$

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

$$= \sum_{x \in \mathbb{R}} \sum_{y \in \mathbb{R}} (1-p)^x p \mathbb{1}_{x \in \mathbb{N}_0} (1-p)^y p \mathbb{1}_{y \in \mathbb{N}_0} \mathbb{1}_{x \geq y+1}$$

$$= p^2 \sum_{y \in \mathbb{N}_0} \sum_{x \in \mathbb{N}_0} (1-p)^x (1-p)^y \mathbb{1}_{x \geq y+1}$$

$$= p^2 \sum_{y \in \mathbb{N}_0} (1-p)^y \sum_{x \in \mathbb{N}_0} (1-p)^x \mathbb{1}_{x \in \{y+1, y}$$

$$= p^2 \sum_{y \in \mathbb{N}_0} (1-p)^y \sum_{x \in \mathbb{N}_0} (1-p)^x \mathbb{1}_{x \in \{y+1, y+2, \dots\}}$$

$$= p^2 \sum_{y \in \mathbb{N}_0} (1-p)^y \sum_{x \in \{y+1, \dots\}} (1-p)^x$$

let $x' = x - (y+1) \Rightarrow x' \in \mathbb{N}_0$

$\Downarrow$

$x = x' + (y+1)$

$$= p^2 \sum_{y \in \mathbb{N}_0} (1-p)^y \sum_{x' \in \mathbb{N}_0} (1-p)^{x' + (y+1)}$$

Geometric Series

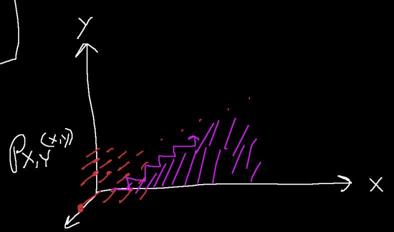
$$= \cancel{p^2} \sum_{y \in \mathbb{N}_0} (1-p)^y (1-p)^y (1-p) \boxed{\sum_{x' \in \mathbb{N}_0} (1-p)^{x'}} = \frac{1}{1-(1-p)} = \frac{1}{\cancel{p}}$$

$$= p(1-p) \sum_{y \in \mathbb{N}_0} (1-p)^{2y} = p(1-p) \boxed{\sum_{y \in \mathbb{N}_0} \left((1-p)^2\right)^y} = \frac{1}{1-(1-p)^2} = \frac{1}{1-(1-2p+p^2)}$$

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

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

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



$P(\sin(X) > 2^Y)$

Bag of fruits: apples, bananas and cantaloupes
p_1: probability of drawing an apple
p_2: probability of drawing a banana
p_3: probability of drawing a cantaloupe
Note: $p_1 + p_2 + p_3 = 1$

Draw n fruits with replacement. Let X_1 be the rv that counts the number of apples, X_2 be the rv that counts the number of bananas and X_3 be the rv that counts the number of cantaloupes.
Note: $X_1 + X_2 + X_3 = n$

$X_1 \sim \text{Bin}(n, p_1), \quad X_2 \sim \text{Bin}(n, p_2), \quad X_3 \sim \text{Bin}(n, p_3)$

What is the probability of drawing 2 apples, 2 bananas and 3 cantaloupes on a draw of n=7?

$$p_{X_1,X_2,X_3}(2,2,3) = P(X_1 = 2, X_2 = 2, X_3 = 3) = \frac{7!}{2!\,2!\,3!} p_1^2 p_2^2 p_3^3$$

$$\parallel$$

$$p_{\vec{X}}(\underbrace{[2\ 2\ 3]^T}_{\vec{x}})$$

A A B B C C C

$$p_{\vec{X}}(\vec{x}) = \frac{n!}{x_1!\, x_2!\, x_3!} p_1^{x_1} p_2^{x_2} p_3^{x_3} \mathbb{1}_{x_1 + x_2 + x_3 = n} \mathbb{1}_{x_1 \in \{0, \dots, n\}} \mathbb{1}_{x_2 \in \{0, \dots, n\}} \mathbb{1}_{x_3 \in \{0, \dots, n\}}$$

$[x_1\ x_2\ x_3]^T$

$$= \binom{n}{x_1\ x_2\ x_3} p_1^{x_1} p_2^{x_2} p_3^{x_3}$$

This is called the multinomial rv. In general consider K categories of items and probabilities of sampling each category $p_1, p_2, \dots, p_K$ where $p_1 + p_2 + \dots + p_K = 1$. The number of samples is n.

$$\vec{X} \sim \text{Multinom}(n, \vec{p}) = \binom{n}{x_1, x_2, \dots, x_K} p_1^{x_1} p_2^{x_2} \cdot \dots \cdot p_K^{x_K}$$

$$\text{Supp}[\vec{X}] = \{\vec{x} : \vec{x} \in \{0, \dots, n\}^K \ \&\ \vec{x} \cdot \vec{1} = n\}$$

Parameter space: $n \in \mathbb{N}, \ \vec{p} \in \{\vec{v} : (0,1)^K \ \&\ \vec{v} \cdot \vec{1} = 1\}$

Let $K = 2$

$$\vec{X} \sim \text{Multi}\left(n, \begin{bmatrix} p \\ 1-p \end{bmatrix}\right) = \underbrace{\binom{n}{x_1, x_2}}_{\frac{n!}{x_1!(n-x_1)!}} p^{x_1} (1-p)^{\overset{n - x_1}{\overset{\parallel}{x_2}}} = \text{Bin}? \text{ No...}$$

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim p_{X_1, X_2}(x_1, x_2) \neq p_{X_1}(x_1)$$

Are X_1 and X_2 independent here? No!

$$P(X_1 = x_1 \mid X_2 = x_2) = P(X_1 = x_1) \quad \forall x_1 \in \text{Supp}[X_1], x_2 \in \text{Supp}[X_2]$$

$$0 = P(X_1 = 1 \mid X_2 = n) \neq P(X_1 = 1) = np(1-p)^{n-1}$$

Some definitions for vector rvs

if $\vec{X} \sim \text{Multin}(n, \vec{p})$

$$\vec{\mu} := E[\vec{X}] := \begin{bmatrix} E[X_1] \\ E[X_2] \\ \vdots \\ E[X_K] \end{bmatrix} = \begin{bmatrix} np_1 \\ np_2 \\ \vdots \\ np_K \end{bmatrix} = n\vec{p}$$

A matrix rv is a matrix whose entries are rv's

$$E[X] = \begin{bmatrix} E[X_{11}] & E[X_{12}] & \dots & E[X_{1n}] \\ E[X_{21}] & E[X_{22}] & & \\ \vdots & & \ddots & \\ E[X_{m1}] & & & E[X_{mn}] \end{bmatrix}$$

What is the "variance" of a vector rv? How do we define $\text{Var}[\vec{X}]$?

Recall the definition for a scalar rv. Var[X] := $E[(X-\mu)^2]$.

What is the "covariance" between two rv's X_1 and X_2?

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

Rules of covariances:

$$\text{Var}[X_1 + X_2] \overset{HW}{=} \sigma_1^2 + \sigma_2^2 + 2\sigma_{12}$$

① $\text{Cov}[X, X] = \text{Var}[X] \geq 0$

② $\text{Cov}[X_1, X_2] = \text{Cov}[X_2, X_1]$

③ $\text{Cov}[X_1 + X_2, X_3] = \text{Cov}[X_1, X_3] + \text{Cov}[X_2, X_3]$

④ $\text{Cov}[a_1 X_1, a_2 X_2] = a_1 a_2 \sigma_{12}$

⑤ $\text{Var}[X_1 + X_2 + \dots + X_n] = \sum_{i=1}^{n} \sum_{j=1}^{n} \text{Cov}[X_i, X_j]$

$n = 2$

$$= \sum_{i=1}^{2} \sum_{j=1}^{2} \text{Cov}[X_i, X_j] = \text{Cov}[X_1, X_1] + \text{Cov}[X_1, X_2] + \text{Cov}[X_2, X_1] + \text{Cov}[X_2, X_2]$$

$$= \sigma_1^2 + \sigma_2^2 + 2\sigma_{12}$$