1

Lec 3 Math 621 9/5/17

$X_1, X_2 \overset{iid}{\sim} Bern(p)$, $T = X_1 + X_2$

$$p(t) = p_{X_1}(x) * p_{X_2}(x) = \sum_{x \in Supp(X_1)} p_{X_1}(x)\, p_{X_2}(t-x) \overset{?}{=} 2p^t (1-p)^{2-t}$$

$$p(2) \overset{?}{=} p_{X_1}(0)\, p_{X_2}(2-0) + p_{X_1}(1)\, p_{X_2}(2-1)$$

$$= p_{X_1}(0) \underbrace{p_{X_2}(2)}_{=0} + p_{X_1}(1) \cdot p_{X_2}(1)$$

$$= p^0 (1-p)^2 p^2 (1-p)^0 + p^1 (1-p)^1 p^1 (1-p)^1$$

$$= 2p^2 (1-p)^2$$

Identity

$$2^n = \sum_{i=0}^{n} \binom{n}{i}$$

let $A = \{w_1, w_2, \ldots, w_n\}$ $|A| = n$

let $2^A = \{B : B \subseteq A\}$

$$= \{B : B \subseteq A \ \& \ |A| = 0\} \cup$$
$$\{B : B \subseteq A \ \& \ |A| = 1\} \cup$$
$$\{B : B \subseteq A \ \& \ |A| = 2\} \cup$$
$$\vdots$$
$$\cup \{B : B \subseteq A \ \& \ |A| = n\}$$

$$\Rightarrow 2^A = \bigcup_{i=1}^{n} \{B : B \subseteq A \ \& \ |A| = i\}$$

Since each set disjoint

$$2^n = |2^A| = \sum_{i=1}^{n} \Big| \underbrace{\{B : B \subseteq A \ \& \ |A| = i\}}_{\binom{n}{i}} \Big|$$

Known relationship

Recall $E(X) = \sum_{x \in Supp(X)} x \, p(x)$ for discrete r.v.'s

Consider a function of a r.v. g.

$$E[g(X)] = \sum_{x \in Supp(X)} g(x)\, p(x)$$ Sometimes expectations don't exist e.g.

Prove $X \sim Poisson(\lambda)$ $E(\ln(X))$ d.n.e.

$$= \sum_{x \in \{0,1,\dots\}} \ln(x) \frac{\lambda^x e^{-\lambda}}{x!} = \underbrace{\ln(0)}_{d.n.e.} \frac{\lambda^0 e^{-\lambda}}{0!} + \sum_{x=1}^{\infty} \ln(x) \frac{\lambda^x e^{-\lambda}}{x!} \Rightarrow d.n.e.$$

let $Z = \mathbb{1}_A$ What is $E[Z]$? Note $Z \sim Bern(P(A)) \Rightarrow E[Z] = P(A)$ Cool Fact!

if $Z = g(X,Y)$ i.e. a function of two r.v.'s

$$E[Z] = E[g(X,Y)] = \sum_{x \in Supp(X)} \sum_{y \in Supp(Y)} g(x,y)\, \overbrace{p_{X,Y}(x,y)}^{J.M.F.}$$

Let $X, Y \overset{iid}{\sim} Geom(p) = (1-p)^x (p)$

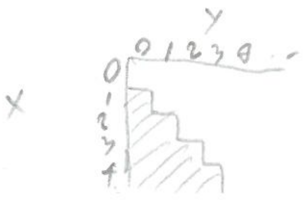
$$F(x) = P(X \le x) = 1 - P(X > x) = 1 - (1-p)^{x+1}$$

e.g. $P(X > 2) = P(X \ge 3)$ i.e. the # failures is at least 3

What is $P(X > Y)$? Let $Z = \mathbb{1}_{X>Y} = g(X,Y)$

$$= E[Z] = \sum_{y \in \mathbb{N}_0} \sum_{x \in \mathbb{N}_0} \mathbb{1}_{x>y}\, p_{X,Y}(x,y) = p^2 \sum_{y \in \mathbb{N}_0} (1-p)^y \sum_{x \in \mathbb{N}_0} (1-p)^x \mathbb{1}_{x>y}$$

Since $X, Y \overset{iid}{\sim}$ $p_{X,Y}(x,y) = p_X(x)\, p_Y(y) = p(1-p)^x p(1-p)^y$



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

let $x' = x-(y+1) = x-y-1 \Rightarrow x = x'+y+1$

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

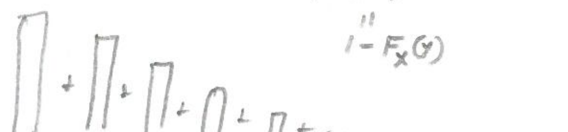
$$= p^2 \sum_{y\in\mathbb{N}_0} (1-p)^{2y+1} \sum_{x'\in\mathbb{N}_0} (1-p)^{x'}$$

Recall $\sum_{k=0}^{\infty} r^k = \frac{1}{1-r}$ if $r\in(0,1)$

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

Another way to think about this...

$$= \sum_{y\in\mathbb{N}_0} \underbrace{p(1-p)^y}_{P(Y=y)} \underbrace{\sum_{x=y+1}^{\infty} p(1-p)^x}_{P(X>y) = 1-F_X(y)}$$



$$\lim_{p\to 0} P(X>Y) = \frac{1}{2}. \text{ Why?}$$

What is $P(X=Y)$? $\quad Z = \mathbb{1}_{X=Y} \quad P(X=Y) = E(Z) = \sum\sum \mathbb{1}_{x=y}\, p_{X,Y}(x,y)$

$$\sum_{y\in\mathbb{N}_0} p(1-p)^y \underbrace{\sum_{x=y}^{y} p(1-p)^x}_{\text{one elem}} = p^2 \sum_{y\in\mathbb{N}_0}(1-p)^{2y} = p^2 \frac{1}{p(2-p)} = \frac{p}{2-p}$$

This is not always pretty...

$X, Y \overset{iid}{\sim} Bin(n,p)$ No closed form for CDF of binomial...

$$P(X>Y) = \sum_{y \in \mathbb{N}_0} P(Y=y)\left(1-F_X(y)\right)$$ $\Rightarrow$ No closed form... must be done numerically. There are approx's but they are complicated...

Basket of apples & bananas. p_1 = prop of apples, p_2 = prop of bananas

Param space $\Rightarrow p_2 = 1-p_1$, $p_1 \in (0,1)$

Experiment: draw one fruit with replacement n times. How many apples?

$X_1 \sim Bin(n, p_1)$

How many bananas? $X_2 = n - X_1$. Together, I can create a vector. $\vec{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$

Not super interesting... What if..

a multidim r.v

Add cantaloupes p_3 = prop of cantaloupes

Param space. $p_1 + p_2 + p_3 = 1$

Let $\vec{X} = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}$

What is $P(\vec{X} = \vec{x})$? This is just the JMF.. Nothing new at all!

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

Multinomial notation!

$\underline{A}$ $\underline{B}$ $\underline{A}$ $\underline{C}$ ..... $\underline{B}$

1 2 3 4 n

~~$X \sim$ Binomial(n, p)~~ ~~Now... more than two classes... Apple, Banana, Cantaloupe~~

~~$\overline{}\ \overline{}\ \overline{}\ \ldots\ \overline{}$~~ ~~$\overline{}\ \overline{}\ \overline{}\ \ldots\ \overline{}$~~

~~1 2 3 ... n~~ ~~n experiments~~

~~$p_A p_B p_C$ s.t. $p_A + p_B + p_C \stackrel{?}{=} 1$~~

~~n experiments, each w.p. p of being a success~~

~~Also read as # of ways of producing x 1's and $n-x$ 0's~~

Generally...

$\vec{X} \sim \text{Multinom}(n, \vec{p}) \iff \binom{n}{x_1, x_2, \ldots x_K} p_1^{x_1} p_2^{x_2} \cdots p_K^{x_K} \mathbb{1}_{\sum x_i = n}$ s.t. $\binom{n}{x_1, \ldots x_K} := \frac{n!}{x_1! \cdots x_K!}$

$\forall x_i \in \mathbb{N},\ \sum x_i \leq n$

o/t $= 0$

$\vec{p}$ ← a multi-dim param of dim K

$\vec{X}$ ← a multi-dim r.v.'s of dim K

$\vec{p} = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_K \end{bmatrix}$

This is the multi-dim'l generalization of the binomial.

Here instead of two categories, success and failure, there are arbitrary K.

$\overline{}\ \overline{}\ \overline{}\ \overline{}\ \ldots\ \overline{}$

1 2 3 4 ... n

$p_1 = \frac{1}{4},\ p_2 = \frac{1}{8},\ p_3 = \frac{5}{8}$

If $K = 3$, $n = 10$. How many ways to have 3 A's, 3 B's & C's

Could be answered by the multinomial distr.

$$P\left(\vec{X} = \begin{bmatrix} 3 \\ 3 \\ 4 \end{bmatrix}\right) = \binom{10}{3, 3, 4} \left(\frac{1}{4}\right)^3 \left(\frac{1}{8}\right)^3 \left(\frac{5}{8}\right)^4$$

What's the param space?

$n \in \mathbb{N}$ same as binomial

$p \in (0,1)^K = \underbrace{(0,1) \times (0,1) \times \ldots \times (0,1)}_{K} \rightarrow$ sets of all K-tuples

s.t. $\vec{p} \cdot \vec{1} = 1$ i.e. $\sum p_k = 1$