Neel Gupta 10/01/2022

Math 505A Homework 5

1) A box has b blue balls and r red balls. ($n = b + r$).
Consider the random vector $X = (X_1, \dots, X_{b+1})$. The range of X are all vectors $(k_1, \dots, k_{b+1})$ with nonnegative integer components that sum to r.

Think of k_i as number of reds before the blue we are at in the i^{th} draw.

Number of ways to order n balls $- n!$

The number of "seats" or events $k_1 - k_{b+1}$ of blue balls is b, so the number of ways to arrange ble balls is $b!$ Similarly number of ways to order red balls is $r!$
Since the location and spacing between balls is represented by vector $\vec{k}$.

Hence, the number of ways to arrange b blue and r red is $(b!)(r!)$.

So for any vector $\vec{k}$, the probability of getting that unique vector $\hat{k}$ is

$$\boxed{\frac{b! \cdot r!}{n!}} = P(k_1, \dots, k_{b+1})$$

HW5

2) a) Given 1000 married couples, find
P("in at least 3, both husband and wife were born on same day")

P(H born on a day) = 1/365

P(W born on a certain day) = 1/365

P(couple born on same day) = $\frac{1}{(365)^2}$

Let X_i be an r.v. representing a couple being born on the same day of the year.

$n = 1000$, $p = \frac{1}{365^2}$ Choose 3 or more
Binom$(1000, \frac{1}{365^2})$ $P(X_i \geq 3)$

$P(X \geq 3) = 1 - P(X < 3)$
num successes $\{0, 1, 2\}$

$P(0) = .99252$, $P(1) = .00745$, $P(2) = .00003$

$P(X < 3) = .99252 + .00745 + .00003 = 1$
$P(X \geq 3) = 1 - P(X < 3) = 1 - 1 = 0$

Poisson distribution: $\lambda = np = \frac{1000}{133225}$

$P(X \geq 3) = 1 - P(X < 3) = 1 - P(2) - P(1) - P(0)$

$P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}$, $P(0) = \frac{e^{-\frac{1000}{365^2}} \left(\frac{1000}{365^2}\right)^0}{0!} = .99252$

$P(X=1) = \frac{e^{-\frac{1000}{365^2}} \left(\frac{1000}{365^2}\right)^1}{1} = .00744$

$P(X=2) = \frac{1}{2} = .00002$

$P(X < 3) = .99252 + .00744 + .00002$
$= .99998$

$P(X \geq 3) = 1 - .99998 = \boxed{.00002}$

Using poisson, $P(X \geq 3) = .00002$
Using binomial, $P(X \geq 3) = 0$

b) When $n = 300000$

Binomial $(300000, 1/365^2)$

$P(X \geq 3) = 1 - P(X < 3)$

$P(X < 3) = P(2, 1, 0)$

$P(X=2) = \text{Binomial}((300000, p)) = .26674$
$P(X=1) = .23691$
$P(X=0) = .10521$
$P(X<3) = .2667 + .23691 + .10521 = .6087$
$P(X \geq 3) = 1 - .6087 = .3913$

Poisson distribution: $\lambda = np = \frac{300000}{(365)^2}$

$P(X \geq 3) = 1 - P(X < 3) = 1 - P(0) - P(1)\ P(2)$
$P(X=0) = e^{\frac{-300000}{365^2}} = .1052$
$P(X=1) = \left(\frac{300000}{365^2}\right) e^{-300000/365^2} = .2369$
$P(X=2) = \frac{(\quad)^2}{2} = .26673$

$P(X<3) = .26673 + .2369 + .1052 = .60883$
$1 - .60883 = .39117$

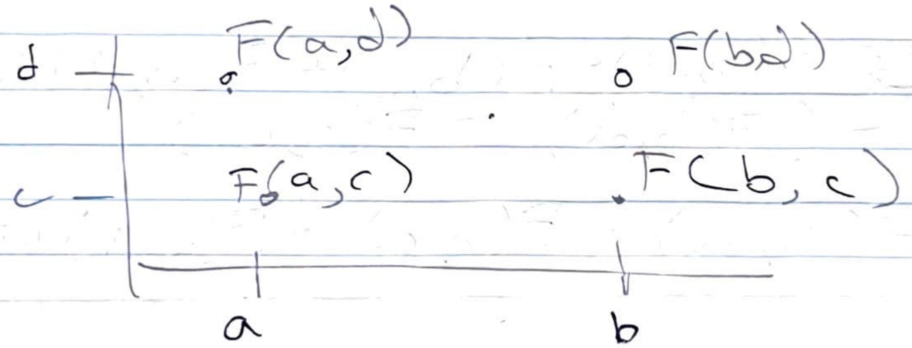
$\boxed{P(X \geq 3) = .39117}$

As n gets very large and p gets very small,
Poisson becomes closer to binomial.

3. a) Let X and Y have joint df F. Show that for any $a < b$ and $c < d$;

$$P(a < X \le b,\ c < Y \le d) = F(b,d) - F(a,d) - [F(b,c) - F(a,c)]$$

$P(a \le X \le b)$ and $P(c < Y \le d)$



Now let $S = \{X \le b,\ c < Y \le d\}$
and $T = \{a < X \le b,\ Y \le d\}$

Then $P(S) = P(\{X \le b, c \le Y \le d\}$

Since F is the df of X and Y, we see that

$F(x) = P(X \le x)$ over $-\infty \le x < \infty$
also $F(b) - F(a) = P(a \le X \le b)$

So $P(S) = F(b,d) - F(b,c)$

and $P(T) = F(b,d) - F(a,d)$

$P(S \cap T) = P(S) + P(T) - P(S \cup T)$

$P(S \cup T) = P(S) + P(T) - P(S \cap T)$

So $P(S \cap T) = \underbrace{F(b,d) - F(b,c)}_{P(S)} +$
$\underbrace{F(b,d) - F(a,d)}_{P(T)} +$
$- P(S \cup T)$

$P(S \cup T) = F(b,d) - F(a,c)$

So $F(b,d) - F(b,c) + F(b,d) - F(a,d)$
$- (F(b,d) - F(a,c))$

$= F(b,d) - F(b,c) - F(a,d) + F(a,c)$

By showing, $P(S \cap T) = \text{eqn }(1)$,

we've shown that $P(a < X \leq b, c < Y \leq d)$

$= F(b,d) - F(b,c) - F(a,d) + F(a,c)$

$P(X = b, Y = d) = P(X \leq b, Y \leq d) -$
$P(X \leq b^{-1}, Y \leq d^{-1})$

Where d^{-1} and b^{-1} are the lefthand limits of F at b, and d.

$F_{xy}(x,y) = P(X \leq x, Y \leq y)$, so

$P(X = b, Y = d) = F(b,d) - F(b^{-1}, d^{-1})$

b) $F(x,y) = 1 - e^{-xy}, \; 0 \leq x, y < \infty$

$F_x(x,y) = ~~+~~ye^{xy}$

$F_{xy}(x,y) = e^{-xy} ~~2~~ xye^{-xy}$

$\lim_{x \to -\infty} F(-\infty, y) = 0 = \lim_{y \to \infty} F(x, -\infty)$

$F(\infty, \infty) = 1 - \frac{1}{e^{\infty}} = 1 - 0 = 1$

$F(x,y) = 1 - \frac{1}{e^{xy}}$

The function e^{xy} is strictly increasing, so $1/e^{xy}$ is strictly decreasing over the domain $(x,y) \in (0, \infty)$, so $F(x,y)$ is a non decreasing function.

Since we've seen that F is twice differential and non decreasing, $F(x,y)$ can be the cdf of any r.v.'s X, Y.

So $F(x,y)$ can be the joint pdf for any variable X, Y s.t.

$F(x,y)$ represents $P(X \leq x, Y \leq y)$,

4. Let $X_1, X_2, \ldots, X_n$ be identical continuous random variables w/ same df F and pdf f.

Assume that $P(X_1 \le x_1, X_2 \le x_2, \ldots, X_n \le x_n)$
$= P(X_1 \le x_1)P(X_2 \le x_2) \ldots P(X_n \le x_n)$

Such a collection is called a random sample w/ n independent "observations" of X w/ df F and pdf f.

a) Show that $(X_1 \ldots X_n)$ is jointly continuous

with the df: $G(x_1, \ldots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f(u_1) \ldots f(u_n)$

Let $G(\vec{x})$ be the joint df of all x values in a vector.

So, $G(\vec{x}) = P(X_1 \le x_1, X_2 \le x_2, X_3 \le x_3 \ldots X_n \le x_n)$

By independence, $G(\vec{x}) = P(X_1 \le x_1) P(X_2 \le x_2) \ldots P(X_n \le x_n)$

By definition of df, $G(\vec{x}) =$

$$\int_{-\infty}^{x_1} f(u_1)\,du_1 \times \int_{-\infty}^{x_2} f(u_2)\,du_2 \times \ldots \times \int_{-\infty}^{x} f(u_n)\,du_n$$

$$= \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_n} f(u_1) f(u_2) \ldots f(u_n)\,du_n \ldots du_1$$

Since $u_1, u_2, \cdots u_n$ are not connected.

Due to independence, $G(\vec{x})$ can further be written $G(\vec{x}) = F(x_1) F(x_2) \ldots F(x_n)$

$G(\vec{x}) = F(x_1) \cdots F(x_n)$

By definition of joind pdf,

$G(\vec{x}) \Rightarrow \left(\begin{array}{c}\text{the derivative} \\ \text{of all } x_1, \dots x_n\end{array}\right) G(\vec{x})$

Where $G(\vec{x})$ is the series of integrals over all $v_1, \dots v_n$.

$G(\vec{x}) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f(v_1) \cdots f(v_n)\, \partial v_n \dots \partial v_1$

$\frac{\partial}{\partial v_1 \cdots \partial v_n} \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f(v_1) \cdots f(v_n)\, \partial v_n \dots \partial v_1$

$= f(x_1) f(x_2) \cdots f(x_n), (x_1 \dots x_n) \in \mathbb{R}^n$

We also see that $X_1, \dots, X_n$ are jointly continuous. $\square$

b) b) $P(X_i = X_j)$ wher $i \neq j$

$P(X_i = X_j) = \int_{\mathbb{R}} P(X_i = X_j \mid X_i = x_i) f_{X_i}(x_i)\, dx_i$

$= \int_{\mathbb{R}} P(X_j = x_i) f_{X_i}(x_i)\, dx_i$, so $P(X_i = X_j) = 0$

$= 0$, $\therefore P(X_j = x_i) = 0$ because of the continuous function that is X_j.

So $P(X_i \neq X_j) = 1 - P(X_i = X_j)$

$= 1 - 0, = 1$, So for all $i \neq j$, the Sample values are distinct.

5. Let X and Y be independent r.v.'s taking positive integers ($\mathbb{Z}$) with the same mass function $f(x) = 2^{-x}$ for all $x \in \{1, 2 \ldots\}$, that is they are geometric with $p = 1/2$.

a) $P(\min(X,Y) \leq x) = 1 - P(\min\{X,Y\} > x)$

$= 1 - P(X > x, Y > \cancel{x}) = 1 - P(X > x) P(Y > x)$

Since both X & Y have pmf $f(x) = 2^{-x}$

$\Rightarrow 1 - 2^{-x}(2^{-x}) = \boxed{1 - 4^{-x}}$

b) $P(Y > X) = P(X > Y)$ by symmetry.

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

$P(X = Y) = \sum_{k=1}^{\infty} P(X = k, Y = k)$

$= \sum_{k=1}^{\infty} P(X=k) P(Y=k)$ by independence

$= \sum_{k=1}^{\infty} (f(k) f(k)) = \sum_{k=1}^{\infty} 4^{-k}$

$= \dfrac{4^{-k}}{1 - 4^{k}} = \dfrac{\frac{1}{4}}{\frac{4}{4} - \frac{1}{4}} = \dfrac{1}{\cancel{4}} \cdot \dfrac{\cancel{4}}{3} = \dfrac{1}{3}$

$P(X \neq Y) = 1 - P(X = Y) = 2/3$, split by 2

$= P(Y > X)$ by symmetry

$P(Y > X)$, $P(Y = X)$, $P(Y < X)$ are all $1/3$

$\boxed{P(Y > X) = 1/3}$

c) From part(b), $P(X=Y) = 1/3$

d) $P(X \geq kY) = \sum_{y=1}^{\infty} P(X \geq kY, Y=y)$

$$= \sum_{y=1}^{\infty} P(X \geq ky) \cdot P(Y=y) \text{ by independence}$$

$$= \sum_{y=1}^{\infty} \sum_{x=0}^{\infty} 2^{-1(ky+x)} 2^{-y}$$

$$= \frac{2}{2^{k+1}-1}$$

e) P("X divides Y") $=$

$$\sum_{k=1}^{\infty} P(Y=kX) = \sum_{k=1}^{\infty} \sum_{x=1}^{\infty} P(X=x, Y=kx)$$

By independence $= \sum_{k=1}^{\infty} \sum_{x=1}^{\infty} f(x) f(kx)$

$$= \sum_{k=1}^{\infty} \sum_{x=1}^{\infty} 2^{-kx} 2^{-x} = \sum_{k=1}^{\infty} \frac{1}{2^{k+1}-1}$$

$$= \boxed{\sum_{k=1}^{\infty} \frac{1}{2^{k+1}-1}}$$