Statistics 100A
Homework 7 Solutions

Ryan Rosario

Chapter 6

11. *A television store owner figures that 45 percent of the customers entering his store will purchase an ordinary television set, 15 percent will purchase a color television set, and 40 percent will just be browsing. If 5 customers enter his store on a given day, what is the probability that he will sell exactly 2 ordinary sets and 1 color set on that day?*

    In this problem, we take a look at a distribution that we have not studied in this course, the **multinomial** distribution. The PDF for the multinomial distribution is below

$$P(X_1 = x_1, \ldots, X_r = x_r) = \frac{n!}{n_1!n_2!\ldots n_r!} p_1^{n_1} p_2^{n_2} \ldots p_r^{n - \sum n_i}$$

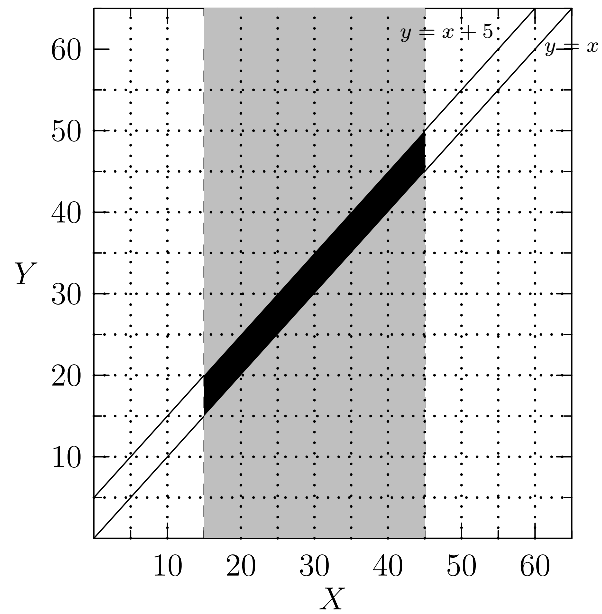
    The probability that exactly 2 ordinary sets ($X_1 = 2$) and 1 plasma set ($X_2 = 1$) are purchased (and the rest $X_3 = 2$ are just browsing) is given by

$$\begin{aligned} P(X_1 = 2, X_2 = 1, X_3 = 2) &= \frac{5!}{2!1!2!}(0.45)^2(0.15)(0.4)^2 \\ &= \boxed{0.1457} \end{aligned}$$

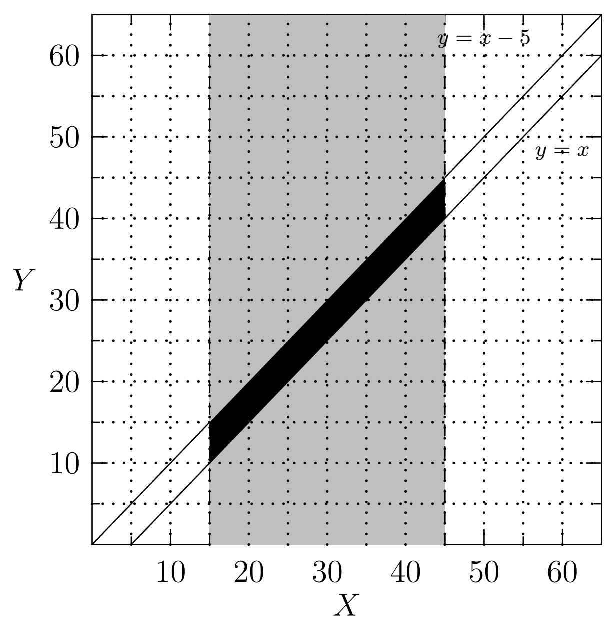
13. *A man and a woman agree to meet at a certain location about 12:30 p.m. If the man arrives at a time uniformly distributed between 12:15 and 12:45 and the woman independently arrives at a time uniformly distributed between 12:30 and 1 pm, find the probability that the man arrives first?*

    The easiest way to illustrate this problem is to consider a geometric representation. Let X be the number of minutes past noon that the man arrives, and let Y represent the number of minutes past noon that the woman arrives. From the problem, we know that $15 \le X \le 45$ and $0 \le Y \le 60$. First we find the probability that the first to arrive waits no longer than 5 minutes for the other person. There are two ways to do this: either the man can arrive first and waits for the woman, or vice-verse.

    First consider the situation that the man arrives first. The man arrives sometime between 15 minutes and 45 minutes past noon. The acceptable time for the woman to arrive is no later than 5 minutes after the man. We can represent this situation geometrically. Given that the man arrives between 15 and 45 minutes after noon, the woman independently arrives some time within the hour. This area is shaded in light grey and has area 1800. So that the man does not get impatient, the woman must arrive before 5 minutes have passed. The minutes for the man and woman to arrive such that the man does not become impatient is shaded in black and has area 150 (a parallelogram). The probability that the woman arrives no more than 5 minutes after the man is $\frac{150}{1800} = \frac{1}{12}$.



Next, consider the situation that the woman arrives first. The woman arrives some time between 0 minutes and 60 minutes past noon. The acceptable time for the man to arrive is no later than 5 minutes after the woman. Given that the woman arrives some time during the hour, the man independently arrives some time between 15 minutes and 45 minutes past the hour. This area is shaded in light grey and has area 1800. So that the woman does not get mad, the man must arrive before 5 minutes have passed. The minutes for the man and woman to arrive such that the woman does not get mad is shaded in black and has area 150 as well. The probability that the man arrives no more than 5 minutes after the woman is $\frac{150}{1800} = \frac{1}{12}$.



$$P\left(|X - Y| \leq 5\right) = P(X < Y < X + 5) + P(Y < X < Y + 5) = \frac{1}{12} + \frac{1}{12} = \boxed{\frac{1}{6}}$$

Notice the probability that the man waits no more than 5 minutes for the woman and vice-verse are the same; equally likely. The probability that the man arrives first is $\boxed{\frac{1}{2}}$.

15. *The random vector (X, Y) is said to be uniformly distributed over a region R in the plane if, for some constant c, its joint density is*

$$f(x,y) = \begin{cases} c & \text{if } (x,y) \in R \\ 0 & \text{otherwise} \end{cases}$$

(a) Show that $\frac{1}{c}$ = area of region R.

Proof. Since $f(x,y)$ is a joint density function, $\int\int_R c\,dydx = 1$. But recall from Math 32B the double integral over a region R is the area of R, call it $A(R)$. Then,

$$\int\int_R c\,dydx = c \cdot A(R) = 1$$

so $A(R) = \frac{1}{c}$. □

(b) *Suppose that (X, Y) is is uniformly distributed over the square centered at $(0, 0)$, whose sides are of length 2. Show that X and Y are independent, with each being distributed uniformly over $(-1, 1)$.*

All we need to do is show that X and Y are independent.

Proof. Since X and Y are uniformly distributed over the same domain $(-1, 1)$, we know that

$$f_X(x) = f_Y(y) = \frac{1}{1-(-1)} = \frac{1}{2}, \qquad -1 < x < 1, -1 < y < 1$$

From part (a), we know that $A(R) = \frac{1}{c}$. If the square is centered at $(0, 0)$ with each side having length 2, then the area of the square is 4. Therefore, $A(R) = \frac{1}{c} = 4$ and $c = \frac{1}{4}$, so then $f(x,y) = c = \frac{1}{4}, (x,y) \in R$.

We see that $f(x,y) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = f_X(x) \cdot f_Y(y)$, thus X and Y are independent.

□

(c) *What is the probability that that (X, Y) lies in the circle of radius 1 centered at the origin? That is, find $P\left(X^2 + Y^2 \le 1\right)$.*

$$\begin{aligned} P(X^2+Y^2 \le 1) &= \int\int_{X^2+Y^2\le 1} c\,dxdy \\ &= c \cdot \text{Area of Circle} \\ &= \frac{\text{Area of Circle}}{\text{Area of Square}} \end{aligned}$$

since $c = \frac{1}{\text{area of square}}$.

The area of the square was 4. The area of the unit circle is $\pi r^2 = \pi$. So,

$$\boxed{P(X^2+Y^2 \le 1) = \frac{\pi}{4}}$$

18. *Two points are selected randomly on a line of length L so as to be on opposite sides of the midpoint of the line. Find the probability that the distance between the two points is greater than $\frac{L}{3}$.*

Let X represent the leftmost point. Its position is uniformly distributed between 0 and $\frac{L}{2}$. Let Y represent the rightmost point. Its position is uniformly distributed between $\frac{L}{2}$ and L. We want to find the probability that $P\left(Y - X > \frac{L}{3}\right)$. Note that $P\left(|X - Y| > \frac{L}{3}\right)$ is the same thing, but since we assigned Y to be the leftmost point, the first representation is also appropriate.

$$\int\int_{Y-X>\frac{L}{3}} \left(\frac{1}{\frac{L}{2}-0}\right)\left(\frac{1}{L-\frac{L}{2}}\right) dydx = \int\int_{Y-X>\frac{L}{3}} \frac{4}{L^2} dydx$$

We have to be careful here. There are two possible scenarios. First, we can fix X in its domain $0 < x < \frac{L}{2}$ and let Y vary within its domain **restricted to the condition that the distance between the points is greater than** $\frac{L}{3}$, so the restricted domain would be $0 < x < \frac{L}{6}$.

$$\begin{aligned}
P\left(Y - X > \frac{L}{3}\right) &= \frac{4}{L^2}\int\int_{Y-X>\frac{L}{3}} dxdy \\
&= \frac{4}{L^2}\left[\int_0^{\frac{L}{6}}\int_{\frac{L}{2}}^{L} dydx + \int_{\frac{L}{6}}^{\frac{L}{2}}\int_{x+\frac{L}{3}}^{L} dydx\right] \\
&= \frac{4}{L^2}\left[\frac{L^2}{12} + \frac{5}{24}L^2 - \frac{7}{72}L^2\right] \\
&= \boxed{\frac{7}{9}}
\end{aligned}$$

19. *Show that $f(x,y) = \frac{1}{x}, 0 < y < x < 1$ is a joint density function. Assuming that f is the joint density function of X, Y, find*

First we show that $f(x,y)$ is a joint density function.

Proof.

$$\int_0^1\int_0^x \frac{1}{x} dydx = \int_0^1 \left.\frac{y}{x}\right|_0^x dx = \int_0^1 dx = x|_0^1 = 1$$

□

(a) *the marginal density of Y.*

$$\begin{aligned}
f_Y(y) &= \int_y^1 \frac{1}{x} dx = \log x|_y^1 \\
&= \boxed{-\log y}
\end{aligned}$$

(b) *the marginal density of X.*

$$\begin{aligned} f_X(x) &= \int_0^x \frac{1}{x} dy \\ &= \left. \frac{y}{x} \right|_0^x \\ &= \boxed{1} \end{aligned}$$

(c) $E(X)$

$$\begin{aligned} E(X) &= \int_0^1 x dx \\ &= \left. \frac{x^2}{2} \right|_0^1 \\ &= \boxed{\tfrac{1}{2}} \end{aligned}$$

(d) $E(Y)$

$$\begin{aligned} E(Y) &= \int_0^1 y\left(-\log y\right) dy \\ &= -\int_0^1 y \log y dy \\ & \text{Integrate by parts: let } u = \log y, dv = y \text{ so } du = \tfrac{1}{y} dy, v = \tfrac{y^2}{2}. \\ &= -\left[\frac{y^2 \log y}{2} - \frac{y^2}{4}\right]_0^1 \\ &= \boxed{\tfrac{1}{4}} \end{aligned}$$

20. *The joint density of X and Y is given by*

$$f(x,y) = \begin{cases} xe^{-(x+y)} & x > 0, y > 0 \\ 0 & \text{otherwise} \end{cases}$$

Are X and Y independent? What if $f(x,y)$ were given by

$$f(x,y) = \begin{cases} 2 & 0 < x < y, 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

X and Y are independent if $f(x,y) = f_X(x) f_Y(y)$. We already have $f(x,y)$ so we find the marginals.

$$\begin{aligned} f_X(x) &= \int_0^\infty xe^{-(x+y)} dy \\ &= \left. -xe^{-(x+y)} \right|_0^\infty \end{aligned}$$

$$= xe^{-x}$$

and,

$$\begin{aligned} f_Y(y) &= \int_0^{\infty} xe^{-(x+y)}dx \\ &= -\left[xe^{-(x+y)} + e^{-(x+y)}\right]_0^{\infty} \\ &= -\left[e^{-(x+y)}(x+1)\right]_0^{\infty} \\ &= -\left[0 - e^{-y}\right] \\ &= e^{-y} \end{aligned}$$

$$f_X(x) \cdot f_Y(y) = xe^{-x}e^{-y} = xe^{-(x+y)} = f(x,y)$$

$\boxed{\text{Yes}}$, X and Y are independent.

For the other PDF given, we do the same thing.

$$\begin{aligned} f_X(x) &= \int_x^1 2dy \\ &= 2y|_x^1 \\ &= 2 - 2x \\ &= 2(1-x) \end{aligned}$$

and

$$\begin{aligned} f_Y(y) &= \int_0^y 2dx \\ &= 2x|_0^y \\ &= 2y \end{aligned}$$

$$f_X(x) \cdot f_Y(y) = 2(1-x) \cdot 2y \neq 2 = f(x,y)$$

$\boxed{\text{No}}$, X and Y are not independent.

23. *The random variables X and Y have joint density function,*

$$f(x,y) = \begin{cases} 12xy(1-x) & 0 < x < 1, 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

Before considering the problems that follow, let's find the marginal densities of X and Y.

We want $f_X(x)$ to be in terms of x, so integrate out y and choose limits such that the result is in terms of x. Here we can simply use the domain of x.

$$\begin{aligned} f_X(x) &= \int_0^1 12xy(1-x)dy \\ &= 12x(1-x)\int_0^1 ydy \\ &= 12(1-x)x\left[\frac{y^2}{2}\right]_0^1 \\ &= 6x(1-x) \end{aligned}$$

We want $f_Y(y)$ to be in terms of y, so integrate out x and choose limits such that the result is in terms of y. Here we can simply use the domain of y.

$$\begin{aligned} f_Y(y) &= \int_0^1 12xy(1-x)dx \\ &= 12y\int_0^1 x - x^2dx \\ &= 12y\left[\frac{x^2}{2} - \frac{x^3}{3}\right]_0^1 = 12y \\ &= 12y\left(\frac{1}{6}\right) \\ &= 2y \end{aligned}$$

(a) *Are X and Y independent?*

Again, we check whether or not $f(x,y) = f_X(x) \cdot f_Y(y)$.

$$f_X(x) \cdot f_Y(y) = 12xy(1-x) = f(x,y)$$

$\boxed{\text{Yes}}$, X and Y are independent.

(b) *Find $E(X)$.*

$$\begin{aligned} E(X) &= \int_a^b xf_X(x)dx \\ &= \int_0^1 6x^2(1-x)dx \\ &= 6\int_0^1 x^2 - x^3dx \end{aligned}$$

$$\begin{aligned}
&= 6\left[\frac{x^3}{3} - \frac{x^4}{4}\right]_0^1 \\
&= 6\left[\frac{1}{3} - \frac{1}{4}\right] \\
&= \boxed{\tfrac{1}{2}}
\end{aligned}$$

(c) *Find $E(Y)$.*

$$\begin{aligned}
E(Y) &= \int_0^1 y f_Y(y)dy \\
&= \int_0^1 2y^2 dy \\
&= \frac{2}{3}y^3\Big|_0^1 \\
&= \boxed{\tfrac{2}{3}}
\end{aligned}$$

(d) *Find Var(X).*

Recall that

$$\text{Var}\,(X) = E(X^2) - E(X)^2$$

We already have $E(X)$. To find $E(X^2)$ (the second moment), we compute

$$E(X^2) = \int_a^b x^2 f_X(x)dx$$

$$\begin{aligned}
E(X^2) &= \int_0^1 6x^3(1-x)dx \\
&= 6\int_0^1 x^3 - x^4 dx \\
&= 6\left[\frac{x^4}{4} - \frac{x^5}{5}\right]_0^1 \\
&= 6\left[\frac{1}{4} - \frac{1}{5}\right] \\
&= \frac{3}{10}
\end{aligned}$$

and we already know that $E(X) = \frac{1}{2}$, so

$$\begin{aligned}
\text{Var}\,(X) &= E(X^2) - (X)^2 \\
&= \frac{3}{10} - \left(\frac{1}{2}\right)^2
\end{aligned}$$

$$
\begin{aligned}
&= \frac{3}{10} - \frac{1}{4} \\
&= \boxed{\tfrac{1}{20}}
\end{aligned}
$$

(e) *Find Var(Y).*

Again, $\operatorname{Var}(Y) = E(Y^2) - E(Y)^2$.

$$
\begin{aligned}
\operatorname{Var}(Y) &= E(Y^2) - E(Y)^2 \\
&= \int_0^1 2y^3 dy - \left(\frac{2}{3}\right)^2 \\
&= \frac{1}{2}y^4\Big|_0^1 - \frac{4}{9} \\
&= \frac{1}{2} - \frac{4}{9} \\
&= \boxed{\tfrac{1}{18}}
\end{aligned}
$$

28. *If X_1 and X_2 are independent exponential random variables with respective parameters λ_1 and λ_2, find the distribution of $Z = \frac{X_1}{X_2}$. Also compute $P(X_1 < X_2)$.*

First, note that

$$f_{xy}(x, y) = \lambda_1 \lambda_2 e^{-\lambda_1 x_1} e^{-\lambda_2 x_2}$$

We start with the CDF for $Z = \frac{X_1}{X_2}$ and take the derivative to get f_Z. Take $a = \frac{X_1}{X_2}$. Then,

$$\begin{aligned}
F_Z(a) &= P\left(\frac{X_1}{X_2} \le a\right) \\
&= P(X_1 \le aX_2) \\
&= \int_0^\infty \int_0^{ax_2} \lambda_1 \lambda_2 e^{-\lambda_1 x_1} e^{-\lambda_2 x_2} dx_1 dx_2 \\
&= \int_0^\infty \lambda_1 \lambda_2 e^{-\lambda_2 x_2} \left[-\frac{1}{\lambda_1} e^{-\lambda_1 x_1}\right]_0^{ax_2} dx_2 \\
&= \lambda_2 \int_0^\infty e^{-\lambda_2 x_2} \left(1 - e^{-\lambda_1 a x_2}\right) dx_2 \\
&= \lambda_2 \int_0^\infty e^{-\lambda_2 x_2} - e^{-x_2(\lambda_2 + \lambda_1 a)} dx_2 \\
&= \lambda_2 \left[\frac{1}{\lambda_2 + \lambda_1 a} e^{-y(\lambda_2 + \lambda_1 a)} - e^{-\lambda_2 x_2}\right]_0^\infty \\
&= -\left[\frac{\lambda_2}{\lambda_2 + \lambda_1} - 1\right] \\
&= 1 - \frac{\lambda_2}{\lambda_2 + \lambda_1 a} \\
&= \frac{\lambda_1 a}{\lambda_2 + \lambda_1 a}
\end{aligned}$$

Then,

$$f_Z(a) = F'_Z(a) = \frac{dF}{da} = \frac{\lambda_1 \lambda_2}{(\lambda_1 a + \lambda_2)^2}$$

Next,

$$\begin{aligned}
P(X_1 < X_2) &= P\left(\frac{X_1}{X_2} < 1\right) \\
&= \boxed{\tfrac{\lambda_1}{\lambda_2 + \lambda_1}}
\end{aligned}$$

30. *The expected number of typographical errors on a page of a certain magazine is .2. What is the probability that an article of 10 pages contains*

We know that the error rate is 0.2 per page. Let X_i be the number of errors on page i. Then X_i is Poisson with rate $\lambda_i = 0.2$. **We are working with a sum of independent Poisson random variables.** The sum, $X = \sum_{i=1}^{10} X_i \sim \text{Poisson}(\lambda = \sum_{i=1}^{10} \lambda_i = 10(0.2) = 2)$. Then,

$$P(X = x) = \frac{2^x e^{-2}}{x!}$$

(a) *0 typographical errors.*

From the above, $P(X = 0) = \frac{2^0 e^{-2}}{0!} = \boxed{e^{-2}}$.

(b) *2 or more typographical errors.*

$$\begin{aligned} P(X \geq 2) &= 1 - P(X = 0) - P(X = 1) \\ &= 1 - e^{-2} - \frac{2e^{-2}}{1!} \\ &= 1 - e^{-2} - 2e^{-2} \\ &= 1 - 3e^{-2} \\ &= \boxed{0.593} \end{aligned}$$

31. *The monthly worldwide average number of airplane crashes of commercial airlines is 2.2. What is the probability that there will be*

Again, we have a monthly average rate of plane crashes. Assuming the probability of a plane crash is very small, we use the **Poisson** distribution. For one month, $\lambda_1 = 2.2$. Let X_1 be the number of airplane crashes in one month. So,

$$P(X_1 = x_1) = \frac{2.2^{x_1} e^{-2.2}}{x_1!}$$

(a) *more than 2 such accidents in the next month?*

We are still considering only one month. So,

$$\begin{aligned} P(X > 2) &= 1 - P(X = 0) - P(X = 1) - P(X = 2) \\ &= 1 - e^{-2.2} - 2.2e^{-2.2} - \frac{2.2^2 e^{-2.2}}{2} \\ &= 1 - e^{-2.2}\left[1 + 2.2 + \frac{2.2^2}{2}\right] \\ &= 1 - 4.41e^{-2.2} \\ &\approx \boxed{0.5113} \end{aligned}$$

(b) *more than 4 such accidents in the next 2 months?*

Now we are considering the next *two* months. We are working with the **sum of two independent Poisson random variables.** Let X_2 be the number of airplane crashes in the second month and let X be the number of airplane crashes in two months, such that $X = X_1 + X_2$. X_2 follows the same distribution with the same mean as X_1. Then, X is a Poisson random variable with $\lambda = \lambda_1 + \lambda_2 = 4.4$. Why?

$$E(X) = E(X_1 + X_2) = E(X_1) + E(X_2) = \lambda_1 + \lambda_2 = 2.2 + 2.2 = 4.4$$

Then,

$$\begin{aligned} P(X > 4) &= 1 - P(X=0) - P(X=1) - P(X=2) - P(X=3) - P(X=4) \\ &= 1 - e^{-4.4}\left[1 + 4.4 + \frac{(4.4)^2}{2!} + \frac{(4.4)^3}{3!} + \frac{(4.4)^4}{4!}\right] \\ &\approx 1 - 0.55e^{-4.4} \\ &\approx \boxed{0.993} \end{aligned}$$

(c) *more than 5 such accidents in the next 3 months.*

Now we are considering the next *three* months. We are working with the **sum of three independent Poisson random variables.** Let X_3 be the number of airplane crashes in the third month and let X be the number of airplane crashes in three months, such that $X = X_1 + X_2 + X_3$. X_3 follows the same distribution with the same mean as X_1 and X_2. Then, X is a Poisson random variable with $\lambda = \lambda_1 + \lambda_2 + \lambda_3 = 6.6$.

Then,

$$\begin{aligned} P(X > 5) &= 1 - P(X=0) - P(X=1) - P(X=2) - P(X=3) - P(X=4) - P(X=5) \\ &= 1 - e^{-6.6}\left[1 + 6.6 + \frac{(6.6)^2}{2!} + \frac{(6.6)^3}{3!} + \frac{(6.6)^4}{4!} + \frac{(6.6)^5}{5!}\right] \\ &\approx 1 - 260.7e^{-6.6} \\ &\approx \boxed{0.645} \end{aligned}$$

32. *The gross weekly sales at a certain restaurant is a normal random varaible with mean* \$2200 *and standard deviation* \$230*. What is the probability that*

(a) *the total gross sales over the next 2 weeks exceeds* \$5000*?*

Let X_1 be the restaurant's profit for week 1. Let X_2 be the total gross sales during week 2. We want to find the probability that the total gross sales over the next 2 weeks exceeds \$5000 so we need a new random variable to represent the total gross sales over *two* weeks. We assume that X_1 and X_2 follow the same distribution with the same mean $\mu = 2200$ and standard deviation $\sigma = 2200$ and note they are independent. Let X be the **sum of two normal random variables**. That is $X = X_1 + X_2$. Then, we need to find the mean and SD of this new random variable.

$$E(X) = E(X_1 + X_2) = E(X_1) + E(X_2) = \mu + \mu = 2\mu = 2(2200) = 4400$$

$$\sigma_X = \sqrt{\text{Var}(X)} = \sqrt{\text{Var}(X_1 + X_2)} = \sqrt{\text{Var}(X_1) + \text{Var}(X_2)} = \sqrt{2\sigma^2} = \sqrt{2}\sigma = 2(230) = 460$$

Thus, $X \sim N(4400, 460)$.

$$
\begin{aligned}
P(X > 5000) &= 1 - P(X \leq 5000) \\
&= 1 - P\left(Z \leq \frac{5000 - 4400}{460}\right) \\
&= 1 - P(Z \leq 1.845) \\
&= 1 - 0.967 \\
&= \boxed{0.0326}
\end{aligned}
$$

(b) *weekly sales exceed* \$2000 *in at least 2 of the next 3 weeks?*

This is a two-step problem. We want to find the probability that sales exceed the given amount in at least 2 of the 3 next weeks, so we will want to use the **binomial** distribution with $n = 3$. The first step involves finding p.

Find p.

To find p, consider the event that weekly sales exceed \$2000 in an arbitrary week. Let X be the total gross sales in some arbitrary week. Then $X \sim N(2200, 230)$, and,

$$
\begin{aligned}
p &= P(X > 2000) \\
&= 1 - P(X \leq 2000) \\
&= 1 - P\left(Z \leq \frac{2000 - 2200}{230}\right) \\
&= 1 - P(Z < -0.87) \\
&= 1 - .1922 \\
&= \boxed{0.808}
\end{aligned}
$$

Find $P(Y \geq 2)$.

Let Y be the number of weeks that total gross sales exceed \$2000. Then,

$$
\begin{aligned}
P(Y \geq 2) &= P(Y = 2) + P(Y = 3) \\
&= \binom{3}{2}(0.808)^2(0.192) + \binom{3}{3}(0.808)^3 \\
&= \boxed{0.9036}
\end{aligned}
$$

34. *According to the U.S. National Center for Health Statistics, 25.2 percent of males and 23.6 percent of females never eat breakfast. Suppose that random samples of 200 men and 200 women are chosen. Approximate the probability that*

This is a **binomial** problem, but since n is large and we are asked for an approximation, we use the **normal approximation to the binomial.** Let X be the number of men that never eat breakfast and let Y be the number of women that never eat breakfast. Note that $n = 200$ for both the men and the women. Then,

$$E(X) = np_X = 200 \cdot 0.252 = 50.4$$

$$E(Y) = np_Y = 200 \cdot 0.236 = 47.2$$

$$\text{Var}(X) = np_X(1-p_X) = 37.7 \Rightarrow \sigma_X = 6.14$$
$$\text{Var}(Y) = np_Y(1-p_Y) = 36.06 \Rightarrow \sigma_Y = 6$$

(a) *at least 110 of those 400 people never eat breakfast.*

In this part we collapse the men and women into one group. That is, we create a new variable $W = X + Y$. We need to find the distribution of W.

$$E(W) = E(X+Y) = E(X) + E(Y) = 50.4 + 47.2 = 97.6$$
$$\sigma_W = \sqrt{\text{Var}(X+Y)} = \sqrt{\text{Var}(X) + \text{Var}(Y)} = \sqrt{37.7 + 36.06} = \sqrt{73.76} = 8.59$$

So $W \sim N(97.6, 8.59)$. Then,

$$\begin{aligned} P(W \geq 110) &= 1 - P(W < 110) \\ &= 1 - P\left(Z < \frac{110 - 0.5 - 97.6}{8.59}\right) \\ &= 1 - P(Z < 1.385) \\ &= 1 - 0.917 \\ &= \boxed{0.0829} \end{aligned}$$

(b) *the number of the women who never eat breakfast is at least as large as the number of the men who never eat breakfast.*

We want to find $P(Y \geq X) = P(Y - X \geq 0)$. We need to find the distribution of $Y - X$. Note that the standard deviation of $X + Y$ and $Y - X$ is the same.

$$E(Y - X) = E(Y) - E(X) = 50.4 - 47.2 = -3.2$$

Then,

$$\begin{aligned} P(Y - X \geq 0) &= P\left(Z \geq \frac{0 - 0.5 + 3.2}{8.59}\right) \\ &= P(Z \geq 0.3143) \\ &= 1 - P(Z < 0.3143) \\ &= \boxed{0.3766} \end{aligned}$$

35. *In Problem 2, calculate the conditional probability mass function of X_1 given that*

(a) $X_2 = 1$

We want to find the conditional probability of X_1 given that the second ball chosen is white ($X_2 = 1$). We must calculate this for both parts of the original problem. We really only need to find the probability for $X_1 = 1$ since the probability of $X_1 = 0$ is just the complement.

i. $P(X_1 = 1|X_2 = 1) = \frac{P(X_1=1,X_2=1)}{P(X_2=1)} = \frac{p(1,1)}{p(0,1)+p(1,1)} = \frac{5\cdot 4}{8\cdot 5+5\cdot 4} = \frac{20}{60} = \boxed{\frac{1}{3}}$

ii. $P(X_1 = 1|X_2 = 1) = \frac{p(1,1,0)+p(1,1,1)}{p(1,1,1)+p(0,1,1)+p(1,1,0)+p(0,1,0)} = \boxed{\frac{1}{3}}$

(b) $X_2 = 0$

i. $P(X_1 = 1|X_2 = 0) = \frac{P(X_1=1,X_2=0)}{P(X_2=0)} = \frac{p(1,0)}{p(1,0)+p(0,0)} = \boxed{0.417}$

ii. $P(X_1 = 1|X_2 = 0) = \frac{P(X_1=1,X_2=0)}{P(X_2=0)} = \frac{p(1,0,0)+p(1,0,1)}{p(1,0,1)+p(1,0,0)+p(0,0,1)+p(0,0,0)} = \boxed{0.417}$

39. *Choose a number X at random from the set of numbers $\{1, 2, 3, 4, 5\}$. Now choose a number at random from the subset no larger than X, that is, from $\{1, , X\}$. Call this second number Y.*

(a) *Find the joint mass function of X and Y.*

Note that this sampling method has two steps. Once we draw a value for X, then we draw a value of Y. That is, Y is conditional on X. Then, by Bayes' rule,

$$P_{Y|X}(x|y) = \frac{P(x, y)}{P_X(x)}$$

Then,

$$P(x, y) = P_{X|Y}(x|y)P_X(x)$$

Start with $P_X(x = i)$. In the set $\{1, 2, 3, 4, 5\}$ each number is equally likely to be chosen for the value of X, so $P_X(x = i) = \frac{1}{5}$.

Once we know which value is chosen for X, we choose a value for Y that is not greater than X. Each number in the subset is equally likely, so

$$P_{Y|X}(y = j|x = i) = \frac{1}{j}, \qquad 1 < j < i, 1 < i < 5$$

Then, $\boxed{P(x, y) = \frac{1}{5j}, \quad 1 < j < i, \quad 1 < i < 5}$.

(b) *Find the conditional mass function of X given that $Y = i$. Do it for $i = 1, 2, 3, 4, 5$.*

$$P_{X|Y}(X = j|Y = i) = \frac{P(X = j, Y = i)}{P_Y(Y = i)}$$

To find the marginal, think about the original problem. If $Y = 1$, this means that $X = 1$ by the specification of the problem, so the probability of drawing $X = 1$ given that $Y=2$ is 1. For $Y=2$, X can be 1 or 2. The probability of drawing $X=j$ given that $Y=2$ is $\frac{2}{5}$ etc. So,

$$P_Y(Y=i)=\sum_{k=1}^{5}\frac{k}{5}$$

Then,

$$\boxed{P_{X|Y}(X=j|Y=i)=\frac{\frac{1}{5j}}{\sum_{k=i}^{5}\frac{k}{5}}}$$

(c) Are X and Y independent? Why?

$\boxed{\text{No}}$, if they were independent, this would not be an exercise. X and Y are not independent because the choice of Y is restricted by the choice of X.

$$P(X=j|Y=i)\neq P(X=j)$$

because if they were equal, that would mean that the probability of X being some value does not at all depend on the value of Y, assuming Y is known.

41. *The joint probability mass function of X and Y is given by*

$$p(1,1)=\frac{1}{8},\quad p(1,2)=\frac{1}{4},\quad p(2,1)=\frac{1}{8},\quad p(2,2)=\frac{1}{2}$$

(a) *Compute the conditional mass function of X given $Y=i, i=1,2$.*

$$\begin{aligned}
P(X=1|Y=1) &= \frac{P(X=1,Y=1)}{P(Y=1)}=\frac{p(1,1)}{p(2,1)+p(1,1)}=\frac{\frac{1}{8}}{\frac{1}{8}+\frac{1}{8}}=\frac{1}{2}\\
P(X=2|Y=1) &= 1-P(X=1|Y=1)=\frac{1}{2}\\
P(X=1|Y=2) &= \frac{P(X=1,Y=2)}{P(Y=2)}=\frac{p(1,2)}{p(1,2)+p(2,2)}=\frac{\frac{1}{4}}{\frac{1}{4}+\frac{1}{2}}=\frac{1}{3}\\
P(X=2|Y=2) &= 1-P(X=1|Y=2)=\frac{2}{3}
\end{aligned}$$

(b) *Are X and Y independent?*

Check if $P(X=j,Y=i)=P(X=j)P(X=i)$. Suppose $i=j=2$.

$$\begin{aligned}
P(X=2,Y=2) &= p(2,2)=\frac{1}{2}\\
P(X=2) &= p(2,1)+p(2,2)=\frac{1}{8}+\frac{1}{2}=\frac{5}{8}\\
P(Y=2)= & \quad p(1,2)+p(2,2)=\frac{1}{4}+\frac{1}{2}=\frac{3}{4}
\end{aligned}$$

Note that $\frac{1}{2}\neq\frac{5}{8}\cdot\frac{3}{4}$. $\boxed{\text{No}}$.

(c) *Compute the following...*

$$
\begin{aligned}
P(X+Y>2) &= p(1,2)+p(2,1)+p(2,2) = \frac{1}{4}+\frac{1}{8}+\frac{1}{2} = \boxed{\tfrac{7}{8}} \\
P(XY \leq 3) &= p(1,1)+p(1,2)+p(2,1) = \frac{1}{8}+\frac{1}{4}+\frac{1}{8} = \boxed{\tfrac{1}{2}} \\
P\left(\frac{X}{Y}>1\right) &= P(X>Y) = p(2,1) = \boxed{\tfrac{1}{8}}
\end{aligned}
$$