40.017 PROBABILITY & STATISTICS

Homework 1

Michael Hoon
1006617
Section 2

February 6, 2024

Question 1

(a)

Let A be the random variable denoting the number of questions chosen from section A. Since the student 'samples' the question without replacement, A follows a hypergeometric distribution $A \sim \text{hypgeo}(12, 9, 9)$. The pmf $f(x)$ is given by:

$$f(x) = \mathbb{P}(X = x) = \frac{\binom{j}{x}\binom{k}{n-x}}{\binom{j+k}{n}}$$

for 6 questions from section A, we have $\mathbb{P}(X = 6)$:

$$\begin{aligned}\mathbb{P}(X = 6) &= \frac{\binom{9}{6}\binom{9}{6}}{\binom{18}{12}} \\ &= \frac{84}{221} \\ &\boxed{\approx 0.38}\end{aligned}$$

(b)

Now we need to consider two cases, one where 5 questions come from Section A and the other 7 from Section B, and vice versa. i.e.

$$\begin{aligned}\mathbb{P}(A = 5) + \mathbb{P}(A = 7) &= \frac{\binom{9}{5}\binom{9}{7}}{\binom{18}{12}} + \frac{\binom{9}{7}\binom{9}{5}}{\binom{18}{12}} \\ &= \frac{108}{221} \\ &\boxed{\approx 0.49}\end{aligned}$$

Question 2

(a)

To find the probability that at least 1 man receives his own hat, we consider the complementary case of no man receiving his own hat, given by D_6. Thus:

$$\begin{aligned}\mathbb{P}(\text{'at least 1 man receives his own hat'}) &= 1 - \frac{D_6}{6!} \\ &= 1 - \frac{6!(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} + \frac{1}{6!})}{6!} \\ &= \boxed{\frac{91}{144}}\end{aligned}$$

(b)

Again we consider the complementary case that no man receives his own hat, or only 1 man receives his own hat. Thus:

$$\begin{aligned}\mathbb{P}(\text{'at least 2 men receives their own hats'}) &= 1 - \underbrace{\frac{D_6}{6!}}_{\substack{\text{no man receives} \\ \text{their own hat}}} - \underbrace{\frac{6 \times D_5}{6!}}_{\substack{\text{1 man receives} \\ \text{his own hat}}} \\ &= \frac{91}{144} - 6 \times \frac{5!(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!})}{6!} \\ &= \boxed{\frac{191}{720}}\end{aligned}$$

(c)

The probability that at least 5 men receive their own hats can be interpreted as every man receiving their own hats, since if 5 men receive their own hats, the last man will automatically also receive his own hat. There is only 1 way for all 6 men to receive their own hats, so we have:

$$\begin{aligned}\mathbb{P}(\text{'at least 5 men receives their own hats'}) &= \frac{1}{6!} \\ &= \boxed{\frac{1}{720}}\end{aligned}$$

Question 3

(a)

Since X is the number of coin tosses upon completion, we model X as a geometric distribution with the "success" being obtaining a `T`: $X \sim \text{geometric}(\frac{1}{2})$, $X \in \{1, 2, \ldots, n\}$. For $x \in \{1, 2, \ldots, n-1\}$, we thus have: $\mathbb{P}(X = x) = (1 - \frac{1}{2})^{x-1}(\frac{1}{2}) = (\frac{1}{2})^x$. As for $x = n$, we take the complement of all the other probabilities (since they sum to 1):

$$\begin{aligned}
\mathbb{P}(X = n) &= 1 - \left(\sum_{i=1}^{n-1} \mathbb{P}(X = i)\right) \\
&= 1 - \left(\sum_{i=1}^{n-1} \left(\frac{1}{2}\right)^i\right) \\
&= 1 - \frac{\frac{1}{2}(1 - (\frac{1}{2})^{n-1-1+1})}{1 - \frac{1}{2}} \\
&= \left(\frac{1}{2}\right)^{n-1}
\end{aligned}$$

The p.m.f. table is shown below:

x	1	2	3	...	n
$\mathbb{P}(X = x)$	$\frac{1}{2}$	$(\frac{1}{2})(\frac{1}{2})$	$(\frac{1}{2})(\frac{1}{2})(\frac{1}{2})$	...	$(\frac{1}{2})^{n-1}$

and in piecewise notation:

$$\boxed{\mathbb{P}(X = x) = \begin{cases} \left(\frac{1}{2}\right)^x, & x \in \{1, 2, \ldots, n-1\} \\ \left(\frac{1}{2}\right)^{x-1}, & x = n \\ 0, & \text{otherwise} \end{cases}}$$

(b)

To find $\mathbb{E}(X)$, we sum over both cases of the p.m.f:

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{i=1}^{n-1} i \cdot \mathbb{P}(X = i) + n \cdot \left(\frac{1}{2}\right)^{n-1} \qquad (1) \\
&= \sum_{i=1}^{n-1} i \cdot \left(\frac{1}{2}\right)^{i} + n \cdot \left(\frac{1}{2}\right)^{n-1} \\
&= \frac{1}{2} \sum_{i=1}^{n-1} i \cdot \left(\frac{1}{2}\right)^{i-1} + n \cdot \left(\frac{1}{2}\right)^{n-1}
\end{aligned}$$

Using the hint, we differentiate the finite geometric series from the fact that $\frac{d}{dx} x^a = a x^{a-1}$. Letting $r = \frac{1}{2}$:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}r}\left[\sum_{i=0}^{n-1} r^i\right] &= \frac{\mathrm{d}}{\mathrm{d}r}\left[\frac{1 - r^n}{1 - r}\right] \\
\sum_{i=1}^{n-1} i \cdot r^{i-1} &= \frac{1 - r^n}{(1 - r)^2} - \frac{n r^{n-1}}{1 - r}
\end{aligned}$$

Now, substituting back into Equation (1), we have:

$$\begin{aligned}
\mathbb{E}(X) &= \frac{1}{2}\left[\frac{1 - (\frac{1}{2})^n}{(1 - \frac{1}{2})^2} - \frac{n (\frac{1}{2})^{n-1}}{1 - \frac{1}{2}}\right] + n \cdot \left(\frac{1}{2}\right)^{n-1} \\
&= \frac{1}{2}\left[4\left(1 - \left(\frac{1}{2}\right)^n\right) - 2n\left(\frac{1}{2}\right)^{n-1}\right] + n \cdot \left(\frac{1}{2}\right)^{n-1} \\
&= 2 - \left(\frac{1}{2}\right)^{n-1} - n \cdot \left(\frac{1}{2}\right)^{n-1} + n \cdot \left(\frac{1}{2}\right)^{n-1} \\
&= \boxed{2 - \left(\frac{1}{2}\right)^{n-1}}
\end{aligned}$$

Question 4

Chebyshev's inequality in the standard form gives the **upper bound** for the probability of being k standard deviations ($\sigma = 5$) away from the mean $\mu = 50$.

$$\mathbb{P}(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2} \tag{2}$$

Let X be the weekly production of the factory (in tons). We want to find the probability $\mathbb{P}(41 \leq X \leq 61)$. Since the bounds are not symmetric about μ, we need the tighter bound $\mathbb{P}(41 \leq X \leq 59)$. Then, the value of k is now $k = \frac{9}{5}$ standard deviations away from the mean.

From Figure 1, we can see that 9 is indeed the "tighter bound":

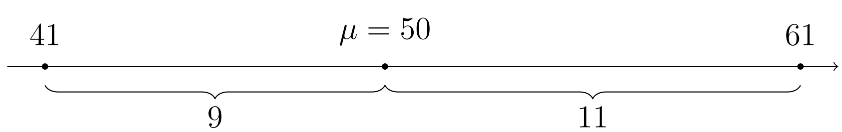

Figure 1: Illustration of bounds

From Equation (2), to find a *lower* bound, we need to take the complement:

$$\mathbb{P}(|X - \mu| < k\sigma) \geq 1 - \frac{1}{k^2} \tag{3}$$

Thus the lower bound for the probability that the weekly production is between 41 and 61 tons is:

$$\begin{aligned}
\mathbb{P}(41 \leq X \leq 61) &\geq \mathbb{P}\left(|X - \mu| < \frac{9}{5}\sigma\right) \\
&\geq 1 - \frac{1}{\left(\frac{9}{5}\right)^2} \\
&= \boxed{\frac{56}{81}}
\end{aligned}$$

Question 5

Let $X \sim \text{negbin}(n, p)$. In the recitation, to find the MGF of X, since the negative binomial distribution is the sum of n independent geometric distributions, let $Y \sim \text{Geometric}(p)$, and $q = 1 - p$.

$$M_Y(t) = \frac{pe^t}{1 - qe^t} \tag{4}$$

We have $Y_1 + Y_2 + \cdots + Y_n \sim \text{negbin}(n, p)$. Since each Y_i's are independent, by the sum property, the MGF of X is given by:

$$\begin{aligned}
M_{Y_1+Y_2+\cdots+Y_n}(t) &= M_{Y_1}(t) \cdot M_{Y_2}(t) \ldots M_{Y_n}(t) \\
&= [M_Y(t)]^n \\
&= M_X(t) \\
&= \left[\frac{pe^t}{1 - qe^t}\right]^n
\end{aligned}$$

Proof. To find $\mathbb{E}(X)$, we need to find the first moment, which requires the derivative:

$$\begin{aligned}
M_X(t) &= \left[\frac{pe^t}{1 - qe^t}\right]^n \\
M_X'(t) &= n\left[\frac{pe^t}{1 - qe^t}\right]^{n-1} \cdot \frac{\mathrm{d}}{\mathrm{dt}}\left[\frac{pe^t}{1 - qe^t}\right] \\
&= n\left[\frac{pe^t}{1 - qe^t}\right]^{n-1} \cdot \left[\frac{pe^t(qe^t + (1 - qe^t))}{(1 - qe^t)^2}\right] \\
&= n\left[\frac{pe^t}{1 - qe^t}\right]^{n-1} \cdot \left[\frac{pe^t}{(1 - qe^t)^2}\right] \\
&= \left[\frac{pe^t}{1 - qe^t}\right] \cdot \frac{n}{(1 - qe^t)}
\end{aligned}$$

Now, to find $\mathbb{E}(X)$, we set $t = 0$:

$$\begin{aligned}
M_X'(0) &= \left[\frac{p}{1 - q}\right]^n \cdot \frac{n}{1 - q} \\
&= \left[\frac{p}{1 - (1 - p)}\right]^n \cdot \frac{n}{1 - (1 - p)} \\
&= \boxed{\frac{n}{p}} = \mathbb{E}(X)
\end{aligned}$$

$\square$

Question 6

(a)

To find the MGF of X, since X is continuous, we need to integrate over all X, taking note that the MGF is only defined on $-1 < t < 1$:

$$
\begin{aligned}
M_X(t) &= \int_{-\infty}^{\infty} e^{tx} f(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty} e^{tx}\frac{1}{2}e^{-|x|}\,\mathrm{d}x \\
&= \int_{-\infty}^{0} e^{tx}\cdot\frac{1}{2}e^{x}\,\mathrm{d}x + \int_{0}^{\infty} e^{tx}\cdot\frac{1}{2}e^{-x} \\
&= \frac{1}{2}\left[\int_{-\infty}^{0} e^{(t+1)x}\,\mathrm{d}x + \int_{0}^{\infty} e^{(t-1)x}\,\mathrm{d}x\right] \\
&= \frac{1}{2}\left[\frac{1}{t+1}e^{(t+1)x}\Bigg|_{-\infty}^{0} + \frac{1}{t-1}e^{(t-1)x}\Bigg|_{0}^{\infty}\right] \\
&= \frac{1}{2}\left[\frac{1}{t+1} - \frac{1}{t-1}\right] \qquad \because (-1 < t < 1) \\
&= \frac{1}{2}\left[\frac{t-1-t-1}{(t+1)(t-1)}\right] \\
&= \frac{1}{2}\left[\frac{-2}{t^2-1}\right] \\
&= \boxed{\frac{1}{1-t^2}}
\end{aligned}
$$

(b)

Note that $\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$, so we need to find the first and second moments. For the first moment we need the derivative:

$$
\begin{aligned}
\mathbb{E}(X) &= M_X'(0) \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{1-t^2}\right)\Bigg|_{t=0} \\
&= -(1-t^2)^{-2}(-2t)\Big|_{t=0} \\
&= \frac{2t}{(1-t^2)^2}\Bigg|_{t=0} \\
&= 0
\end{aligned}
$$

For the second moment:

$$\begin{aligned}
\mathbb{E}(X^2) &= M_X''(0) \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{2t}{(1-t^2)^2}\right)\bigg|_{t=0} \\
&= 2t(-2(1-t^2)^{-3}(-2t)) + (1-t^2)^{-2}(2)\Big|_{t=0} \\
&= \left(\frac{8t^3}{(1-t^2)^3}\right) + \frac{2}{(1-t^2)^2}\bigg|_{t=0} \\
&= 2
\end{aligned}$$

Thus, we have:

$$\begin{aligned}
\mathrm{Var}(X) &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 \\
&= 2 - 0^2 \\
&= \boxed{2}
\end{aligned}$$

Question 7

(a)

Manipulating the right tail probability of a standard normal r.v. Z, we have:

$$\mathbb{P}(Z \geq t) = \mathbb{P}(tZ \geq t^2) = \mathbb{P}(e^{tZ} \geq e^{t^2}) \tag{5}$$

Markov's Inequality gives an upper bound on the tail probability, subject to any constant $a > 0$:

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a} \tag{6}$$

Proof. Applying Equation (6) on Equation (5), we have $a = e^{tZ}$.

$$\begin{aligned}
\mathbb{P}(e^{tZ} \geq e^{t^2}) &\leq \frac{\mathbb{E}(e^{tZ})}{e^{t^2}} \\
&= \frac{M_Z(t)}{e^{t^2}} \\
&= \frac{e^{\frac{t^2}{2}}}{e^{t^2}} \\
&= e^{-\frac{t^2}{2}}
\end{aligned}$$

Where $M_Z(t) = e^{t^2}$ was obtained from Week 2 Class 2 Example 3 (too lazy to typeset). From here, we have:

$$\boxed{\mathbb{P}(Z \geq t) = \mathbb{P}(e^{tZ} \geq e^{t^2}) \leq e^{-\frac{t^2}{2}}} \tag{7}$$

□

(b)

From Equation (7), we substitute $t = -7$, and the fact that the standard normal Z is symmetric about $\mu = 0$, to get:

$$\begin{aligned}
\mathbb{P}(Z \leq -7) &= \mathbb{P}(Z \geq 7) \\
&\leq e^{-\frac{7^2}{2}} \\
&\boxed{\leq e^{-\frac{49}{2}}}
\end{aligned}$$