Statistics 134 - Instructor: Adam Lucas

MIDTERM

Friday, October 5, 2018

SOLUTIONS

Exam Information and Instructions:

- You will have 45 minutes to take this exam. Closed book/notes/etc. No calculator or computer.
- We will be using Gradescope to grade this exam. Write any work you want graded on the front of each page, in the space below each question. Additionally, write your SID number in the top right corner on every page.
- Please use a dark pencil (mechanical or #2), and bring an eraser. *If you use a pen and make mistakes, you might run out of space to write in your answer.*
- Provide calculations or brief reasoning in every answer.
- Unless stated otherwise, you may leave answers as unsimplified numerical and algebraic expressions, and in terms of the Normal c.d.f. Φ. Finite sums are fine, but simplify any infinite sums.
- Do your own unaided work. Answer the questions on your own. The students around you have different exams.

I certify that all materials in the enclosed exam are my own original work.

Sign your name: ________________________________

GOOD LUCK!

SID: ____________________

1. (5 pts) Suppose that on average, 2 moths per 12-hour night are killed by a particular hanging bug zapper. Assume that conditions are the same across different nights and different times of the night, and that moths arrive independently of one another. Find the chance that more than 7 moths are killed in a period of three nights.

Under these assumptions, the number of moths killed in a certain amount of time follows a Poisson Scatter. Let X represent the number of moths killed over three nights. We are told that in one night, 2 moths on average appear, so over three nights, 6 moths on average should appear; thus $X \sim \text{Pois}(6)$. And so,

$$\begin{aligned} P(X > 7) &= 1 - P(X \leq 7) \\ &= 1 - \sum_{k=0}^{7} P(X = k) \\ &= 1 - \sum_{k=0}^{7} \frac{e^{-6} 6^k}{k!} \end{aligned}$$

2. (5 pts) You are given a fair coin and a coin which lands heads with probability $\frac{1}{3}$. Unsure which coin is which, you select one of the coins and decided to toss it until you observe 4 heads; this takes 10 tosses. Given this information, what is the chance you selected the fair coin?

Let A be the event that the fair coin is chosen, and X represent the number of tosses needed to get 4 heads. We want to find $P(A|X = 10)$. To do this, we use Bayes' Rule and condition on whether we selected the fair coin or not. If we selected the fair coin, $X \sim \text{NegBin}\left(4, \frac{1}{2}\right)$ on the range $\{4, 5, 6, \ldots\}$, whereas if the biased coin was chosen the p parameter is instead $\frac{1}{3}$.

$$\begin{aligned} P(A|X = 10) &= \frac{P(A)P(X = 10|A)}{P(X = 10)} \\ &= \frac{\frac{1}{2}\left(\binom{9}{3}\left(\frac{1}{2}\right)^{10}\right)}{\frac{1}{2}\left(\binom{9}{3}\left(\frac{1}{2}\right)^{10}\right) + \frac{1}{2}\left(\binom{9}{3}\left(\frac{1}{3}\right)^{4}\left(\frac{2}{3}\right)^{6}\right)} \end{aligned}$$

An interesting side note is that the $\binom{9}{3}$ term counting for where the successes are allowed to occur disappears when simplifying this fraction. So all that matters is the total number of heads and tails, not where they occurred! This is not immediately obvious however, and would require justification.

SID: ____________________

3. (10 pts total) There are 60 marbles in a bag, of which there are 10 each of the colors red, orange, yellow, green, blue, and violet. Let X denote the number of different colors appearing among 5 marbles selected at random from the bag. Find:

   (a) (3 pts) $P(X = 2)$;

   We consider the probability that two specified colors appear (and no other colors), and multiply by the amount of possible pairs of colors.

$$\begin{aligned} P(X=2) &= \binom{6}{2} P(\text{red and blue both appear, no other colors}) \\ &= \binom{6}{2} \left(P(\text{all red or blue}) - P(\text{all blue}) - P(\text{all red})\right) \\ &= \binom{6}{2} \left(\frac{\binom{20}{5}}{\binom{60}{5}} - 2\left(\frac{\binom{10}{5}}{\binom{60}{5}}\right)\right) \end{aligned}$$

   (b) (3 pts) $\mathbb{E}(X)$;

   Here we use indicators. Let $\mathbb{I}_j$ be the indicator that color j appears among the 5 marbles in our sample. Note that these indicators are identical but not independent. So for all j we have

$$\begin{aligned} P(\mathbb{I}_j = 1) &= 1 - P(\text{none of color } j) \\ &= 1 - \frac{\binom{50}{5}}{\binom{60}{5}} \end{aligned}$$

   And so, by linearity of expectations,

$$\begin{aligned} \mathbb{E}(X) &= \mathbb{E}(\mathbb{I}_1 + \mathbb{I}_2 + \ldots + \mathbb{I}_6) \\ &= 6\mathbb{E}(\mathbb{I}_1) = 6\left(1 - \frac{\binom{50}{5}}{\binom{60}{5}}\right) \end{aligned}$$

(c) (4 pts) $Var(X)$.

Again we proceed using indicators. We know $Var(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$; the second term we already have. We focus on the computation of $\mathbb{E}(X^2)$:

$$\begin{aligned}
\mathbb{E}(X^2) &= \mathbb{E}\left(\left(\sum_{j=1}^{6} \mathbb{I}_j\right)^2\right) \\
&= \mathbb{E}\left(\sum_{j=1}^{6} \mathbb{I}_j^2 + \sum_{j \neq k} \mathbb{I}_j \mathbb{I}_k\right) \\
&= 6 \cdot \mathbb{E}(\mathbb{I}_1) + 6 \cdot 5 \cdot \mathbb{E}(\mathbb{I}_1 \mathbb{I}_2)
\end{aligned}$$

Here we note that $\mathbb{I}_1\mathbb{I}_2 = 1$ when both colors appear in the sample. The calculation of this probability is as follows (assuming color 1 is red and color 2 is blue for simplicity):

$$\begin{aligned}
P(\mathbb{I}_1\mathbb{I}_2 = 1) &= P(\text{red appears} \cap \text{blue appears}) \\
&= 1 - P(\text{red doesn't appear} \cup \text{blue doesn't appear}) \quad \text{(by DeMorgan's Law)} \\
&= 1 - (P(\text{no red}) + P(\text{no blue}) - P(\text{no red \& no blue})) \\
&= 1 - \left(\frac{\binom{50}{5}}{\binom{60}{5}} \cdot 2 - \frac{\binom{40}{5}}{\binom{60}{5}}\right)
\end{aligned}$$

Putting it all together, we have the final answer:

$$Var(X) = \mathbb{E}(X) + 30\left(1 - \left(\frac{\binom{50}{5}}{\binom{60}{5}} \cdot 2 - \frac{\binom{40}{5}}{\binom{60}{5}}\right)\right) - \mathbb{E}(X)^2$$

4. (5 pts) Suppose that bundles of yarn are 60 meters long on average, with an SD of 5 meters, and that bundles are independent of one another. In terms of n, find an upper bound (less than 1) on the probability that the total length of n bundles is less than 200 meters, for $n \geq 4$.

Let X_n denote the total length of n bundles of yarn. We start by observing that $\mathbb{E}(X_n) = 60n$, and $Var(X_n) = n \cdot Var(X_1) = 25n$.

Note that for $n \geq 4$, $\mathbb{E}(X_n) \geq 240$, so we are trying to bound a left tail probability. Markov's Inequality only works for right tail probabilities, so we have no choice but to use Chebyshev's. We proceed as follows, first standardizing the random variable and then manipulating it into the form required for Chebyshev's Inequality:

$$\begin{aligned}
P(X_n < 200) &= P(X_n - 60n < 200 - 60n) \\
&= P\left(\frac{X_n - 60n}{\sqrt{25n}} \leq \frac{200 - 60n}{\sqrt{25n}}\right) \\
&\leq P\left(\frac{X_n - 60n}{\sqrt{25n}} \leq \frac{200 - 60n}{\sqrt{25n}} \text{ or } \frac{X_n - 60n}{\sqrt{25n}} \geq \frac{60n - 200}{\sqrt{25n}}\right) \\
&= P\left(\left|\frac{X_n - 60n}{\sqrt{25n}}\right| \geq \frac{60n - 200}{\sqrt{25n}}\right) \\
&\leq \frac{1}{\left(\frac{60n-200}{\sqrt{25n}}\right)^2}
\end{aligned}$$

SID: ______________________

5. (5 pts) Three couples attend a dinner. Each of the six people chooses a seat randomly from a round table with six seats. What is the probability that no couple sits together? (Hint: use the inclusion-exclusion rule.)

This is a reworked version of the example with Democrats, Republicans, and Independents from Lecture #8. Let A_i be the event that couple i sits together. We observe that

$$\begin{aligned} P(\text{no couple sits together}) &= P(A_1^c \cap A_2^c \cap A_3^c) \\ &= 1 - P(A_1 \cup A_2 \cup A_3) \\ &= 1 - \left(3P(A_1) - \binom{3}{2} P(A_1A_2) + P(A_1A_2A_3)\right), \end{aligned}$$

where

$$\begin{aligned} P(A_1) &= \frac{6 \cdot 2! \cdot 4!}{6!} \\ P(A_1A_2) &= \frac{6 \cdot 2! \cdot 3 \cdot 2! \cdot 2!}{6!} \\ P(A_1A_2A_3) &= \frac{6 \cdot 2! \cdot 2 \cdot 2! \cdot 2!}{6!} \end{aligned}$$