Problem Sheet 11 [1]

Based on Chapters 9.5 and 10.1-10.3 the course book.

Optional Revision Problems

Exercise 1. While Fred is sleeping one night, X legitimate emails and Y spam emails are sent to him. Suppose that X and Y are independent, with $X \sim Pois(10)$ and $Y \sim Pois(40)$. When he wakes up, he observes that he has 30 new emails in his inbox. Given this information, what is the expected value of how many new legitimate emails he has?

Hint: You might find Theorem 4.8.1 and 4.8.2 from the course book useful.

Exercise 2. Let $X_1, X_2, \ldots$ be i.i.d. r.v.s with mean 0, and let $S_n = X_1 + \cdots + X_n$. As shown in Example 9.3.6 in the book, the expected value of the first term given the sum of the first n terms is

$$E(X_1|S_n) = \frac{S_n}{n}$$

Generalize this result by finding $E(S_k|S_n)$ for all positive integers k and n.

Note: There is no restrictions which one of k and n is the greater number, so you should consider all cases.

Week 11 exercises

Exercise 3. Joe will read $N \sim Pois(\lambda)$ books next year. Each book has a $G \sim Pois(\mu)$ number of pages, with book lengths independent of each other and independent of N. Find the variance of the number of book pages Joe will read next year.

Hint: Remember from last week that $E(T|N) = N\mu$

Exercise 4. Let X_1, X_2, and Y be random variables, such that Y has finite variance. Let

$$A = E(Y|X_1) \text{ and } B = E(Y|X_1, X_2).$$

Show that

$$Var(A) \leq Var(B)$$

Also, check that this make sense in the extreme cases where Y is independent of X_1 and where $Y = h(X_2)$ for some function h.

Hint: Use Eve's law on B, with conditioning on X_1 (Theorem 9.3.8).

[1] Exercises are based on the coursebook Statistics 110: Probability by Joe Blitzstein

Exercise 5. One of two identical-looking coins is picked from a hat randomly, where one coin has probability p_1 of Heads and the other has probability p_2 of Heads. Let X be the number of Heads after flipping the chosen coin n times. Find the mean and variance of X.

Exercise 6. Let X and Y be i.i.d. positive r.v.s, and let $c > 0$. For each part below, fill in the appropriate equality or inequality symbol: write $=$ if the two sides are always equal, $\leq$ if the left-hand side is less than or equal to the right-hand side (but they are not necessarily equal), and similarly for $\geq$. If no relation holds in general, write ?.

1. $E(\ln(X))$ ____ $\ln(E(X))$
2. $E(X)$ ____ $\sqrt{E(X^2)}$
3. $P(X > c)$ ____ $\frac{E(X^3)}{c^3}$
4. $P(X \leq Y)$ ____ $P(X \geq Y)$
5. $E(XY)$ ____ $\sqrt{E(X^2)E(Y^2)}$
6. $P(X + Y > 10)$ ____ $P(X > 5 \text{ or } Y > 5)$
7. $E(\min(X, Y))$ ____ $\min(E(X), E(Y))$
8. $E(X/Y)$ ____ $E(X)/E(Y)$
9. $E(X^2(X^2 + 1))$ ____ $E(X^2(Y^2 + 1))$
10. $E\left(\frac{X^3}{X^3+Y^3}\right)$ ____ $E\left(\frac{Y^3}{X^3+Y^3}\right)$

Exercise 7. For i.i.d. r.v.s $X_1, \ldots, X_n$ with mean μ and variance σ^2, give a value of n (as a specific number) that will ensure that there is at least a 99% chance that the sample mean, defined as $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ will be within 2 standard deviations of the true mean μ.

Exercise 8. In a national survey, a random sample of people are chosen and asked whether they support a certain policy. Assume that everyone in the population is equally likely to be surveyed at each step, and that the sampling is with replacement (sampling without replacement is typically more realistic, but with replacement will be a good approximation if the sample size is small compared to the population size). Let n be the sample size, and let $\hat{p}$ and p be the proportion of people who support the policy in the sample and in the entire population, respectively. Show that for every $c > 0$,

$$P(|\hat{p} - p| > c) \leq \frac{1}{4nc^2}.$$