APMA 1650

Homework 5

Instructions: Homework is due by 11:59pm EST in Gradescope on the day listed on the course webpage. You can use calculators for this assignment. Solutions must be written independently and *cannot* be shared with any other students.

You must *show all work* and *explain your answers thoroughly* to get full credit. **You will be graded partly on how well you explain the answers**.

There will be a 5pt penalty for homework submitted with problems incorrectly assigned to a page. A 10pt penalty will be applied for homework submitted during the late window.

Problem 1. (18 pts) You own a moderately successful small-town grocery store. You have collected enough data to determine that the weekly demand Y for potatoes (measured in hundreds of pounds of potatoes purchased) has a PDF

$$f_Y(y) = \begin{cases} 3y^2 & 0 \le y \le 1, \\ 0 & \text{otherwise} \end{cases}$$

Here we are assuming that you can't stock more than 100 pounds of potatoes at once. At the beginning of each week, you purchase up to 100 pounds of potatoes at 75 cents per pound and proceed to sell them at 100 cents per pound. At the end of the week you donate any remaining potatoes to a local food bank.

a. (4 pts) Suppose that you buy a fixed number $a \in [0, 1]$ of hundreds of pounds of potatoes a week. What is your weekly profit in dollars as a function of a and the demand Y? (Keep in mind, if the demand exceeds a, you run out of stock)

b. (8 pts) What are your expected profits in a given week? This will be a function of a. Simplify your answer. (Hint: divide your integral into $0 \le y \le a$ and $a \le y \le 1$)

c. (6 pts) How many hundreds of pounds of potatoes should you buy every week to maximize your expected profits? Give an exact answer. What are the expected profits in this case?

Problem 2. (26 pts) Let X be a continuous random variable with range $[a, \infty)$, for some $a > 0$ and PDF given by

$$f_X(x) = \begin{cases} cx^{-p-1} & x \ge a \\ 0 & \text{otherwise} \end{cases}$$

for some $p > 0$.

a. (4 pts) Find the value of c that makes this a valid PDF.

b. (4 pts) What is the CDF of X?

c. (5 pts) Find the expectation of X. (You may get very different answers depending on whether $p \leq 1$ or $p > 1$)

d. (5 pts) Find $P(X > 4a | X > 3a)$.

e. (8 pts) What is distribution of the random variable $Y = \ln(X/a)$ (i.e. range and PDF)? Do you recognize this distribution? Interpret you answer to part d in terms of this.

Problem 3. (15 pts) Let $U \sim \text{Uniform}([1, 2])$ and let X be the largest root of the quadratic equation

$$x^2 - 2Ux + 1 = 0.$$

a. (5 pts) Give a formula for X in terms of U. What is the range of X?

b. (10 pts) What are the CDF and PDF of X? (Hint: use the fact that for $x \geq 1$, the event $\{X \leq x\}$ is the same as the event $\{x^2 - 2Ux + 1 \geq 0\}$)

Problem 4. (15 pts) The Normal distribution is commonly used in practice to approximate the Binomial distribution for large n. In fact, it also works pretty well for fairly "small" values of n as long as p isn't too close to 0 or 1. Lets see how well it does for $n = 25$ and $p = .6$. (You may use a calculator and CDF calculator in the book)

a. (2 pts) Suppose that $X \sim \text{Binomial}(n, p)$, n and p as above. Compute the probability $P_X(7)$.

b. (5 pts) Suppose that $Y \sim N(\mu, \sigma^2)$, with the same mean and variance as X, $\mu = np$ and $\sigma^2 = npq$. Compute $P(6.5 \leq Y \leq 7.5)$. Compare your answer with part a.

c. (4 pts) The general principle is that the normal approximation is good as long as the values $p \pm 3\sqrt{\frac{pq}{n}}$ are between 0 and 1. Show that this is the case if and only if

$$n > \frac{9p}{q}, \quad \text{and} \quad n > \frac{9q}{p}$$

or in other words if

$$n > 9\left(\frac{\max\{p, q\}}{\min\{p, q\}}\right).$$

d. (4 pts) How large should n be taken to approximate the Binomial distribution by a Normal for $p = .5, .8, .99, .999$?

Problem 5 (26 pts) Our book doesn't have a CDF calculator for the Gamma distribution (bummer!). However, we have seen that the Gamma distribution is closely related to the Poisson distribution. In fact, it is possible to calculate all the probabilities for the Gamma distribution using only the Poisson distribution. Let $X \sim \text{Gamma}(n, \lambda)$ for $\lambda > 0$ and $n \in \mathbb{N}$ and let $Y \sim \text{Poisson}(\lambda)$.

a. (8 pts) Use the following formula

$$\frac{1}{\Gamma(n)} \int_{\lambda}^{\infty} x^{n-1} e^{-x} \, \mathrm{d}x = \sum_{k=0}^{n-1} \frac{\lambda^k e^{-\lambda}}{k!}$$

to show that

$$P(X > 1) = P(Y \leq n - 1).$$

b. (5 pts) What does the result of part a give you for $P(X > 1)$ when $\lambda = 1$ and $n = 5$? (Use the Poisson CDF calculator)

c. (5 pts) Recall that the Poisson distribution can be interpreted as the number of occurrences of a certain event that occurs in the time window $[0, 1]$, where the times between each event are independent and Exponential(λ) distributed. We also know that the gamma distribution is a sum of n independent Exponential(λ) RVs. Using words, explain why $P(X > 1) = P(Y \leq n - 1)$ makes sense.

d. (8 pts) Now for each $t > 0$, let $Y^{(t)} \sim \text{Poisson}(\lambda t)$ (the number of occurrences of events in the time window $[0, t]$). Show that

$$P(X > t) = P(Y^{(t)} \leq n - 1).$$