University of Illinois Spring 2020

ECE 313: Midterm Exam II (Conflict)

Wednesday, April 09, 2020

The exam consists of **6** problems worth a total of 100 points. The problems are not weighted equally, so it is best for you to pace yourself accordingly. SHOW YOUR WORK. Answers without appropriate justification will receive very little credit. Reduce common fractions to lowest terms, but DO NOT convert them to decimal fractions (for example, write $\frac{3}{4}$ instead of $\frac{24}{32}$ or 0.75).

The solutions must be handwritten on pieces of paper or electronic notebook. Please start a new page for each sub-problem. Please scan your solutions and upload them onto Gradescope before the deadline.

1. [**15 points**] Suppose that X is a random variable with CDF

$$F_X(x) = \begin{cases} 0 & x < a \\ (x-1)^2 - 1 & a \leq x < b \\ 1 & x \geq b \end{cases} . \tag{1}$$

Assume $a < b$.

(a) [**7 points**] Find the smallest number a and largest number b such that $F_X(x)$ is a valid CDF.
Solution: A valid CDF is non-negative, monotonically increasing, and satisfying $F_X(+\infty) = 1$. To satisfy these properties, the smallest value of a will be when $0 = (a-1)^2 - 1$, which is $a = 2$; the largest value of b will be when $1 = (b-1)^2 - 1$, which is $b = 1 + \sqrt{2}$.

(b) [**8 points**] Suppose $a = 2.1$ and $b = 2.2$. What is $P\{X < 2.1\}$? What is $P\{X \leq 2.1\}$? $P\{X < 2.2\}$? What is $P\{X \leq 2.2\}$?
Solution: If $a = 2.1$ we have a discontinuity at $x = 2.1$ since $(2.1-1)^2 - 1 = .21 \neq 0$. So $P\{X < 2.1\} = 0$ and $P\{X \leq 2.1\} = .21$.
If $b = 2.2$ we will have a discontinuity at $x = 2.2$ since $(2.2-1)^2 - 1 = .44 \neq 1$. So $P\{X < 2.2\} = .44$ and $P\{X \leq 2.2\} = 1$.

2. [**20 points**] Suppose that the number of pitches thrown in a baseball game are described by a Poisson process of rate r pitches per minute. Let N_t denote the number of pitches thrown by time t and U_j the time delay between pitch $j-1$ and j.

(a) [**6 points**] A manager plans for a pitcher to enter the game and exit after he throws 50 pitches. Let T be the random variable for the total time the pitcher is in the game. Express $E[T]$ as a function of r.
Solution: Since the time delay variables are independent, we have

$$E[T] = E\left[\sum_{i=1}^{50} U_i\right] = 50E[U_i] = \frac{50}{r}. \tag{2}$$

(b) [**7 points**] What is the pdf for the time when the fourth pitch is thrown?
Solution: For a fixed time t,

$$P\{N_t \leq 3\} = \sum_{k=0}^{3} P\{N_t = k\} = e^{-tr} \sum_{k=0}^{3} \frac{(tr)^k}{k!}.$$

Since $P\{N_t \leq 3\}$ is the probability that the fourth pitch comes after time t, we have $1 - P\{N_t \leq 3\}$ is the probability that the fourth pitch comes at time t or sooner; hence it is the cdf of the arrival time of the fourth pitch. Thus, the pdf is

$$-\frac{d}{dt}P\{N_t \leq 3\} = re^{-tr}\sum_{k=0}^{3}\frac{(tr)^k}{k!} - e^{-tr}\sum_{k=0}^{3}k\frac{t^{k-1}r^k}{k!} = e^{-tr}\frac{(tr)^3}{3!}. \tag{3}$$

(c) [**7 points**] Each pitch in a baseball game is either a "strike" or a "ball". Let us model each pitch as a Bernoulli trial in which a "strike" is thrown with probabily s and a "ball" is thrown with probability $1 - s$, for some $0 \leq s \leq 1$. If it is known that no more than five pitches are thrown in the first four minutes of the game, what is the probability that exactly three of those pitches are "strikes"? Express your answer as a function or r and s.

Solution: For $k \leq 5$, we have

$$P[N_4 = k | N_4 \leq 5] = \frac{P[N_4 = k]}{P[N_4 \leq 5]} = \frac{\frac{(4r)^k}{k!}e^{-4r}}{\sum_{j=0}^{5}\frac{(4r)^j}{j!}e^{-4r}}.$$

We are interested in the events of throwing 3, 4, and 5 pitches and 3 of them being strikes. This is given by

$$\frac{\sum_{k=3}^{5}\frac{(4r)^k}{k!}e^{-4r}\binom{k}{3}s^3(1-s)^{k-3}}{\sum_{j=0}^{5}\frac{(4r)^j}{j!}e^{-4r}}. \tag{4}$$

3. [**15 points**] Alice plays the same game at a casino 100 times. In the ith game, her earnings are X_i, where $P\{X_i = 1\} = 0.3$ and $P\{X_i = -1\} = 0.7$. The random variables $X_1, X_2, ..., X_{100}$ are all independent. Let $X = \sum_{i=1}^{n} X_i$ be her total earnings (which may be negative, if she loses money).

(a) [**5 points**] Let $Z_i = (X_i + 1)/2$, for $i = 1, ..., 100$ and $Z = \sum_{i=1}^{100} Z_i$. What is the distribution of Z?

Solution: Since $P\{Z_i = 1\} = P\{X_i = 1\} = 0.3$ and $P\{Z_i = 0\} = P\{X_i = -1\} = 0.7$, Z_i is Bernoulli(0.3). Hence $Z = \sum_{i=1}^{100} Z_i$ has a Binomial$(100, 0.3)$ distribution.

(b) [**5 points**] Express the event $\{X = 0\}$ in terms of Z and compute $P\{X = 0\}$.

Solution: In order for X to be 0, Alice must win exactly 50 out of the 100 games. Since Z is the number of games won, $\{X = 0\} = \{Z = 50\}$, and

$$P\{X = 0\} = P\{Z = 50\} = \binom{100}{50}(0.3)^{50}(0.7)^{50}.$$

(c) [**5 points**] Use the Gaussian approximation with continuity correction to compute $P\{|X| \geq 10\}$. Express your answer in terms of the Q function. (Hint: Express the event $\{|X| \geq 10\}$ in terms of Z.)

Solution: Since $X_i = 2Z_i - 1$, we can write

$$X = \sum_{i=1}^{100}(2Z_i - 1) = 2Z - 100.$$

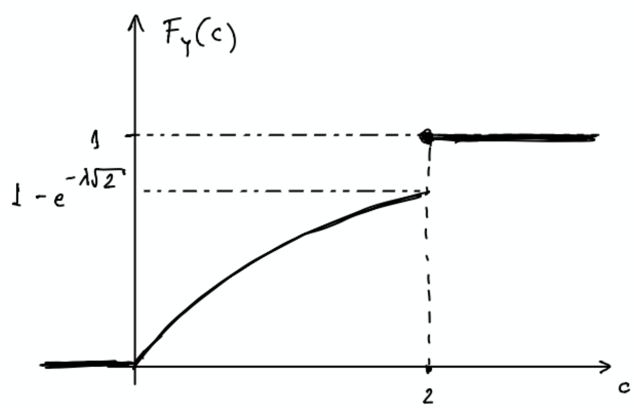
Since Z is Binomial$(100, 0.3)$, $E[Z] = 100 \cdot 0.3 = 30$ and $\text{Var}(Z) = 100(0.3)(0.7) = 21$. Hence, we have

$$\begin{aligned} P\{|X| \geq 10\} &= P\{X \geq 10\} + P\{X \leq -10\} = P\{Z \geq 55\} + P\{Z \leq 45\} \\ &= P\{Z \geq 54.5\} + P\{Z \leq 45.5\} = P\left\{\frac{Z-30}{\sqrt{21}} \geq \frac{24.5}{\sqrt{21}}\right\} + P\left\{\frac{Z-30}{\sqrt{21}} \leq \frac{15.5}{\sqrt{21}}\right\} \\ &\approx Q\left(\frac{24.5}{\sqrt{21}}\right) + \Phi\left(\frac{15.5}{\sqrt{21}}\right) = Q\left(\frac{24.5}{\sqrt{21}}\right) + Q\left(\frac{-15.5}{\sqrt{21}}\right) \end{aligned}$$

4. **[15 points]** Suppose the random variable X has an exponential distribution with parameter λ. Let $Y = \min(X^2, 2)$.

   (a) [**7 points**] Compute and sketch the CDF of Y.
   Solution: Since $X^2 \geq 0$, $Y = \min(X^2, 2) \geq 0$. Moreover, since $\min(x^2, 2) \leq 2$ for any x, $Y \in (0, 2)$. Hence, for $c < 0$, $F_Y(c) = 0$ and for $c \geq 2$, $F_Y(c) = 1$. For $0 \leq c < 2$,

   $$F_Y(c) = P\{\min(X^2, 2) \leq c\} = P\{X^2 \leq c\} = P\{X \leq \sqrt{c}\} = 1 - e^{-\lambda\sqrt{c}}.$$



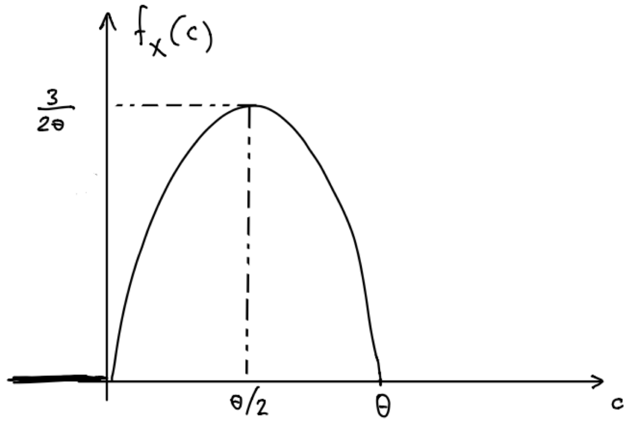
   (b) [**8 points**] Is Y a continuous-type random variable? If so, sketch its pdf, $f_Y(y)$. If not, explain why not.
   Solution: Because the CDF $F_Y(c)$ has a jump at $c = 2$, $P\{Y = 2\} = P\{X^2 \geq 2\} = e^{-\lambda\sqrt{2}} > 0$. Therefore, Y is not a continuous-type random variable and does not have a pdf.

5. **[15 points]** Let X be a continuous-type random variable with pdf given by

$$f_X(x) = \begin{cases} \frac{6}{\theta^3} x(\theta - x) & 0 \leq x \leq \theta \\ 0 & \text{otherwise} \end{cases}$$

   for an unknown parameter $\theta > 0$.

   (a) [**7 points**] Sketch $f_X(x)$ for a general $\theta > 0$. Make sure to label the axes and important points, including maxima/minima, as a function of θ.
   Solution:



(b) [**8 points**] You observe that $X = 10$. Compute the maximum likelihood parameter estimate $\hat{\theta}_{ML}$.

Solution: We need to find the value of θ that maximizes $f_X(u)$. To do that, we can differentiate $f_X(u)$ with respect to θ and set it equal to 0. Clearly, the maximizing value must satisfy $\theta > u$ (otherwise $f_X(u) = 0$). So we can focus on $0 < u < \theta$.

$$\frac{d}{d\theta}\left(\frac{6}{\theta^3}u(\theta - u)\right) = 6\left(-\frac{2u}{\theta^3} + \frac{3u^2}{\theta^4}\right) = 0$$
$$\Leftrightarrow \quad \frac{3u^2}{\theta^4} = \frac{2u}{\theta^3} \Leftrightarrow \theta = \frac{3u}{2}.$$

Since our observation is $u = 10$, we have $\hat{\theta}_{ML} = \frac{3u}{2} = 15$.

6. [**20 points**] A company has developed a video game app and wants to assess whether the app is popular or not. The game downloads follow the Poisson process. They observe how much time it takes to reach 100 downloads after launching it. For popular games, the rate of downloads is $\lambda_1 = 6$ per minute. For unpopular games, the rate of downloads is $\lambda_0 = 2$ per minute. They observe T and would like to determine between H_0: it is an unpopular app; and H_1: it is a popular app.

(a) [**9 points**] Find the ML decision rule.

Solution: The time T it takes to reach $r = 100$ downloads follow the Erlang distribution. With the rate of downloads λ, the likelihood is $f(t) = \frac{\exp(-\lambda t)\lambda^r t^{r-1}}{(r-1)!}$. The likelihood ratio between H_1 and H_0 is

$$\Lambda = \frac{f_1(t)}{f_0(t)} = \frac{\exp(-\lambda_1 t)\lambda_1^r}{\exp(-\lambda_0 t)\lambda_0^r} \tag{5}$$

The ML decision rule is if $\Lambda > 1$, choose H_1, and if $\Lambda < 1$, choose H_0. Take the log likelihood,

$$\log\Lambda = -(\lambda_1 - \lambda_0)t + r\left(\log(\lambda_1) - \log(\lambda_0)\right)\begin{cases} > 0, & \text{Declare } H_1, \\ < 0, & \text{Declare } H_0 \end{cases} \tag{6}$$

Therefore, the ML decision rule is the following:

$$\begin{cases} \text{Declare } H_1, & \text{if } t < \frac{r(\log(\lambda_1) - \log(\lambda_0))}{\lambda_1 - \lambda_0} = 25\log(3) \\ \text{Declare } H_0, & \text{if } t > \frac{r(\log(\lambda_1) - \log(\lambda_0))}{\lambda_1 - \lambda_0} = 25\log(3) \end{cases} \tag{7}$$

(b) [**6 points**] What are the p_{fa} and p_{miss} under the ML rule? You are not required to simplify the integral.

Solution:

$$p_{fa} = \int_0^{25\log(3)} f_0(t)dt = \frac{\lambda_0^r}{(r-1)!}\int_0^{25\log(3)} e^{-\lambda_0 t}t^{r-1}dt \tag{8}$$

$$p_{miss} = \int_{25\log 3}^{\infty} f_1(t)dt = \frac{\lambda_1^r}{(r-1)!}\int_{25\log(3)}^{\infty} e^{-\lambda_1 t}t^{r-1}dt \tag{9}$$

(c) [**5 points**] Suppose we know that this company is very successful and the prior probability π_1 is three times of π_0. Find the MAP decision rule.

Solution: The MAP decision rule is if $\Lambda > \frac{\pi_0}{\pi_1}$, choose H_1, and if $\Lambda < \frac{\pi_0}{\pi_1}$, choose H_0. Take the log likelihood,

$$\log\Lambda = -(\lambda_1-\lambda_0)t + r\left(\log(\lambda_1)-\log(\lambda_0)\right)\begin{cases} > \log\left(\frac{\pi_0}{\pi_1}\right), & \text{Declare } H_1, \\ < \log\left(\frac{\pi_0}{\pi_1}\right), & \text{Declare } H_0 \end{cases} \tag{10}$$

Therefore, the ML decision rule is the following:

$$\begin{cases} \text{Declare } H_1, & \text{if } t < \frac{r(\log(\lambda_1)-\log(\lambda_0))-\log(\pi_0/\pi_1)}{\lambda_1-\lambda_0} = \frac{101\log 3}{4} \\ \text{Declare } H_0, & \text{if } t > \frac{r(\log(\lambda_1)-\log(\lambda_0))-\log(\pi_0/\pi_1)}{\lambda_1-\lambda_0} = \frac{101\log 3}{4} \end{cases} \tag{11}$$