Homework # 1

Elliott Pryor

31 January 2022

Problem 1

Statement

The amount of time (in hours), X, needed by a local repair shop, Rusts 'R Us, to repair a randomly selected piece of equipment is assumed to be an exponential random variable with pdf

$$f_X(x) = \begin{cases} 1/\lambda e^{-x/\lambda} & x > 0; \lambda > 0 \\ 0 & otherwise \end{cases}$$

and mgf $m_X(t) = (1 - \lambda t)^{-1}$.

1. Explain in words what the random variable, X, represents in this scenario. Why is X considered a random variable?
2. Identify the unknown parameter and explain in words what the parameter represents in this scenario
3. In a random sample of 5 repair times, Rusts 'R Us observes the following values (in hours): 14, 6, 3, 4, 1.5. Explain in words what n, $X_1 \dots X_n$, λ, and $x_1, \dots, x_n$ are in this situation, and, if applicable, give the numerical values of each.
4. Find the joint distribution of the random sample in part (3). List all assumptions made when finding this joint distribution and explain whether or not each is reasonable in this scenario.

Solution

1. X is the amount of time Rusts 'R Us takes to repair a piece of equipment. X is a random variable because the amount of time it takes depends on the equipment which is randomly selected.
2. The unknown parameter is λ which is the average amount of time (mean) it takes Rusts 'R Us to repair a piece of equipment.
3. n is the total number of samples, $n = 5$. $X_1 \dots X_n$ is the time it takes to repair each piece of equipment (random variable), X_i = time to repair ith piece of equipment. λ is the parameter of exponential distribution, it is the mean value of the distribution or the mean amount of time it takes to repair; we do not know the real value of λ, so the closest we can do is to estimate based on our 5 observations: $\hat{\lambda} = 5.7$. $x_1, \dots, x_n$ are the observed amount of time it took to repair each piece of equipment, $x_1, x_2, x_3, x_4, x_5 = 14, 6, 3, 4, 1.5$
4. Assume that amount of time it takes to repair each piece of equipment $(X_1, X_2, X_3, X_4, X_5)$ is a random sample from the same exponential distribution (iid) that is given above.

Then the joint distribution is:

$$f_{X_1,X_2,X_3,X_4,X_5}(x_1, x_2, x_3, x_4, x_5) = \prod_{i=1}^{5} f_{X_i}(x_i) = \prod_{i=1}^{5} 1/\lambda e^{-x_i/\lambda} = 1/\lambda^5 e^{-28.5/\lambda} I_{(\lambda,\infty)}(x)$$

Problem 2

Statement

Suppose the length of time, X, it takes a worker at the UPS store to assemble a box has the following pdf :

$$f_X(x) = \begin{cases} e^{-x-\Theta} & x > 0; \Theta > 0 \\ 0 & otherwise \end{cases}$$

Where Θ represents the minimum time until task completion. Let $X_1, \ldots, X_n$ be a random sample of completion times from this distribution.

1. Derive (i.e, go through the steps, don't use the shortcut THM) the CDF of the smallest order statistic, $X_{(1)}$
2. Find the pdf of the smallest order statistic, $X_{(1)}$
3. Give an example of how the pdf of $X_{(1)}$ might be used by a UPS store manager and explain why it would be useful.

Solution

1. We first need to find $F_X(x)$ by integration:

$$F_X(x) = \int_{-\infty}^{x} f_X(y)dy = \int_{\Theta}^{x} e^{-y-\Theta}dy = \left(-e^{-y-\Theta}\right]_{\Theta}^{x} = -e^{-x-\Theta} + e^{-2\Theta}$$

   Then we can compute $F_{X_{(1)}}(x)$:

$$\begin{aligned} F_{X_{(1)}}(x) &= P(X_{(1)} \leq x) = 1 - P(X_{(1)} \geq x) = 1 - P(X_i \geq x \; \forall i) \\ &= 1 - \prod_{i=1}^{n} P(X_i \geq x) = 1 - \prod_{i=1}^{n} 1 - P(X_i \leq x) = 1 - \prod_{i=1}^{n} 1 - F_{X_i}(x) \\ &= 1 - (1 - F_{X_i}(x))^n \\ &= 1 - (1 + e^{-x-\Theta} - e^{-2\Theta})^n \end{aligned}$$

2.

$$f_{X_{(1)}}(x) = \frac{d}{dx}F_{X_{(1)}}(x) = \frac{d}{dx}(1 + e^{-x-\Theta} - e^{-2\Theta})^n) = n(1 + e^{-x-\Theta} - e^{-2\Theta})^{n-1}e^{-x-\Theta}I_{(\Theta,\infty)}(x)$$

3. First we ask what does $X_{(1)}$ physically represent? The minimum. So $X_{(1)}$ is the minimum amount of time it takes any of the (sampled) UPS store employees to assemble a box. Presumably UPS wants to assemble the most boxes possible, and thus wants to reduce the amount of time it takes to assemble a box, so the manager could use this information to compare say different training programs or something. It also could be useful for time estimation, the manager should not expect that workers can go faster than the minimum time and could use $\mathbb{E}[X_{(1)}]$ (which we can get from $f_{X_{(1)}}(x)$) to make more realistic deadlines for their workers.