Nathan Bickel
STAT 509: Section 003
Professor: Dr. Stutz
September 14, 2022

STAT 509 Homework 2

Problem 1 You have the following PDF: $f(x) = \begin{cases} \frac{1}{9} & 1 \leq x \leq c. \end{cases}$
(a) Find the value of c such that this is a valid PDF.
(b) Calculate $\mathbb{E}(X)$.
(c) Calculate $\mathbb{V}(x)$.
(d) Calculate $F(x)$.
(e) Calculate $P(2 < X < 5)$.
(f) Calculate $P(X = 4)$.
(g) Calculate the median of X.

Solution.

(a) We solve for c:

$$\begin{aligned}
\sum_{x=1}^{c} \frac{1}{9} &= 1 && \text{(PDFs must add to 1)} \\
\frac{1}{9}c &= 1 && \text{(constant rule)} \\
c &= 9
\end{aligned}$$

(b) Using the definition of $\mathbb{E}(X)$:

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{x=1}^{9} xf(x) \\
&= \sum_{x=1}^{9} \frac{x}{9} \\
&= \frac{1}{9} \sum_{x=1}^{9} x \\
&= \frac{1}{9} \left(\frac{9(9+1)}{2} \right) = 5.
\end{aligned}$$

(c) Using the definition of $\mathbb{V}(x)$:

$$\begin{aligned}
\mathbb{V}(x) &= \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 \\
&= \sum_{x=1}^{9} \left(x^2 f(x) \right) - 5^2 \\
&= \frac{1}{9} \sum_{x=1}^{9} \left(x^2 \right) - 5^2 \\
&= \frac{1}{9} \left(\frac{9(9+1)(2(9)+1)}{6} \right) - 5^2 = \frac{20}{3}.
\end{aligned}$$

(d) Since the distribution is uniform, for $1 \le x \le 9$ we have

$$F(x) = \frac{x}{9}.$$

(e) By definition, we have

$$P(2 < X < 5) = \sum_{x=3}^{4} \frac{1}{9} = \frac{2}{9}.$$

(f) Since the distribution is uniform, $P(X = 4) = \frac{1}{9}$.

(g) Since the distribution is perfectly symmetrical, the median is equal to the mean. So the median is 5.

Problem 2 Find the CDF of an exponential random variable with mean $\frac{1}{\lambda}$.

Solution.

If the mean is $\frac{1}{\lambda}$, then $f(x) = \lambda e^{-\lambda x}$. Thus,

$$\begin{aligned} F(x) &= \int_0^x \lambda e^{-\lambda t} dt \\ &= -e^{-\lambda t}\Big|_0^x \\ &= -e^{-\lambda x} + e^{-\lambda(0)} \\ &= 1 - e^{-\lambda x}. \end{aligned}$$

Problem 3 Let X be a random variable which represents the lifetime in years of a particular battery. We are given that X has an exponential distribution with rate $\lambda = .15$.
(a) What is the expected value of X?
(b) What is the variance of X?
(c) What is the CDF of X?
(d) What is the probability that the battery fails between the fifth and sixth year? Show by hand **and** provide the appropriate R code.
(e) What is the probability that the battery is still working after three years? Show by hand **and** provide the approriate R code.
(f) What is the probability that the battery is still working after five years, given that the battery is still working after two years?
(g) You observe a battery who's lifetime is in the 99th percentile. How long has this battery lasted? Show by hand **and** provide the appropriate R code.

Solution.

(a) The expected value of X is $\frac{1}{\lambda} = \frac{1}{.15} = \frac{20}{3}$.

(b) Using the definition of $\mathbb{V}(x)$:

$$\begin{aligned} \mathbb{V}(x) &= \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 \\ &= \int_0^\infty x^2(0.15)e^{-(0.15)x} dx - \left(\frac{20}{3}\right)^2 \\ &= \frac{800}{9} - \frac{400}{9} = \frac{400}{9}. \end{aligned}$$

(c) We integrate:

$$\begin{aligned} F(x) &= \int_0^x \frac{3}{20} e^{-3t/20} dt \\ &= -e^{-3t/20}\Big|_0^x \\ &= -e^{-3x/20} + e^0 \\ &= 1 - e^{-3x/20}. \end{aligned}$$

(d) We integrate:

$$\begin{aligned} P(5 < X < 6) &= \int_5^6 \frac{3}{20} e^{-3t/20} dt \\ &= -e^{-3t/20}\Big|_5^6 \\ &= -e^{-3(6)/20} + e^{-3(5)/20} \\ &= e^{-3/4} - e^{-9/10} \approx 0.0658. \end{aligned}$$

```
> pexp(6,0.15)-pexp(5,0.15)
[1] 0.06579689
```

(e) We integrate:

$$\begin{aligned} P(X > 3) &= \int_3^\infty \frac{3}{20} e^{-3t/20} dt \\ &= -e^{-3t/20}\Big|_3^\infty \\ &= -e^{-3(\infty)/20} + e^{-3(3)/20} \\ &= e^{-9/20} \approx 0.638. \end{aligned}$$

```
> 1-pexp(3,0.15)
[1] 0.6376282
```

(f) By the memoryless property, this is equivalent to $P(X > 3)$. So $P(X > 5)|P(X > 2) = P(X > 3) = e^{-9/20} \approx 0.638$.

(g) We integrate:

$$\begin{aligned} P(X \le a) &= \frac{99}{100} \\ \int_0^a \frac{3}{20} e^{-3x/20} dx &= \frac{99}{100} \\ -e^{-3t/20}\Big|_0^a &= \frac{99}{100} \\ -e^{-3a/20} + e^{-3(0)/20} &= \frac{99}{100} \\ 1 - \frac{99}{100} &= e^{-3a/20} \\ \ln \frac{1}{100} &= -3a/20 \\ a &= -\frac{20}{3} \ln \frac{1}{100} \approx 30.7. \end{aligned}$$

```
> qexp(0.99,0.15)
[1] 30.70113
```

Problem 4 The number of earthquakes that occur per week in California follows a Poisson distribution with a mean of 1.5.
(a) What is the probability that an earthquake occurs within the first week? Show by hand **and** provide the appropriate R code.
(b) What is the expected amount of time until an earthquake occurs?
(c) What is the standard deviation of the amount of time until two earthquakes occur?
(d) What is the probability that it takes more than a month to observe 4 earthquakes? Show by hand (you may simply leave it as an integral) **and** provide the appropriate R code.
(e) What is the median amount of time it takes for 5 earthquakes to occur? Show by hand (you may simply leave it as an integral, but be sure to explain how to find the median) **and** provide the appropriate R code.

Solution.

(a) We can model this as a gamma distribution with $\alpha = 1$ and $\lambda = 1.5$. So the probability of the first earthquake occurring within the first week is

$$\begin{aligned}\int_0^1 \frac{\lambda^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\lambda y} dy &= \int_0^1 \frac{3}{2} e^{-3y/2} dy \\ &= -e^{-3y/2}\Big|_0^1 \\ &= -e^{-3/2} + e^0 \\ &= 1 - e^{-3/2} \approx 0.777.\end{aligned}$$

(b) We can again use a gamma distribution and the definition of expected value:

$$\begin{aligned}\int_0^\infty y \frac{\lambda^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\lambda y} dy &= \int_0^\infty \frac{3}{2} y e^{-3y/2} dy \\ &= -ye^{-3y/2}\Big|_0^\infty - \int_0^\infty -e^{-3y/2} dy \\ &= -ye^{-3y/2}\Big|_0^\infty - \frac{2}{3} e^{-3y/2}\Big|_0^\infty \\ &= 0 - 0 + 0 - \left(-\frac{2}{3}\right) = \frac{2}{3}.\end{aligned}$$

So on average, it will take two thirds of a week for an earthquake to happen. This agrees with our formula that $E(Y) = \frac{\alpha}{\lambda} = \frac{1}{3/2}$.

(c) We can model this as a gamma distribution with $\lambda = 1.5$ and $\alpha = 2$. We have that $Var(Y) = \frac{\alpha}{\lambda^2}$, so the standard deviation is $\sqrt{\frac{2}{(3/2)^2}} = \frac{8}{9}$ weeks.

(d) We can model this as a gamma distribution with $\lambda = 1.5$ and $\alpha = 4$. Since a month is roughly 4 weeks, the probability that it takes more than 4 weeks to observe 4 earthquakes is

$$\int_4^\infty \frac{\lambda^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\lambda y} dy = \int_4^\infty \frac{(3/2)^4}{\Gamma(4)} y^{4-1} e^{-3y/2} dy = \int_4^\infty \frac{27}{32} y^3 e^{-3y/2} dy \approx 0.151.$$

```
> 1-pgamma(4,4,1.5)
[1] 0.1512039
```

(e) We can model this as a gamma distribution with $\lambda = 1.5$ and $\alpha = 5$. We call the median time it takes for 5 earthquakes to occur t. So we can solve

$$\int_0^t \frac{\lambda^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\lambda y} dy = \int_0^t \frac{(3/2)^5}{\Gamma(5)} y^{5-1} e^{-3y/2} dy = \int_0^t \frac{81}{256} y^4 e^{-3y/2} dy$$

for t. This gives a value of $t \approx 3.11$ weeks.

```
> qgamma(0.5,5,1.5)
[1] 3.113939
```

Problem 5 The load required to degrade the structural integrity of a foundation is normally distributed with a mean of 22 tons and a standard deviation of 5 tons.
(a) While at the site, a fellow engineer asks you the approximate probability that it takes between 17 and 27 tons to degrade the foundation. Being knowledgeable in statistics, what you are able to tell him off the top of your head?
(b) Verify that your estimate in (a) is accurate using R.
(c) A similar foundation is to be constructed so that the load required to degrade its integrity is again normally distributed with a standard deviation of 5 tons, but the improved foundation is required to only have a 20% chance to degrade under a 22 ton load or less. What must the mean of this distribution be so that the foundation is constructed to the given specifications? Use R.

Solution.

(a) By the 68-95-99.7% rule, the probability that it takes between 17 and 27 tons to degrade the foundation is approximately 0.68 because this is ± 1 standard deviations away from the mean.

(b)

```
> pnorm(27,22,5)-pnorm(17,22,5)
[1] 0.6826895
```

(c)

```
> 22+qnorm(1-0.2,0,5)
[1] 26.20811
> pnorm(22,22+qnorm(1-0.2,0,5),5)
[1] 0.2
```