Homework 5

MSSC 6010- Computational Probability

Henri Medeiros Dos Reis

October 4, 2023

Question 1. (4.4.42) From book. The Laplace distrivution, also known as a double exponential has a **pdf** given by

$$f(x) = \frac{\lambda}{2}e^{-\lambda|x-\mu|}, \text{ where } -\infty < x < \infty, -\infty < \mu < \infty, \lambda > 0$$

- *(a)* Find the theoretical mean and variance of a Laplace distribution. (Hint: Integrals of absolute values should be done as a positive and a negative part, in this case, with limits from $-\infty$ to μ and from μ to ∞.)

First, lets just integrate $f(x)$, since we will also need it later.

$$\begin{aligned}
\int_{-\infty}^{\infty} \frac{\lambda}{2}e^{-\lambda|x-\mu|}dx &= \frac{\lambda}{2}\left(\int_{-\infty}^{\mu} e^{-\lambda(\mu-x)}dx + \int_{\mu}^{\infty} e^{-\lambda(x-\mu)}dx\right) \\
&= \frac{\lambda}{2}\left(\frac{1}{\lambda}e^{-\lambda(\mu-x)}\Bigg|_{-\infty}^{\mu} + \frac{-1}{\lambda}e^{-\lambda(x-\mu)}\Bigg|_{\mu}^{\infty}\right) \\
&= \frac{\lambda}{2}(\frac{1}{\lambda}e^{0} + \frac{1}{\lambda}e^{0}) \\
&= 1
\end{aligned}$$

Now

$$
\begin{aligned}
E[X] &= \int_{-\infty}^{\infty} x\frac{\lambda}{2}e^{-\lambda|x-\mu|}dx \text{ let } y = x-\mu \\
&= \frac{\lambda}{2}\left(\int_{-\infty}^{\infty} ye^{-\lambda|y|}dy + \mu\int_{-\infty}^{\infty} e^{-\lambda|y|}dy\right) \text{ from above, the second integral is 1} \\
&= \frac{\lambda}{2}\left(\int_{-\infty}^{0} ye^{\lambda(y)}dy + \int_{0}^{\infty} ye^{-\lambda(y)}dy\right) + \mu \\
&= \frac{\lambda}{2}\left[\left.\frac{ye^{\lambda y}}{\lambda}\right|_{-\infty}^{0} - \int_{-\infty}^{0}\frac{1}{\lambda}e^{\lambda y}dy + \left.\frac{ye^{-\lambda y}}{-\lambda}\right|_{0}^{\infty} - \int_{0}^{\infty}\frac{1}{-\lambda}e^{-\lambda y}dy\right] + \mu \\
&= \frac{\lambda}{2}\left[-\int_{-\infty}^{0}\frac{1}{\lambda}e^{\lambda y}dy + \int_{0}^{\infty}\frac{1}{\lambda}e^{-\lambda y}dy\right] + \mu \\
&= \frac{\lambda}{2}\left[-\frac{1}{\lambda^2}e^{0} + \frac{1}{\lambda^2}e^{0}\right] + \mu \\
&= \mu
\end{aligned}
$$

Now, Variance

$$
\begin{aligned}
E[X^2] &= \int_{-\infty}^{\infty} x^2 \frac{\lambda}{2} e^{-\lambda|x-\mu|} dx \\
&= \int_{-\infty}^{\mu} x^2 \frac{\lambda}{2} e^{-\lambda(\mu-x)} dx + \int_{\mu}^{\infty} x^2 \frac{\lambda}{2} e^{-\lambda(x-\mu|)} dx \\
&= \frac{\lambda}{2} \left[x^2 \frac{1}{\lambda} e^{-\lambda(x-\mu)} \Bigg|_{-\infty}^{\mu} - \int_{-\infty}^{\mu} 2x \frac{1}{\lambda} e^{-\lambda(\mu-x)} dx + \int_{\mu}^{\infty} x^2 e^{-\lambda(x-\mu)} dx \right] \\
&= \frac{\lambda}{2} \left[\frac{\mu^2}{\lambda} - \left(2x \frac{1}{\lambda^2} e^{-\lambda(\mu-x)} \Bigg|_{-\infty}^{\mu} - \int_{-\infty}^{\mu} 2 \frac{1}{\lambda^2} e^{-\lambda(\mu-x)} dx \right) + \int_{\mu}^{\infty} x^2 e^{-\lambda(x-\mu)} dx \right] \\
&= \frac{\lambda}{2} \left[\frac{\mu^2}{\lambda} - \left(2\mu \frac{1}{\lambda^2} e^{0} - \frac{2}{\lambda^3} e^{-\lambda(\mu-x)} \Bigg|_{-\infty}^{\mu} \right) + \int_{\mu}^{\infty} x^2 e^{-\lambda(x-\mu)} dx \right] \\
&= \frac{\lambda}{2} \left[\frac{\mu^2}{\lambda} - \left(2\mu \frac{1}{\lambda^2} - \frac{2}{\lambda^3} e^{0} \right) + \int_{\mu}^{\infty} x^2 e^{-\lambda(x-\mu)} dx \right] \\
&= \frac{\mu^2}{2} - \frac{\mu}{\lambda} + \frac{1}{\lambda^2} + \frac{\lambda}{2} \int_{\mu}^{\infty} x^2 e^{-\lambda(x-\mu)} dx \\
&= \frac{\mu^2}{2} - \frac{\mu}{\lambda} + \frac{1}{\lambda^2} + \frac{\lambda}{2} \left[x^2 \frac{1}{-\lambda} e^{-\lambda(x-\mu)} \Bigg|_{\mu}^{\infty} - \int_{\mu}^{\infty} 2x \frac{1}{-\lambda} e^{-\lambda(x-\mu)} dx \right] \\
&= \frac{\mu^2}{2} - \frac{\mu}{\lambda} + \frac{1}{\lambda^2} + \frac{\lambda}{2} \left[-\mu^2 \frac{1}{-\lambda} e^{0} - \left(2x \frac{1}{\lambda^2} e^{-\lambda(x-\mu)} \Bigg|_{\mu}^{\infty} - \int_{\mu}^{\infty} 2 \frac{1}{\lambda^2} e^{-\lambda(x-\mu)} dx \right) \right] \\
&= \frac{\mu^2}{2} - \frac{\mu}{\lambda} + \frac{1}{\lambda^2} + \frac{\lambda}{2} \left[\frac{\mu^2}{\lambda} - \left(\frac{-2\mu}{\lambda^2} e^{0} - \left(\frac{2}{-\lambda^3} e^{-\lambda(x-\mu)} \Bigg|_{\mu}^{\infty} \right) \right) \right] \\
&= \frac{\mu^2}{2} - \frac{\mu}{\lambda} + \frac{1}{\lambda^2} + \frac{\lambda}{2} \left[\frac{\mu^2}{\lambda} - \left(\frac{-2\mu}{\lambda^2} - \left(\frac{2}{\lambda^3} e^{0} \right) \right) \right] \\
&= \frac{\mu^2}{2} - \frac{\mu}{\lambda} + \frac{1}{\lambda^2} + \frac{\lambda}{2} \left[\frac{\mu^2}{\lambda} + \frac{2\mu}{\lambda^2} + \frac{2}{\lambda^3} \right] \\
&= \frac{\mu^2}{2} - \frac{\mu}{\lambda} + \frac{1}{\lambda^2} + \frac{\mu^2}{2} + \frac{\mu}{\lambda} + \frac{1}{\lambda^2} \\
&= \mu^2 + \frac{2}{\lambda^2}
\end{aligned}
$$

Finally

$$
\begin{aligned}
Var(X) &= E[X^2] - E[X]^2 \\
&= \mu^2 + \frac{2}{\lambda^2} - \mu^2 \\
&= \frac{2}{\lambda^2}
\end{aligned}
$$

- *(b)* Let X_1, and X_2, be independent exponential random variables, each with parameter λ. The distribution of $Y = X_1 - X_2$ is a Laplace distribution with a mean of zero and a standard deviation of $\frac{\sqrt{2}}{\lambda}$. Set the seed equal to 3, and generate $25000 X_1$ values from an $Exp(\lambda = \frac{1}{2}$ and $25000 X_2$ values from another $Exp(\lambda = \frac{1}{2}$ distribution. Use these values to create the simulated distribution of $Y = X_1 - X_2$

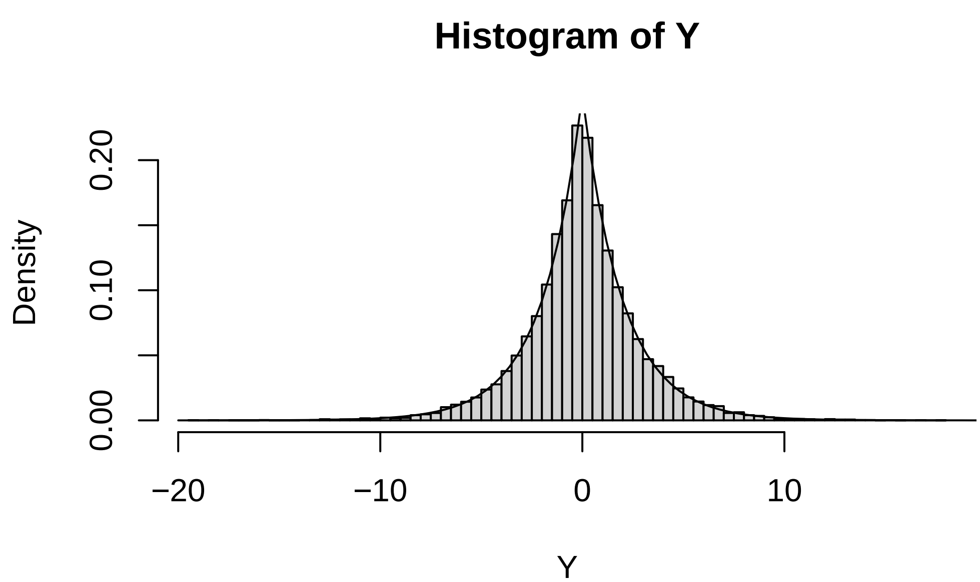
(i) Superimpose a Laplace distribution over a density histogram of the Y values (Hint: the R function `curve()` can be used to superimpose the Laplace distribution over the density histogram.)

(ii) Is the mean of Y within 0.02 of the theoretical mean?

(iii) Is the variance of Y within 2% of the theoretical answer?

```
# (b)
set.seed(3)
n <- 25000
X_1 <- rexp(n, 1/2)
X_2 <- rexp(n,1/2)
Y <- X_1-X_2
```

```
# (i)
hist(Y, freq= FALSE, breaks = 80)
curve(0.5/2*exp(-0.5*abs(x)),from=-20,to=20, add=TRUE)
```



```r
library(dplyr)
```

```r
# ii
theoretical_mean <- 0
between(theoretical_mean,
        theoretical_mean-0.02,theoretical_mean+0.2)
```

```
[1] TRUE
```

As we can see, the mean is within 0.02 of the theoretical.

```r
# iii
theoretical_var <-2/0.5**2
between(theoretical_var, theoretical_var-0.02*theoretical_var,
        theoretical_var+0.02*theoretical_var)
```

```
[1] TRUE
```

Also, the variance is within 2% of the theoretical answer.

Question 2. (4.4.24) Give a general expression to calculate the quantiles of a Weibull random variable.

$$f(X|\alpha,\beta) = \begin{cases} \alpha\beta^{-\alpha}x^{\alpha-1}e^{-(x/\beta)^\alpha} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

The p^{th} percentile is the value of x_p such that

$$\int_{-\infty}^{x_p} f(x)dx = \frac{p}{100} = \int_0^{x_p} \frac{\alpha}{\beta^\alpha} x^{\alpha-1} e^{-(x/\beta)^\alpha} dx$$

Then we can use a u-sub to solve the integral

$$\begin{aligned} \Rightarrow \int_0^{(x_p/\beta)^\alpha} e^{-u} du &= \frac{p}{100} \\ -e^{-u}\Big|_0^{(x_p/\beta)^\alpha} &= \\ -e^{(x_p/\beta)^\alpha} + 1 &= \\ 1 - \frac{p}{100} &= e^{(x_p/\beta)^\alpha} \\ -ln(1 - \frac{p}{100}) &= (x_p/\beta)^\alpha \\ \beta(-ln(1 - \frac{p}{100}))^{\frac{1}{\alpha}} &= x_p \end{aligned}$$

Question 3. (4.4.29) Let X be a random vatiable with probability density function

$$f(x) = 3(\frac{1}{x})^4, x \geq 1$$

- *(a)* Find the cumulative density function .

$$F(X) = \int_1^x 3t^{-4}dt = -x^3 - (-1)^3 = -x^3 + 1$$

- *(b)* Fix the seed at 98 (`set.seed(98)`) and generate a random sample of size $n = 100,000$ from X's distribution. Compute the mean, variance, and coefficient of skewness for the random sample.

```
set.seed(98)
n <- 100000
t <- runif(n)
X <- (1-t)**(-1/3)
mean(X)
```

```
[1] 1.50013
```

```
var(X)
```

```
[1] 0.7145073
```

```
X_star <- X-mean(X)
mean(X_star**3/(sqrt(var(X))**3))
```

```
[1] 10.63284
```

- *(c)* Obtain the theoretical mean, variance and coefficient of skewness of X.

First, mean

$$\begin{aligned} E[X] &= \int_1^\infty x3x^{-4}dx \\ &= \frac{3}{2} \end{aligned}$$

Variance

$$
\begin{aligned}
Var(X) &= E[X^2] - E[X]^2 \\
&= \int_1^\infty x^2 3x^{-4} dx - \frac{9}{4} \\
&= 3 - \frac{9}{4} \\
&= \frac{3}{4}
\end{aligned}
$$

Skewness

$$
\begin{aligned}
\gamma_1 &= \frac{E[(X-\mu)^3]}{\sigma^3} \\
&= \int_1^\infty (x-\mu)^3 3x^{-4} dx \\
&= \int_1^\infty \left(3x^{-1} - 9x^{-2}\mu + 9x^{-3}\mu^2 - 3x^{-4}\mu^3\right) dx \\
&= \infty \text{ , since } \lim_{x\to\infty} ln|x| = \infty
\end{aligned}
$$

- *(d)* How close are the estimates in (b) to the theoretical values in (c)?

```
cat("Mean is ",(mean(X)-1.5)/1.5*100, "% off of theoretical\n")
```

```
Mean is  0.008686381 % off of theoretical
```

```
cat("Variance is ", (var(X)-0.75)/0.75*100,"% off of theoretical")
```

```
Variance is  -4.732365 % off of theoretical
```

While skewness is not even close to theoretical.

Question 4. (4.4.32) Let X be a random variable with probability denstity function

$$
f(x) = (\theta+1)(1-x)^\theta, 0 \le x \le 1, \theta \ge 0
$$

- *(a)* Verify that the area under $f(x)$ is 1.

$$
\begin{aligned}
\int_{-\infty}^{\infty} (\theta+1)(1-x)^\theta dx &= (\theta+1)\int_0^1 (1-x)^\theta dx \\
&= -((1-1)^{\theta+1} - 1^{\theta+1}) \\
&= 1
\end{aligned}
$$

- *(b)* Find the cumulative density function.

$$\begin{aligned} F(X) &= \int_0^x (\theta+1)(1-t)^\theta dt \\ &= -(1-t)^{\theta+1}|_0^x \\ &= -(1-x)^{\theta+1}+1 \end{aligned}$$

- *(c)* What is $\mathbb{P}(X \leq .25|\theta = 2)$?

$$\begin{aligned} \mathbb{P}(X \leq 0.25|\theta=2) &= F(0.25) \\ &= -(1-0.25)^{2+1}+1 \\ &= 0.578125 \end{aligned}$$

- *(d)* Fix the seed at 80, and generate $100,000$ realizations of X with $\theta = 2$. What are the mean and variance of the random sample?

```r
set.seed(80)
n = 100000
t = runif(n)
x = 1-(-t+1)**(1/3)
cat("mean = ",mean(x),"\n")
```

```
mean =  0.2500348
```

```r
cat("variance = ",var(x))
```

```
variance =  0.03769803
```

- *(e)* Calculate the theoretical mean and variance of X when $\theta = 2$

mean when $\theta = 2$

$$\begin{aligned} E[X] &= \int_0^1 x3(1-x)^2 dx \\ &= 3\int_0^1 x - 2x^2 + x^3 dx \\ &= 3(\frac{1}{2} - \frac{2}{3} + \frac{1}{4}) \\ &= \frac{1}{4} \end{aligned}$$

variance when $\theta = 2$

$$
\begin{aligned}
Var(x) &= E[X^2] - E[X]^2 \\
&= \int_0^1 x^2 3(1-x)^2 dx - (\frac{1}{4})^2 \\
&= 3\int_0^1 x^2 - 2x^3 + x^4 dx - (\frac{1}{4})^2 \\
&= 3(\frac{1}{3} - \frac{2}{4} + \frac{1}{5}) - (\frac{1}{4})^2 \\
&= 0.0375
\end{aligned}
$$

- *(f)* How close are the estimates in (d) to the theoretical values in (e)?

```
cat("Mean is ",(mean(x)-1/4)/(1/4)*100, "% off of theoretical\n")
```

```
Mean is  0.01390041 % off of theoretical
```

```
cat("Variance is ", (var(X)-0.0375)/0.0375*100,"% off of theoretical")
```

```
Variance is  1805.353 % off of theoretical
```

Question 5. (4.4.34) A copper wire manufacturer produces conductor cables. These cables are of practical use if their resistance lies between 0.10 and 0.13 ohms per meter. The resistance of the cable follows a normal distribution, where 50% of the cables have a resistance under 0.11 ohms and 10% have a resistance over 0.13 ohms.

- *(a)* Determine the mean and the standard deviation for cable resistance.

Since the median is 0.11, and it follows a normal distribution, then the mean also has to be 0.11.

Then for standard deviation, since 10% of the cables have a resistance of 0.13 ohms, we have

$$
\begin{aligned}
\mathbb{P}(X \leq 0.13) &= 0.9 \\
\Rightarrow \mathbb{P}(Z = \frac{X - 0.11}{\sigma} \leq \frac{0.13 - 0.11}{\sigma}) &= 0.9
\end{aligned}
$$

Because $\mathbb{P}(Z \leq 1.28) = 0.9$, set

$$
\begin{aligned}
\frac{0.13 - 0.11}{\sigma} &= 1.28 \\
\Rightarrow \sigma &= \frac{0.13 - 0.11}{1.28} \\
&= 0.015625
\end{aligned}
$$

- *(b)* Find the probability that a randomly chosen cable can be used.

$$
\mathbb{P}(0.10 \leq X \leq 0.13) = \mathbb{P}(X \leq 0.13) - \mathbb{P}(X \leq 0.10)
$$

```
p<-pnorm(0.13,0.11,0.015625)-pnorm(0.10,0.11,0.015625)
cat("Probability that it can be used is ", p)
```

```
Probability that it can be used is  0.6386411
```

- *(c)* Find the probability that at least 3 out of 5 randomly chosen cables can be used.

We can use the binomial distribution to help us calculate this

$$P(X \geq 3) = \sum_{i=3}^{5} \binom{5}{i} \pi^i (1-\pi)^{5-i}$$

```
cat("Probability that at least 5 can be used is ",sum(dbinom(3:5,5,p)))
```

```
Probability that at least 5 can be used is  0.7469351
```