CS6190: Probabilistic Modeling Homework 0

Gopal Menon

22 January, 2018

1. Let X and Y be continuous real-valued random variables. Prove the following:

   (a) $E[E[X|Y]] = E[X]$

$$\begin{aligned}
E[E[X|Y]] &= \int_{-\infty}^{\infty} E[X|Y=y]\, p_Y(y) dy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\; p_{X|Y=y}(x) dx\; p_Y(y) dy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\; p_{X|Y=y}(x) p_Y(y) dx dy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\; p_{XY}(x,y) dx dy \\
&= \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} p_{XY}(x,y) dy\; dx \\
&= \int_{-\infty}^{\infty} x\; p_X(x) dx \\
&= E[X]
\end{aligned}$$

   Here $p_X(x)$, $p_Y(y)$ and $p_{X|Y=y}(x)$ are the probability density functions for X, Y, and the conditional probability of X given the value of Y. $p_{XY}(x,y)$ is the joint probability distribution function of X and Y. I needed to look up the proof [1] as I was not aware that the first step shown above could be done and due to this, I was getting an extra y in the final result. The result I got before I looked up the proof was $yE[X]$.

   (b) $Var(X) = E[Var(X|Y)] + Var(E[X|Y])$

$$\begin{aligned}
E[Var(X|Y)] + Var(E[X|Y]) &= E\left[E\left[X^2|Y\right] - E[X|Y]\, E[X|Y]\right] \\
&\quad + E[E[X|Y]\, E[X|Y]] - E[E[X|Y]]\, E[E[X|Y]] \\
&= E\left[E\left[X^2|Y\right]\right] - E[E[X|Y]\, E[X|Y]] \\
&\quad + E[E[X|Y]\, E[X|Y]] - E[X]\, E[X] \\
&= E\left[X^2\right] - E[X]\, E[X]\,, \text{ using the result from (a)} \\
&= Var(X)
\end{aligned}$$

   Before I worked out the solution above, I was using the definition $Var(X) = E\left[X^2\right] - \mu_X^2$ and that did not help. After I realized that $\mu_X = E[X]$, the solution fell in place.

2. Consider random variables X and Y with joint pdf

$$p(x,y) = \begin{cases} x+y, & \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, & \text{otherwise} \end{cases}$$

(a) What is the marginal pdf $p(x)$?

$$\begin{aligned}
p(x) &= \int_{-\infty}^{\infty} p(x,y)dy \\
&= \begin{cases} \int_0^1 (x+y)dy, \text{for } 0 \le x \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \left(xy + \frac{1}{2}y^2\right)\Big|_0^1, \text{for } 0 \le x \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} x + \frac{1}{2}, \text{for } 0 \le x \le 1 \\ 0, \text{otherwise} \end{cases}
\end{aligned}$$

(b) What is the conditional pdf $p(y|x)$?

$$\begin{aligned}
p(y|x) &= \frac{p(x,y)}{p(x)} \\
&= \begin{cases} \frac{x+y}{x+\frac{1}{2}}, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases}
\end{aligned}$$

(c) What is the conditional expectation $E[Y|X]$?

$$\begin{aligned}
E[Y|X] &= \int_{-\infty}^{\infty} y p_{Y|X}(y)dy \\
&= \begin{cases} \int_0^1 \frac{xy+y^2}{x+\frac{1}{2}}dy, \text{for } 0 \le x \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \frac{\frac{1}{2}xy^2 + \frac{1}{3}y^3}{x+\frac{1}{2}}\Big|_0^1, \text{for } 0 \le x \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \frac{\frac{1}{2}x + \frac{1}{3}}{x+\frac{1}{2}}, \text{for } 0 \le x \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \frac{\frac{3x+2}{6}}{\frac{2x+1}{2}}, \text{for } 0 \le x \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \frac{3x+2}{6x+3}, \text{for } 0 \le x \le 1 \\ 0, \text{otherwise} \end{cases}
\end{aligned}$$

(d) What is the covariance $Cov(X,Y)$?

$$Cov(X,Y) = E\left[XY\right] - E\left[X\right]E\left[Y\right]$$

$$\begin{aligned}
E\left[XY\right] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyp(x,y)dxdy \\
&= \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} yp(x,y)dydx
\end{aligned}$$

$$\begin{aligned}
\int_{-\infty}^{\infty} yp(x,y)dy &= \begin{cases} \int_0^1 y(x+y)dy, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \int_0^1 (xy+y^2)dy, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \frac{1}{2}xy^2 + \frac{1}{3}y^3\Big|_0^1, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \frac{1}{2}x + \frac{1}{3}, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases}
\end{aligned}$$

$$\begin{aligned}
\int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} yp(x,y)dydx &= \begin{cases} \int_0^1 x\left(\frac{1}{2}x + \frac{1}{3}\right)dx, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \int_0^1 \left(\frac{1}{2}x^2 + \frac{1}{3}x\right)dx, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \left(\frac{1}{2}\frac{1}{3}x^3 + \frac{1}{3}\frac{1}{2}x^2\right)\Big|_0^1, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \left(\frac{1}{6} + \frac{1}{6}\right), \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \frac{1}{3}, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= E\left[XY\right]
\end{aligned}$$

$$
\begin{aligned}
E\left[X\right] &= \int_{-\infty}^{\infty} xp(x)dx \\
&= \begin{cases} \displaystyle\int_{-\infty}^{\infty} x\left(x+\frac{1}{2}\right)dx, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \displaystyle\int_{-\infty}^{\infty} \left(x^2+\frac{1}{2}x\right)dx, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \left(\dfrac{1}{3}x^3+\dfrac{1}{2}\dfrac{1}{2}x^2\right)\Big|_0^1, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \dfrac{1}{3}+\dfrac{1}{4}, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \dfrac{7}{12}, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
p(y) &= \int_{-\infty}^{\infty} p(x,y)dx \\
&= \begin{cases} \displaystyle\int_{0}^{1} (x+y)dx, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \left(\dfrac{1}{2}x^2+yx\right)\Big|_0^1, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} y+\dfrac{1}{2}, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases}
\end{aligned}
$$

$$
\begin{aligned}
E\left[Y\right] &= \int_{-\infty}^{\infty} yp(y)dy \\
&= \begin{cases} \displaystyle\int_{-\infty}^{\infty} y\left(y+\frac{1}{2}\right)dy, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \displaystyle\int_{-\infty}^{\infty} \left(y^2+\frac{1}{2}y\right)dy, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \left(\dfrac{1}{3}y^3+\dfrac{1}{2}\dfrac{1}{2}y^2\right)\Big|_0^1, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \dfrac{1}{3}+\dfrac{1}{4}, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \dfrac{7}{12}, \text{for } 0 \le x \le 1, \text{and } 0 \le y \le 1 \\ 0, \text{otherwise} \end{cases}
\end{aligned}
$$

$$
\begin{aligned}
Cov(X,Y) &= E\left[XY\right] - E\left[X\right]E\left[Y\right] \\
&= \begin{cases} \frac{1}{3} - \frac{7}{12}\frac{7}{12}, \text{for } 0 \leq x \leq 1, \text{and } 0 \leq y \leq 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} \frac{48}{144} - \frac{49}{144}, \text{for } 0 \leq x \leq 1, \text{and } 0 \leq y \leq 1 \\ 0, \text{otherwise} \end{cases} \\
&= \begin{cases} -\frac{1}{144}, \text{for } 0 \leq x \leq 1, \text{and } 0 \leq y \leq 1 \\ 0, \text{otherwise} \end{cases}
\end{aligned}
$$

3. Let $X \sim Exp(\lambda)$, i.e., the exponential distribution with pdf $p(x) = \lambda exp(-\lambda x)$. Let $Y = \sqrt{X}$.
   (a) What is the density function $p(y)$?

$$
\begin{aligned}
Y &= \sqrt{X} \\
f(x) &= \sqrt{x} \\
f^{-1}(y) &= y^2 \\
p(y) &= \left| \frac{d}{dy}(f^{-1}(y)) \right| p(f^{-1}(y)) \\
&= \left| \frac{d}{dy} y^2 \right| p(y^2) \\
&= \begin{cases} 2y\lambda exp(-\lambda y^2), \text{for } y \geq 0 \\ 0, \text{otherwise} \end{cases}
\end{aligned}
$$

   (b) What is the cdf, $F(y) = P(Y \leq y)$? Verify that $F(0) = 0$ and $F(\infty) = 1$.

The cdf can be found by integrating the pdf.

$$
\begin{aligned}
F(y) = P(Y \leq y) &= \int_0^y 2y\lambda exp(-\lambda y^2)dy \\
\text{Substituting } u = -\lambda y^2 \text{, we get } dy &= -\frac{1}{2\lambda y}du \\
\int 2y\lambda exp(-\lambda y^2)dy &= \int 2y\lambda exp(u) \times -\frac{1}{2\lambda y}du \\
&= -\int exp(u)du \\
&= -exp(u) \\
&= -exp(-\lambda y^2) \\
F(y) &= -exp(-\lambda y^2)\Big|_0^y \\
&= -exp(-\lambda y^2) - (-exp(-\lambda 0^2)) \\
&= -exp(-\lambda y^2) + 1 \\
&= 1 - exp(-\lambda y^2) \\
F(0) &= 1 - exp(-\lambda 0^2) \\
&= 1 - 1 \\
&= 0 \\
F(\infty) &= 1 - exp(-\infty) \\
&= 1 - 0 \\
&= 1
\end{aligned}
$$

(c) What is the quantile function F^{-1}?

If q_p is the quantile value given by the quantile function $F^{-1}(q_p)$, based on the cdf above:

$$
\begin{aligned}
q_p &= 1 - exp(-\lambda(F^{-1}(q_p))^2) \\
exp(-\lambda(F^{-1}(q_p))^2) &= 1 - q_p \\
-\lambda(F^{-1}(q_p))^2 &= \ln(1 - q_p) \\
(F^{-1}(q_p))^2 &= \frac{-1}{\lambda}\ln(1 - q_p) \\
F^{-1}(q_p) &= \sqrt{\frac{-1}{\lambda}\ln(1 - q_p)}
\end{aligned}
$$

(d) Compute the mean, $E[Y]$, and variance, $Var(Y)$. **Hint:** Use integration by parts.

$$
\begin{aligned}
E[Y] &= E\left[\sqrt{X}\right] \\
&= \int_{-\infty}^{\infty} \sqrt{x}p(x)dx \\
&= \int_0^{\infty} \sqrt{x}\lambda exp(-\lambda x)dx \\
&= \frac{\sqrt{\pi}}{2\sqrt{\lambda}}
\end{aligned}
$$

$$
\begin{aligned}
Var(Y) &= E\left[Y^2\right] - E\left[Y\right]E\left[Y\right] \\
&= E\left[X\right] - E\left[Y\right]E\left[Y\right] \\
E\left[X\right] &= \int_{-\infty}^{\infty} x\lambda exp(-\lambda x)dx \\
&= \frac{1}{\lambda} \\
Var(Y) &= \frac{1}{\lambda} - \frac{\sqrt{\pi}}{2\sqrt{\lambda}}\frac{\sqrt{\pi}}{2\sqrt{\lambda}} \\
&= \frac{1}{\lambda} - \frac{\pi}{4\lambda} \\
&= \frac{4-\pi}{4\lambda}
\end{aligned}
$$

The two definite integrals were done online [2] as I could not figure out how to do the integration by parts. I did not write down all the intermediate steps.

4. Given a realization $y_1, y_2, \ldots, y_n$ from your random variable Y in the previous problem, what is the maximum likelihood estimate for λ?

   The likelihood function will be

$$
\begin{aligned}
\mathcal{L}(\lambda) &= p_{Y_1,\ldots,Y_n}(y_1, y_2, \ldots, y_n; \lambda) \\
&= \prod_{i=1}^{n} p_{Y_i}(y_i; \lambda), && \text{assuming } Y_i \text{ are independently distributed} \\
&= \prod_{i=1}^{n} \left(2y\lambda exp(-\lambda y_i^2)\right) \\
\ln(\mathcal{L}(\lambda)) = \ell(\lambda) &= \sum_{i=1}^{n} \ln\left(2y_i\lambda exp(-\lambda y_i^2)\right) \\
&= \sum_{i=1}^{n} \ln(2y_i) + \sum_{i=1}^{n} \ln(\lambda) - \sum_{i=1}^{n} \lambda y_i^2
\end{aligned}
$$

   In order to find the maximum likelihood, we can find the maximum log likelihood since log is a monotonously increasing function. The first term above is a constant and we can find the maximum likelihood by setting the derivative with respect to λ of the log likelihood to zero.

$$
\begin{aligned}
\frac{d\ell(\lambda)}{d\lambda} &= \frac{d\sum_{i=1}^{n}\ln(2y_i)}{d\lambda} + \frac{d\sum_{i=1}^{n}\ln(\lambda)}{d\lambda} - \frac{d\sum_{i=1}^{n}\lambda y_i^2}{d\lambda} \\
&= 0 + n\frac{d\ \ln(\lambda)}{d\lambda} - \sum_{i=1}^{n} y_i^2 \frac{d\lambda}{d\lambda} \\
&= n\frac{1}{\lambda} - \sum_{i=1}^{n} y_i^2 = 0 \\
\frac{n}{\lambda} &= \sum_{i=1}^{n} y_i^2 \\
\lambda &= \frac{n}{\sum_{i=1}^{n} y_i^2}, && \text{which is the maximum likely estimate for } \lambda.
\end{aligned}
$$

R Coding Part

4. Recall that if F is a cdf and $U \sim \mathrm{Unif}(0,1)$, then $Y = F^{-1}(U)$ will be a random variable with cdf F. Write a function that uses this method to generate n random numbers from the distribution of Y in

Problem 3 above (n and λ should be parameters to the function).
Hint: Look at the `runif` function in R for simulating uniform random varibles.

(a) Generate 10,000 realizations of the random variable Y with $\lambda = 2$.

(b) Plot a histogram of these numbers, using the `hist` function with option `freq = FALSE`.

(c) Use the `lines` command to plot the pdf of Y on top.

(d) Compute the sample mean and variance of your 10,000 realizations. Do they roughly match the values for $\mathrm{E}[Y]$ and $\mathrm{Var}(Y)$ you calculated above?

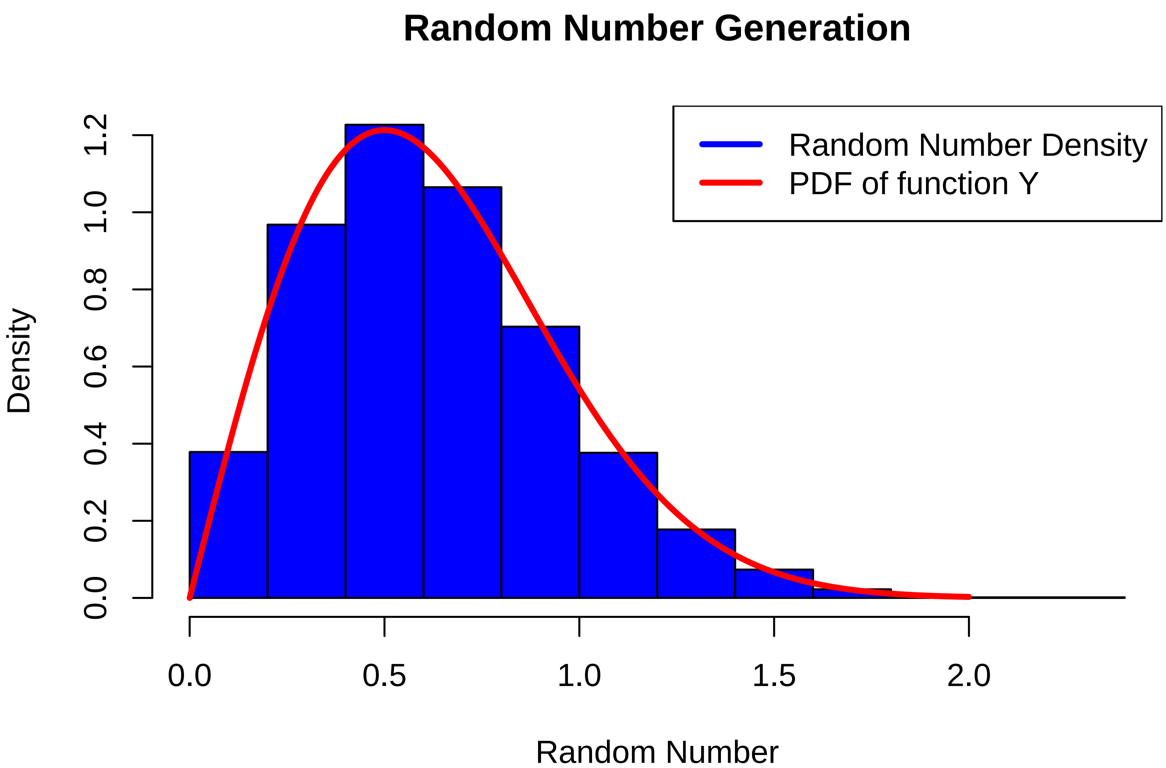
(a)

```
# Return the pth quantile for distribution of Y
get_Y_quantile <- function(quantile_value, exponential_rate) {
  return(sqrt(-1* log(1 - quantile_value) / exponential_rate))
}

# Random number generation function
generate_random_number <- function(how_many, exponential_rate) {

  # Uniform distribution between 0 and 1
  uniform_distribution = runif(how_many, min=0, max=1)

  # Declare an array to store the random numbers from the distribution of Y
  ramdoms_from_exp_distr = numeric(how_many)

  # For each number in the uniform distribution find the quantile of Y
  for (unif_distr_counter in 1:length(uniform_distribution)) {
    ramdoms_from_exp_distr[unif_distr_counter] =
      get_Y_quantile(uniform_distribution[unif_distr_counter], exponential_rate)
  }

  return(ramdoms_from_exp_distr)

}

# Return probability density of Y
y_probability_density <- function(real_number, exponential_rate) {

  return (2*real_number*exponential_rate*exp(-1 * exponential_rate * real_number^2))

}

#Generate 10,000 realizations of the random variable Y with lambda = 2
realizations_needed = 10000
exponential_rate_lambda = 2

random_numbers = generate_random_number(how_many = realizations_needed,
                                        exponential_rate = exponential_rate_lambda)
```

(b)

```
# Plot a histogram of these numbers, using the hist function with option freq = FALSE.
hist(random_numbers, freq=FALSE, main= "Random Number Generation", xlab="Random Number",
     ylab="Density", col="blue")
# Use the lines command to plot the pdf of Y on top.
```

```r
number_sequence = seq(0,2,.001)
y_probability_density_trend = y_probability_density(real_number = number_sequence,
                              exponential_rate = exponential_rate_lambda)

lines(number_sequence, y_probability_density_trend, lty=1, lwd=3, col="red")
legend("topright", legen=c("Random Number Density", "PDF of function Y"),
       col=c("blue", "red"), lwd=3, lty=1)
```



(c)

(d)

```r
## Compute the sample mean and variance of your 10,000 realizations
mean(random_numbers)
```

```
## [1] 0.6258609
```

```r
var(random_numbers)
```

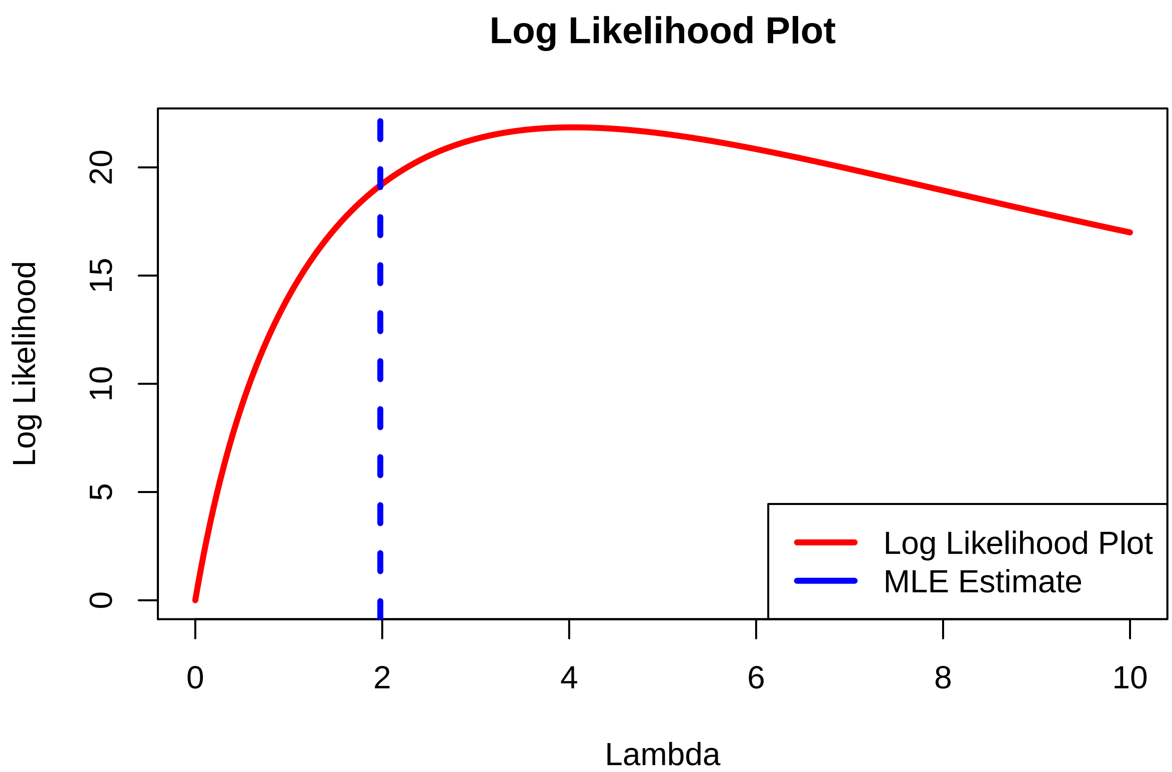
```
## [1] 0.1069342
```

$$
\begin{aligned}
E[Y] &= \frac{\sqrt{\pi}}{2\sqrt{\lambda}} \\
&= \frac{\sqrt{\pi}}{2\sqrt{2}} \\
&= 0.6267, \text{ which roughly matches the experimental value above}
\end{aligned}
$$

$$
\begin{aligned}
Var(Y) &= \frac{4-\pi}{4\lambda} \\
&= \frac{4-\pi}{4\times 2} \\
&= 0.1073, \text{ which roughly matches the experimental value above}
\end{aligned}
$$

5. Now generate 20 realizations from Y with $\lambda = 2$. Plot the log likelihood function for this data. Plot a vertical line at the maximum likelihood estimate for lambda (using the equation you got in Problem 3).

```
## Generate 20 realizations from Y with lambda= 2
new_realizations_needed = 20
random_numbers = generate_random_number(how_many = new_realizations_needed,
                                        exponential_rate = exponential_rate_lambda)
random_numbers = sort(random_numbers)

## Compute the likelihood distribution
log_likelihood_distribution = numeric(length(random_numbers))
lambda_values = seq(0,10,.01)

# Compute the log likelihood for at each lamda point
lambda_value_counter = 0
for (lambda_value in lambda_values) {

  lambda_value_counter = lambda_value_counter + 1
  log_likelihood = 0

  # Compute log likelihood at the point corresponding to the random number
  for(realization_counter in 1 : new_realizations_needed) {
    log_likelihood = log_likelihood +
      y_probability_density(real_number = random_numbers[realization_counter],
                            exponential_rate = lambda_values[lambda_value_counter])
  }

  log_likelihood_distribution[lambda_value_counter] = log_likelihood

}

# Plot the likelihood function
plot(lambda_values, log_likelihood_distribution, type='l', lwd=3, col="red",
     main= "Log Likelihood Plot", xlab="Lambda", ylab="Log Likelihood")

## Maximum likelihood estimate for lambda
random_number_distributions_squared = random_numbers^2
lambda_hat = new_realizations_needed / sum(random_number_distributions_squared)

## Draw a vertical line at the computed value of maximum likelihood
abline(v = lambda_hat, col='blue', lwd = 3, lty = 2)
legend("bottomright", legen=c("Log Likelihood Plot", "MLE Estimate"),
       col=c("red", "blue"), lwd=3, lty=1)
```



References

[1] *Conditional Expectation*, www.statlect.com/fundamentals-of-probability/conditional-expectation.

[2] "Integral Calculator." *Integral Calculator With Steps!*, www.integral-calculator.com/.