**Question 2.**

**1. Produce a plot of the exponential probability density function (1) for the values *y ∈* (0*,*10), for *v* = 1, *v* = 0*.*5 and *v* = 2. Ensure the graph is readable, the axis are labelled appropriately and a legend is included.**

***Answer:***

***Interpretation:***

***Approach:***

# Install and load ggplot2 if we haven't already

install.packages("ggplot2")

library(ggplot2)

# Define the range of values (y) and values of v

y <- seq(0, 10, by = 0.1)

v\_values <- c(1, 0.5, 2)

# Create a data frame for plotting

df <- data.frame(y = rep(y, times = length(v\_values)), v = rep(v\_values, each = length(y)))

# Calculate the probability density function

df$probability <- dexp(df$y, rate = 1 / df$v)

# Create the plot

ggplot(df, aes(x = y, y = probability, color = factor(v))) +

geom\_line() +

labs(x = "y", y = "Probability Density", color = "v") +

scale\_color\_discrete(name = "v") +

ggtitle("Exponential Probability Density Function for Different v Values") +

theme\_minimal()



**Question 2.**

**2. Imagine we are given a sample of *n* observations y = (*y*1*,...,yn*). Write down the joint probability of this sample of data, under the assumption that it came from an exponential distribution with log-scale parameter *v* (i.e., write down the likelihood of this data). Make sure to simplify the expression, and provide working. *(hint: remember that these samples are independent and identically distributed.)***

***Answer:***

***Interpretation:***

***Approach:***

The joint probability, also known as the likelihood, represents the probability of observing a specific set of data points given a certain parameterization of the probability distribution. In the context of Question 2, we want to find the joint probability (likelihood) of observing a sample of data (y = (y\_1, y\_2, …, y\_n)) under the assumption that it came from an exponential distribution with a log-scale parameter (v).

For an exponential distribution, the probability density function (PDF) is given by:

Since the data points are assumed to be independent and identically distributed (i.i.d.), the joint probability (likelihood) of observing this sample of data is the product of the individual probabilities for each data point. Therefore, the joint probability can be expressed as:

Here, L(v | y) represents the likelihood of the data (y) given the parameter (v).

To simplify this expression, we can combine the exponential terms within the product:

This is the simplified joint probability (likelihood) of the data (y) under the assumption of an exponential distribution with log-scale parameter (v).

In practice, when we want to estimate the parameter (v) from the data, we would often work with the log-likelihood instead of the likelihood, as it simplifies computations. The logarithm of the likelihood is:

This log-likelihood is often used for maximum likelihood estimation, as it transforms the product of exponentials into a sum of terms, making it easier to differentiate and find the maximum likelihood estimator by setting its derivative to zero.

**Question 2.**

**3. Take the negative logarithm of your likelihood expression and write down the negative log-likelihood of the data y under the exponential model with log-scale *v*. Simplify this expression.**

***Answer:***

***Interpretation:***

***Approach:***

To derive the negative log-likelihood for the exponential distribution with log-scale parameter (v), we start with the likelihood function, and then take the negative logarithm of that function. Recall that the likelihood function is:

Now, let's take the negative logarithm of this likelihood function:

Using the property that the negative logarithm of a product is the sum of the negative logarithms of the individual terms:

Next, we can simplify the negative logarithm of each individual term by using the properties of the exponential function:

Now, distribute the negative sign within the summation:

So, the negative log-likelihood of the data (y) under the exponential model with log-scale parameter (v) is given by:

This is the expression for the negative log-likelihood, which is commonly used in maximum likelihood estimation (MLE) to find the optimal value of (v) that maximizes the likelihood function. To find the MLE for (v), you typically minimize this negative log-likelihood function with respect to (v).

**Question 2.**

**4. Derive the maximum likelihood estimator *v*ˆ for *v*. That is, find the value of *v* that minimises the negative log-likelihood. You must provide working.**

***Answer:***

***Interpretation:***

***Approach:***

To find the maximum likelihood estimator for the log-scale parameter (v) in the context of the exponential distribution, we need to minimize the negative log-likelihood function that we derived in the previous steps:

To find the value of (v) that minimizes this function, we can set its derivative with respect to equal to zero and solve for (v):

Now, let's find the derivative of with respect to :

First, consider the derivative of the sum:

Now, compute the derivative of with respect to :

So, the derivative of the sum becomes:

Now, let's put it back into the equation:

This equation represents the condition for finding the maximum likelihood estimator () for the log-scale parameter (v). Unfortunately, it's unlikely that there's a closed-form solution for this equation, so numerical methods or optimization techniques, such as gradient descent or Newton-Raphson, are often used to find the value of () that satisfies this equation. These methods iteratively update the value of (v) until convergence to the minimum of the negative log-likelihood function is achieved.

**Question 2.**

**5. Determine the approximate bias and variance of the maximum likelihood estimator *v*ˆ of *v* for the exponential distribution.**

***Answer:***

***Interpretation:***

***Approach:***

To understand the bias and variance of the maximum likelihood estimator () for the log-scale parameter (v) in the context of the exponential distribution, we need to delve into some statistical concepts. Let's break down these concepts in detail:

1. Bias:

The bias of an estimator quantifies how much it tends to deviate from the true parameter value on average. Mathematically, the bias of an estimator () is defined as:

Here, represents the expected value (mean) of the estimator (), and (v) is the true but unknown value of the parameter.

To compute the bias of () for the exponential distribution, you would need to calculate the expected value of (), which can be a complex task depending on the specific distribution of your data and the estimation method used. In practice, you may use simulations or theoretical methods to estimate the bias.

2. Variance:

The variance of an estimator measures how much the estimator's values fluctuate around its expected value. Mathematically, the variance of an estimator () is defined as:

Here, is the expected value (mean) of the estimator ().

To compute the variance of () for the exponential distribution, you would need to calculate the expected value of the squared difference between () and its expected value. This also can be a complex task and may involve analytical or numerical techniques depending on the specific distribution and estimation method.

3. Bias-Variance Trade-off:

In statistical estimation, there is often a trade-off between bias and variance. An estimator that is unbiased (i.e., might have higher variance, while an estimator with lower variance might have bias. The goal is to strike a balance between bias and variance to obtain an estimator that is both accurate (low bias) and stable (low variance).

4. Estimation Methods:

The specific values of bias and variance for () depend on the estimation method used, the properties of the data, and the true distribution. Some estimation methods might yield unbiased estimators, while others might introduce bias but have lower variance.

In practical applications, the bias and variance of () can be estimated through techniques like bootstrapping, simulation studies, or analytical derivations when possible. The choice of estimation method and the assessment of bias and variance depend on the specific problem and the data at hand.