

1: Problem 1

Let X_1, X_2, \dots, X_n be independent geometric distribution. Consider the likelihood function

$$T(X_1, \dots, X_n) = \sum_{k=1}^n X_k$$

Then

$$\begin{aligned} P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n | T = t) &= \frac{P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n, T = t)}{P(T = t)} \\ &= \frac{P(X_1 = x_1) \cdots P(X_n = x_n)}{P(T = t)} \\ &= \frac{p^n (1-p)^{\sum_{i=1}^n (x_i - 1)}}{\binom{t-1}{n-1} p^n (1-p)^{t-n}} \text{Negative Binomial Distribution} \\ &= \frac{p^n (1-p)^{t-n}}{\binom{t-1}{n-1} p^n (1-p)^{t-n}} \\ &= \frac{1}{\binom{t-1}{n-1}} \end{aligned}$$

which does not depend on p .

To construct a $(1 - \alpha) \times 100\%$ confidence interval for the parameter p in the geometric distribution, we can use the maximum likelihood estimator (MLE) of p , which is given by:

$$\hat{p} = \frac{1}{\bar{x}},$$

where \bar{x} is the sample mean. The distribution of \hat{p} is approximately normal for large sample sizes, and we can use this fact to construct a confidence interval.

The standard error of \hat{p} is given by:

$$SE(\hat{p}) = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}},$$

where n is the sample size. We can use this to construct a confidence interval for p as:

$$\hat{p} \pm z_{\alpha/2} SE(\hat{p}),$$

where $z_{\alpha/2}$ is the $(1 - \alpha/2)$ -th quantile of the standard normal distribution.

Substituting the expression for \hat{p} and $SE(\hat{p})$ and simplifying, we get the confidence interval:

$$\left(\frac{1}{\bar{x}} - z_{\alpha/2} \sqrt{\frac{1 - 1/n}{\bar{x}}}, \frac{1}{\bar{x}} + z_{\alpha/2} \sqrt{\frac{1 - 1/n}{\bar{x}}} \right).$$

Therefore, a $(1 - \alpha) \times 100\%$ confidence interval for the parameter p in the geometric distribution is given by:

$$\left(\frac{1}{\bar{x}} - z_{\alpha/2} \sqrt{\frac{1 - 1/n}{\bar{x}}}, \frac{1}{\bar{x}} + z_{\alpha/2} \sqrt{\frac{1 - 1/n}{\bar{x}}} \right).$$

In my derivation, the geometric distribution was used to obtain the maximum likelihood estimator of p , which is given by $\hat{p} = 1/\bar{x}$, where \bar{x} is the sample mean.

To see how this comes about, let X_1, X_2, \dots, X_n be independent and identically distributed random variables from a geometric distribution with parameter p . The probability mass function of X is given by:

$$P(X = k) = p(1 - p)^{k-1}, \quad k = 1, 2, \dots$$

The likelihood function for this sample is:

$$L(p; x_1, x_2, \dots, x_n) = \prod_{i=1}^n p(1 - p)^{x_i-1} = p^n (1 - p)^{\sum_{i=1}^n (x_i-1)},$$

where x_i is the observed value of X_i .

To obtain the maximum likelihood estimator of p , we take the derivative of the likelihood function with respect to p and set it equal to zero:

$$\frac{d}{dp} L(p; x_1, x_2, \dots, x_n) = np^{n-1}(1 - p)^{\sum_{i=1}^n (x_i-1)} - p^n \sum_{i=1}^n (x_i - 1)(1 - p)^{\sum_{j=1}^n (x_j-1)} = 0.$$

Solving for p , we obtain:

$$\hat{p} = \frac{1}{\bar{x}},$$

where $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ is the sample mean. Therefore, \hat{p} is the maximum likelihood estimator of p based on the sample x_1, x_2, \dots, x_n from the geometric distribution.

(C) Consider $H_0 : p = p_0$ and $H_A : p = p_1$, where $p_1 > p_0$. Use Neyman Pearson Lemma to find the most powerful test. Is this test a UMP test? Justify your answer.

Rayleigh distribution:

Consult page 333 for example of usage

MLE of this distribution:

To find the maximum likelihood estimator (MLE) of the scale parameter σ in the Rayleigh distribution, we start with the probability density function (pdf) of the Rayleigh distribution:

$$f(x; \sigma) = \frac{x}{\sigma^2} e^{-x^2/(2\sigma^2)}, \quad x \geq 0,$$

where $\sigma > 0$ is the scale parameter.

Suppose we have a random sample X_1, X_2, \dots, X_n from the Rayleigh distribution. The likelihood function is:

$$L(\sigma; x_1, x_2, \dots, x_n) = \prod_{i=1}^n f(x_i; \sigma) = \frac{1}{\sigma^{2n}} \exp \left[-\frac{1}{2\sigma^2} \sum_{i=1}^n x_i^2 \right] \prod_{i=1}^n x_i,$$

where x_i is the observed value of X_i . Taking the logarithm of the likelihood function, we obtain:

$$\log L(\sigma; x_1, x_2, \dots, x_n) = -2n \log \sigma - \frac{1}{2\sigma^2} \sum_{i=1}^n x_i^2 + \sum_{i=1}^n \log x_i.$$

To find the MLE of σ , we differentiate this expression with respect to σ , set the derivative equal to zero, and solve for σ . We get:

$$\hat{\sigma} = \sqrt{\frac{1}{2n} \sum_{i=1}^n X_i^2}.$$

Therefore, the MLE of the scale parameter σ in the Rayleigh distribution is $\hat{\sigma} = \sqrt{\frac{1}{2n} \sum_{i=1}^n X_i^2}$.

Sufficient Statistic

To find a sufficient statistic for $\lambda = \sigma^2$ in the Rayleigh distribution, we can use the factorization theorem, which states that a statistic $T(X)$ is sufficient for a parameter θ if and only if the likelihood function can be factorized as:

$$L(\theta; x) = g[T(x), \theta]h(x),$$

where g and h are functions that do not depend on θ and $T(x)$ is a statistic.

For the Rayleigh distribution, the likelihood function is:

$$L(\sigma; x_1, x_2, \dots, x_n) = \prod_{i=1}^n \frac{x_i}{\sigma^2} e^{-x_i^2/(2\sigma^2)} = \frac{1}{\sigma^{2n}} \exp \left[-\frac{1}{2\sigma^2} \sum_{i=1}^n x_i^2 \right] \prod_{i=1}^n x_i.$$

If we define $T(x) = \sum_{i=1}^n x_i^2$, then the likelihood function can be written as:

$$L(\sigma; x_1, x_2, \dots, x_n) = \frac{1}{\sigma^{2n}} \exp \left[-\frac{T(x)}{2\sigma^2} \right] \left(\prod_{i=1}^n x_i \right),$$

which can be written in the form required by the factorization theorem, with $g[T(x), \lambda] = \exp \left[-\frac{T(x)}{2\lambda} \right]$ and $h(x) = (\prod_{i=1}^n x_i) / \lambda^n$. Therefore, $T(x) = \sum_{i=1}^n x_i^2$ is a sufficient statistic for $\lambda = \sigma^2$ in the Rayleigh distribution.

Confidence Interval

To construct a $(1 - \alpha) \times 100\%$ confidence interval for the parameter $\lambda = \sigma^2$ in the Rayleigh distribution using the MLE, we can use the asymptotic normality of the MLE, as $n \rightarrow \infty$. Specifically, the MLE $\hat{\lambda}$ is asymptotically normal with mean λ and variance $\frac{\lambda^2}{2n}$.

Using this fact, a $(1 - \alpha) \times 100\%$ confidence interval for λ is given by:

$$\left(\hat{\lambda} - z_{\alpha/2} \sqrt{\frac{\hat{\lambda}^2}{2n}}, \hat{\lambda} + z_{\alpha/2} \sqrt{\frac{\hat{\lambda}^2}{2n}} \right),$$

where $z_{\alpha/2}$ is the $(1 - \alpha/2)$ -th quantile of the standard normal distribution.

Substituting the MLE $\hat{\lambda} = \frac{1}{2n} \sum_{i=1}^n X_i^2$, we get:

$$\left(\frac{1}{2n} \sum_{i=1}^n X_i^2 - z_{\alpha/2} \sqrt{\frac{1}{2n} \left(\frac{1}{2n} \sum_{i=1}^n X_i^2 \right)^2}, \frac{1}{2n} \sum_{i=1}^n X_i^2 + z_{\alpha/2} \sqrt{\frac{1}{2n} \left(\frac{1}{2n} \sum_{i=1}^n X_i^2 \right)^2} \right).$$

Therefore, a $(1 - \alpha) \times 100\%$ confidence interval for the parameter $\lambda = \sigma^2$ in the Rayleigh distribution is given by:

$$\left(\frac{1}{2n} \sum_{i=1}^n X_i^2 - z_{\alpha/2} \sqrt{\frac{1}{2n} \left(\frac{1}{2n} \sum_{i=1}^n X_i^2 \right)^2}, \frac{1}{2n} \sum_{i=1}^n X_i^2 + z_{\alpha/2} \sqrt{\frac{1}{2n} \left(\frac{1}{2n} \sum_{i=1}^n X_i^2 \right)^2} \right).$$

Hypothesis testing

To find the most powerful test for the hypotheses $H_0 : \lambda = \lambda_0$ and $H_A : \lambda = \lambda_1$, where $\lambda_1 > \lambda_0$, we can use the Neyman-Pearson lemma, which states that the most powerful test at a given significance level α is the likelihood ratio test.

The likelihood ratio test compares the likelihood of the observed data under the null and alternative hypotheses. Specifically, the likelihood ratio test statistic is:

$$\Lambda(x) = \frac{L(\lambda_1; x)}{L(\lambda_0; x)},$$

where $L(\lambda; x)$ is the likelihood function for the parameter λ based on the observed data x .

If $\Lambda(x)$ is greater than a threshold value k , we reject the null hypothesis in favor of the alternative hypothesis. The threshold value k is chosen such that the test has a significance level of α .

To find the threshold value k , we can use the fact that the likelihood ratio test is a monotone test. That is, for any two values of λ and λ' , if $\lambda > \lambda'$, then the likelihood ratio $\Lambda(x)$ is a decreasing function of x . Therefore, we can choose k such that the probability of rejecting the null hypothesis when the alternative hypothesis is true (i.e., the power of the test) is equal to α , or:

$$P(\Lambda(x) > k | \lambda = \lambda_1) = \alpha.$$

This probability can be calculated using the distribution of the likelihood ratio test statistic under the alternative hypothesis. Specifically, if X_1, X_2, \dots, X_n are independent and identically distributed random variables from a Rayleigh distribution with parameter λ_1 , then the distribution of $-2 \log \Lambda(X_1, X_2, \dots, X_n)$ is chi-squared with one degree of freedom.

Using this, we can find the threshold value k such that

$P(-2 \log \Lambda(X_1, X_2, \dots, X_n) > k | \lambda = \lambda_1) = \alpha$. Equivalently, k satisfies:

$$P(-2 \log \Lambda(X_1, X_2, \dots, X_n) \leq k | \lambda = \lambda_0) = 1 - \alpha.$$

Therefore, the most powerful test is to reject the null hypothesis $H_0 : \lambda = \lambda_0$ in favor of the alternative hypothesis $H_A : \lambda = \lambda_1$ if $-2 \log \Lambda(X_1, X_2, \dots, X_n) \geq k$, where k is chosen such that $P(-2 \log \Lambda(X_1, X_2, \dots, X_n) \leq k | \lambda = \lambda_0) = 1 - \alpha$.

Note that the likelihood ratio test is not a uniformly most powerful (UMP) test in general. To determine whether the likelihood ratio test is a UMP test, we need to consider other possible test statistics and compare their power functions to the likelihood ratio test. If the likelihood ratio test has the highest power function among all possible test statistics, then it is a UMP test. However, this is a difficult and non-trivial task in general, and it is not always possible to find a UMP test.

In the solution I provided, the likelihood ratio test statistic is $\Lambda(x) = \frac{L(\lambda_1; x)}{L(\lambda_0; x)}$, where $L(\lambda; x)$ is the likelihood function for the parameter λ based on the observed data x .

For the Rayleigh distribution with scale parameter σ , the likelihood function for a random sample X_1, X_2, \dots, X_n is given by:

$$L(\sigma; X_1, X_2, \dots, X_n) = \prod_{i=1}^n \frac{X_i}{\sigma^2} e^{-X_i^2/(2\sigma^2)} = \frac{1}{\sigma^{2n}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n X_i^2\right) \prod_{i=1}^n X_i.$$

To test the hypotheses $H_0 : \lambda = \lambda_0$ and $H_A : \lambda = \lambda_1$, where $\lambda_1 > \lambda_0$, we can set $\lambda = \sigma^2$ and write the likelihood functions under the null and alternative hypotheses as:

$$L(\lambda_0; X_1, X_2, \dots, X_n) = \frac{1}{(\lambda_0)^{n/2}} \exp\left(-\frac{1}{2\lambda_0} \sum_{i=1}^n X_i^2\right) \prod_{i=1}^n X_i,$$

and

$$L(\lambda_1; X_1, X_2, \dots, X_n) = \frac{1}{(\lambda_1)^{n/2}} \exp\left(-\frac{1}{2\lambda_1} \sum_{i=1}^n X_i^2\right) \prod_{i=1}^n X_i,$$

respectively. Therefore, the likelihood ratio test statistic is:

$$\Lambda(X_1, X_2, \dots, X_n) = \frac{L(\lambda_1; X_1, X_2, \dots, X_n)}{L(\lambda_0; X_1, X_2, \dots, X_n)} = \left(\frac{\lambda_0}{\lambda_1}\right)^{n/2} \exp\left(\frac{1}{2} \left(\frac{\sum_{i=1}^n X_i^2}{\lambda_0} - \frac{\sum_{i=1}^n X_i^2}{\lambda_1}\right)\right).$$

We reject the null hypothesis $H_0 : \lambda = \lambda_0$ in favor of the alternative hypothesis $H_A : \lambda = \lambda_1$ if $\Lambda(X_1, X_2, \dots, X_n)$ is greater than a threshold value k . The threshold value k is chosen such that the probability of rejecting H_0 when H_A is true (i.e., the power of the test) is equal to the

significance level α . The threshold value k depends on the distribution of the likelihood ratio test statistic, which under certain conditions can be approximated by a chi-squared distribution with one degree of freedom.

In my solution, I applied the Neyman-Pearson lemma to find the most powerful test for the hypotheses $H_0 : \lambda = \lambda_0$ and $H_A : \lambda = \lambda_1$, where $\lambda_1 > \lambda_0$.

According to the Neyman-Pearson lemma, the most powerful test for a given significance level α is the likelihood ratio test. The likelihood ratio test compares the likelihood of the observed data under the null and alternative hypotheses, and the test statistic is the likelihood ratio:

$$\Lambda(X_1, X_2, \dots, X_n) = \frac{L(\lambda_1; X_1, X_2, \dots, X_n)}{L(\lambda_0; X_1, X_2, \dots, X_n)},$$

where $L(\lambda; X_1, X_2, \dots, X_n)$ is the likelihood function for the parameter λ based on the observed data X_1, X_2, \dots, X_n .

To apply the Neyman-Pearson lemma in my solution, I calculated the likelihood ratio test statistic for the Rayleigh distribution with scale parameter σ , which is given by:

$$\Lambda(X_1, X_2, \dots, X_n) = \left(\frac{\lambda_0}{\lambda_1}\right)^{n/2} \exp\left(\frac{1}{2} \left(\frac{\sum_{i=1}^n X_i^2}{\lambda_0} - \frac{\sum_{i=1}^n X_i^2}{\lambda_1}\right)\right).$$

Then, I showed that the likelihood ratio test is a monotone test, meaning that the likelihood ratio is an increasing function of the test statistic. This allows us to choose a threshold value k such that the test has a significance level of α and the power of the test is maximized. Specifically, we reject the null hypothesis $H_0 : \lambda = \lambda_0$ in favor of the alternative hypothesis $H_A : \lambda = \lambda_1$ if the likelihood ratio test statistic $\Lambda(X_1, X_2, \dots, X_n)$ is greater than or equal to the threshold value k .

Finally, I noted that the likelihood ratio test is not always a uniformly most powerful (UMP) test, as determining whether a test is UMP requires comparing it to all other possible tests.

Problem 3: page 328

$$f(x; k, \theta) = \frac{1}{\theta^k \Gamma(k)} x^{k-1} e^{-x/\theta}, \quad x > 0,$$

where $\Gamma(k)$ is the gamma function, which is defined as:

$$\Gamma(k) = \int_0^\infty t^{k-1} e^{-t} dt, \quad k > 0.$$

Part 1

To show that $\prod_{i=1}^n X_i$ is a sufficient statistic for the gamma distribution, we can use the factorization theorem. Specifically, we need to show that the joint probability density function (PDF) of the sample X_1, X_2, \dots, X_n can be written as the product of a function that depends only on $\prod_{i=1}^n X_i$ and a function that depends only on the parameters k and θ .

The joint PDF of X_1, X_2, \dots, X_n for a gamma distribution with parameters k and θ is given by:

$$f(x_1, x_2, \dots, x_n; k, \theta) = \frac{1}{\theta^{nk} \Gamma(k)^n} \prod_{i=1}^n x_i^{k-1} e^{-\frac{x_i}{\theta}}.$$

Taking the logarithm of this joint PDF, we obtain:

$$\log f(x_1, x_2, \dots, x_n; k, \theta) = -nk \log \theta - n \log \Gamma(k) + (k-1) \sum_{i=1}^n \log x_i - \sum_{i=1}^n \frac{x_i}{\theta}.$$

By the factorization theorem, we can write the joint PDF of X_1, X_2, \dots, X_n as:

$$f(x_1, x_2, \dots, x_n; k, \theta) = g\left(\prod_{i=1}^n x_i; k, \theta\right) h(x_1, x_2, \dots, x_n; k, \theta),$$

where $g(\prod_{i=1}^n x_i; k, \theta) = \exp((k-1) \log(\prod_{i=1}^n x_i) - nk \log \theta)$ depends only on $\prod_{i=1}^n X_i$ and the parameters k and θ , and $h(x_1, x_2, \dots, x_n; k, \theta) = \frac{1}{\Gamma(k)^n} \prod_{i=1}^n x_i^{k-1} e^{-\frac{x_i}{\theta}}$ depends only on x_1, x_2, \dots, x_n and the parameters k and θ .

Therefore, we can apply the factorization theorem and conclude that $\prod_{i=1}^n X_i$ is a sufficient statistic for the gamma distribution with parameters k and θ .

Part 2

To show that $\sum_{i=1}^n X_i$ is a sufficient statistic for the gamma distribution, we need to show that the joint probability density function (PDF) of the sample X_1, X_2, \dots, X_n can be written as the product of a function that depends only on $\sum_{i=1}^n X_i$ and a function that depends only on the parameters k and θ .

The joint PDF of X_1, X_2, \dots, X_n for a gamma distribution with parameters k and θ is given by:

$$f(x_1, x_2, \dots, x_n; k, \theta) = \frac{1}{\theta^{nk} \Gamma(k)^n} \prod_{i=1}^n x_i^{k-1} e^{-\frac{x_i}{\theta}}.$$

We can write the PDF of X_1, X_2, \dots, X_n in terms of $\sum_{i=1}^n X_i$ and X_n as follows:

$$\begin{aligned}
f(x_1, x_2, \dots, x_n; k, \theta) &= \frac{1}{\theta^{nk} \Gamma(k)^n} \prod_{i=1}^n x_i^{k-1} e^{-\frac{x_i}{\theta}} \\
&= \frac{1}{\theta^{nk} \Gamma(k)^n} \prod_{i=1}^{n-1} x_i^{k-1} e^{-\frac{x_i}{\theta}} \cdot x_n^{k-1} e^{-\frac{x_n}{\theta}} \\
&= \frac{1}{\theta^{(n-1)k} \Gamma(k)^{n-1}} \prod_{i=1}^{n-1} x_i^{k-1} e^{-\frac{x_i}{\theta}} \cdot \frac{x_n^{k-1} e^{-\frac{x_n}{\theta}}}{\theta^k \Gamma(k)} \\
&= \frac{1}{\theta^{(n-1)k} \Gamma(k)^{n-1}} \prod_{i=1}^{n-1} x_i^{k-1} e^{-\frac{x_i}{\theta}} \cdot \frac{(n-1)!}{\theta^k \Gamma(k) (n-1)!} \left(\frac{\sum_{i=1}^n x_i}{n} \right)^{k-1} e^{-\frac{\sum_{i=1}^n x_i}{n\theta}} \\
&= \frac{(n-1)!}{\theta^{nk} \Gamma(k)^{n-1} (n-1)!} \left(\frac{\sum_{i=1}^n x_i}{n} \right)^{k-1} e^{-\frac{\sum_{i=1}^n x_i}{n\theta}} \cdot \frac{1}{\theta^k \Gamma(k)} \prod_{i=1}^{n-1} x_i^{k-1} e^{-\frac{x_i}{\theta}} \\
&= \frac{(n-1)!}{\theta^{nk} \Gamma(k)^{n-1} (n-1)!} \left(\frac{\sum_{i=1}^n x_i}{n} \right)^{k-1} e^{-\frac{\sum_{i=1}^n x_i}{n\theta}} \cdot h(x_1, x_2, \dots, x_{n-1}; k, \theta)
\end{aligned}$$

where $h(x_1, x_2, \dots, x_{n-1}; k, \theta)$ is a function that depends only on x_1, x_2, \dots, x_{n-1} and the parameters k and θ .

Therefore, we have shown that the joint PDF of X_1, X_2, \dots, X_n can be written as the product of a function that depends only on $\sum_{i=1}^n X_i$ and a function that depends only on the parameters k and θ . Hence, $\sum_{i=1}^n X_i$ is a sufficient statistic for the gamma distribution with parameters k and θ .

Exponential Family

The gamma distribution with shape parameter $k = 2$ has probability density function:

$$f(x|\theta) = \frac{1}{\theta^2} \cdot \frac{x^{k-1}}{(k-1)!} e^{-x/\theta} = \frac{1}{\theta^2} \cdot \frac{x^1}{\Gamma(2)} e^{-x/\theta},$$

where $x > 0$ and $\theta > 0$.

To express this distribution in the form of an exponential family, we can rewrite the PDF in terms of the natural parameters and sufficient statistics. First, we can factorize the PDF as follows:

$$f(x|\theta) = \frac{1}{\theta^2} \cdot \frac{x^1}{\Gamma(2)} e^{-x/\theta} = \exp\left(-\frac{x}{\theta}\right) \cdot \frac{1}{\theta^2} \cdot x \cdot \frac{1}{\Gamma(2)},$$

where we have identified $w_1(\theta) = -1/\theta$, $w_2(\theta) = -1/\theta^2$, $t_1(x) = x$, $t_2(x) = 1$, $h(x) = 1$, and $A(\theta) = -\log(\Gamma(2)\theta^2)$.

Substituting these into the general form of an exponential family, we have:

$$f(x|\theta) = \exp\left(\sum_{i=1}^2 w_i(\theta) t_i(x) - A(\theta)\right) = \exp\left(-\frac{x}{\theta} - \frac{1}{\theta^2} + \log(\theta^2 \Gamma(2))\right) \cdot 1.$$

Therefore, we see that the gamma distribution with $k = 2$ belongs to the exponential family with natural parameters $w_1(\theta) = -1/\theta$ and $w_2(\theta) = -1/\theta^2$, sufficient statistics $t_1(x) = x$ and $t_2(x) = 1$, and log-partition function $A(\theta) = -\log(\theta^2\Gamma(2))$.

Sufficient statistic

To find a sufficient statistic for the gamma distribution with $k = 2$, we can use the factorization theorem. From the exponential family form of the gamma distribution with $k = 2$:

$$f(x|\theta) = \exp\left(-\frac{x}{\theta} - \frac{1}{\theta^2} + \log(\theta^2\Gamma(2))\right) \cdot 1,$$

we see that the sufficient statistic for θ must be of the form $\sum_{i=1}^n t_1(x_i)$, where $t_1(x) = x$ is the sufficient statistic corresponding to the natural parameter $w_1(\theta) = -1/\theta$.

Therefore, the sample sum $\sum_{i=1}^n X_i$ is a sufficient statistic for the gamma distribution with $k = 2$.