

Mathematical Statistics 2021/2022, Problem set 7

Estimator properties, part II

There is more material in this problem set than can be done during a single class. The material which is not solved during classes may serve as additional training material before the test.

1. Let us assume that the number of claims from a single yearly insurance policy follows a Poisson distribution with an unknown parameter θ , and let X_1, X_2, \dots, X_n denote the number of claims from independent policies of a given insurance company. We want to estimate the probability that there will be no claims from a policy.
 - (a) Let $\hat{g}_1 = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{X_i=0}$ be an estimator of the probability of no claims. Verify whether the estimator is unbiased and find the MSE of \hat{g}_1 . Find the asymptotic distribution of \hat{g}_1 and calculate the asymptotic efficiency.
 - (b) Find \hat{g}_{MLE} , the m.l.e. estimator of $e^{-\theta}$ (for the Poisson distribution, the probability of the variable being equal to 0 is equal to $e^{-\theta}$). Verify whether the estimator is unbiased and find the MSE (*Hint: if X_1, X_2, \dots, X_n are independent random variables from a Poisson distribution with parameter θ , then $\sum X_i \sim \text{Poiss}(n\theta)$. Hint2: $\sum_{k=0}^{\infty} \frac{x^k}{k!} = e^x$). Find the asymptotic distribution of \hat{g}_{MLE} and calculate the asymptotic efficiency.*
 - (c) Which of the two estimators would you use and why?
2. Let us assume that the distribution of genotypes in a population is multinomial, with probabilities $\theta^2, 2\theta(1-\theta)$ and $(1-\theta)^2$. Let n_1, n_2 and n_3 denote the population numbers for the three genotypes, respectively, in a population of size n . We want to estimate θ (the probability that a single gene will be of a dominant version), and we consider three estimators $\hat{\theta}_1 = \sqrt{\frac{n_1}{n}}$, $\hat{\theta}_2 = 1 - \sqrt{\frac{n_3}{n}}$, and the maximum likelihood estimator $\hat{\theta}_{MLE}$. Find $\hat{\theta}_{MLE}$. Compare the asymptotic efficiency of the three estimators.
3. Let X_1, \dots, X_n be a sample of independent observations from an exponential distribution with an unknown parameter $\lambda > 0$.
 - Find $\hat{\lambda}_{ML}$, the maximum likelihood estimator for λ .
 - Verify that $\hat{\lambda}_{ML}$ is biased, and propose $\hat{\lambda}_U$, an unbiased estimator on the base of $\hat{\lambda}_{ML}$.
Hint. If X_1, \dots, X_n are independent random variables from an exponential distribution with parameter λ , $Z = \sum_{i=1}^n X_i$ has a distribution with density $f_Z(x) = \frac{\lambda^n}{(n-1)!} x^{n-1} e^{-\lambda x}$ for $x > 0$, i.e. a gamma distribution.
Hint 2. $\int_0^{\infty} x^k e^{-\lambda x} dx = \frac{k!}{\lambda^{k+1}}$ for integer values of k .
 - Compare $\hat{\lambda}_{ML}$ and $\hat{\lambda}_U$ on the base of the MSE.
 - Calculate the efficiency of $\hat{\lambda}_U$. Is this estimator efficient?
 - Verify whether $\hat{\lambda}_{ML}$ and $\hat{\lambda}_U$ are consistent.
 - Is $\hat{\lambda}_{ML}$ asymptotically normal? If yes, is it asymptotically efficient?
4. Let X_1, \dots, X_n be a sample of independent observations from an exponential distribution with parameter $\frac{1}{\lambda}$, where $\lambda > 0$ is unknown. Find a such that the estimator $\hat{\lambda}_a = a \sum_{i=1}^n X_i$ has the smallest MSE. Is this estimator biased? Is this estimator consistent?

5. Let X_1, \dots, X_n be a sample of independent observations from a distribution with density equal to $f_\theta(x) = \frac{\theta}{x^{\theta+1}}$ for $x > 1$ and 0 otherwise, where $\theta > 2$ is an unknown parameter. Find the ML estimator of θ . Determine whether this estimator is: consistent? asymptotically normal? If yes, find the normal distribution that best resembles the distribution of the estimator for large n . *Hint. The expected value for a random variable with density f_θ is equal to $\frac{\theta}{\theta-1}$.*
6. Let X_1, \dots, X_n be a sample of independent observations from a geometric distribution with unknown parameter $\theta \in (0, 1)$, defined in the following way: $P(X = x) = \theta(1 - \theta)^x$ for $x = 0, 1, \dots$. Find the maximum likelihood estimator for θ and the method of moments estimator for θ (based on the mean) and compare the two estimators.

Hint. For a geometric distribution in this parametrisation, $EX = \frac{1-\theta}{\theta}$.