Statistics

List 2

Exercise 1.

Generate n observations from a binomial $b(5, p)$ distribution.

(a) $n = 50, p = 0.1$,

(b) $n = 50, p = 0.3$,

(c) $n = 50, p = 0.5$,

(d) $n = 50, p = 0.7$,

(e) $n = 50, p = 0.9$.

Calculate a value of the maximum likelihood estimator of the quantity $P(X \geq 3)$, where $X \sim b(5, p)$. Repeat the experiment 10 000 times. Estimate the variance, mean squared error (MSE), and bias of the estimator. Discuss the influence of the parameter p on the outcomes.

Exercise 2.

Generate n observations from a Poisson distribution with the parameter λ.

(a) $n = 50, \lambda = 0.5$,

(b) $n = 50, \lambda = 1$,

(c) $n = 50, \lambda = 2$,

(d) $n = 50, \lambda = 5$.

Calculate a value of the maximum likelihood estimator of the quantity $P(X = x)$, $x = 0, 1, \ldots, 10$, where $X \sim \pi(\lambda)$. Repeat the experiment 10 000 times. Estimate the variance, mean squared error (MSE), and bias of the estimator. Discuss the influence of the parameter λ on the outcomes.

Exercise 3.

Generate n observations from a beta distribution with the parameters θ and 1.

(a) $n = 50, \theta = 0.5$,

(b) $n = 50, \theta = 1$,

(c) $n = 50, \theta = 2$,

(d) $n = 50, \theta = 5$.

Repeat the experiment 10 000 times. Calculate a value of the maximum likelihood estimator, say $\widehat{I(\theta)}$, of the Fisher information for the parameter θ. Remember the score.

Generate, independently, n observations from a beta distribution with the parameters θ and 1. Calculate a value of the maximum likelihood estimator of the parameter θ. Define a new variable $Y = \sqrt{n\widehat{I(\theta)}}(\hat{\theta} - \theta)$. Compute its value. Repeat the experiment 10 000 times. Draw a histogram and Q-Q plot. Discuss a method of the selection of the number of classes in the histogram as well as a method of finding theoretical quantiles in the Q-Q plot. Is the distribution of Y normal? Justify the answer.

Exercise 4.

Generate n observations from a Laplace distribution with the parameters θ and σ.

(a) $n = 50, \theta = 1, \sigma = 1$,

(b) $n = 50, \theta = 4, \sigma = 1$,

(c) $n = 50, \theta = 1, \sigma = 2$.

Calculate a value of an estimator of the parameter θ of the form

(i) $\hat{\theta}_1 = \overline{X} = (1/n)\sum_{i=1}^{n} X_i$,

(ii) $\hat{\theta}_2 = Me\{X_1, \ldots, X_n\}$,

(iii) $\hat{\theta}_3 = \sum_{i=1}^{n} w_i X_i$, $\sum_{i=1}^{n} w_i = 1$, $0 \leq w_i \leq 1$, $i = 1, \ldots, n$, with an arbitrary weights' selection,

(iv) $\hat{\theta}_4 = \sum_{i=1}^{n} w_i X_{i:n}$, where $X_{1:n} \leq \cdots \leq X_{n:n}$ are the order statistics from the sample $X_1, \ldots, X_n$,

$$w_i = \varphi\Big(\Phi^{-1}(\frac{i-1}{n})\Big) - \varphi\Big(\Phi^{-1}(\frac{i}{n})\Big),$$

while φ is the density and Φ is the cumulative distribution function of the standard normal $N(0,1)$ distribution.

Repeat the experiment 10 000 times. Estimate the variance, mean squared error (MSE), and bias of the estimators under consideration. Discuss the outcomes. Which estimator is optimal and why? Confront the outcomes with the results of Exercise 1, List 1.

Exercise 5.

Repeat the numerical experiment from Exercises 1, 2, 3, and 4, for $n = 20$ and $n = 100$. Discuss the results in comparison to the previous outcomes.