Homework Assignment for Chapters 20 and 21 (Due by 3pm on Apr. 22)

Reference Exercise Problems: Text Book, 20.5 and 21.6 Exercises.

Homework problems

Problem 1 Given is a random sample X_1, X_2,..., X_n from a distribution with finite variance σ^2. We estimate the expectation of the distribution with the sample mean $\overline{X}_n$. Argue that the larger our sample, the more efficient our estimator. What is the relative efficiency $Var(\overline{X}_n)/Var(\overline{X}_{10n})$ of $\overline{X}_{10n}$ with respect to $\overline{X}_n$?

Problem 2 Leaves are divided into four different types: starchy-green, sugary-white, starchy-white, and sugary-green. According to genetic theory, the types occur with probabilities $\frac{1}{4}(\theta+2)$, $\frac{1}{4}\theta$, $\frac{1}{8}(1-\theta)$, and $\frac{3}{8}(1-\theta)$, respectively, where $0<\theta<1$. Suppose one has n leaves. Then the number of starchy-green leaves is modeled by a random variable N_1 with a $Bin(n, p_1)$ distribution, where $p_1 = \frac{1}{4}(\theta+2)$, and the number of sugary-white leaves is modeled by a random variable N_2 with a $Bin(n, p_2)$ distribution, where $p_2 = \frac{1}{8}\theta$. Consider the following two estimators for θ:

$$T_1 = \frac{4}{n}N_1 - 2 \text{ and } T_2 = \frac{8}{n}N_2.$$

In previous homework you showed that both T_1 and T_2 are unbiased estimators for θ. Which estimator would you prefer? Motivate your answer.

Problem 3 A certain type of plant can be divided into four types: starchy-green, starchy-white, sugary-green, and sugary-white. The following table lists the counts of the various types among 3839 leaves.

Type	Count
Starchy-green	1997
Sugary-white	32
Starchy-white	906
Sugary-green	904

Source: R.A. Fisher. Statistical methods for research workers. Hafner, New York, 1958; Table 62 on page 299.

Setting

$$X = \begin{cases} 1 & \text{if the observed leave is of type starchy-green} \\ 2 & \text{if the observed leave is of type sugary-white} \\ 3 & \text{if the observed leave is of type starchy-white} \\ 4 & \text{if the observed leave is of type sugary-green} \end{cases} \quad (1)$$

,

the probability mass function p of X is given by

a	1	2	3	4
$p(a)$	$\frac{1}{4}(\theta+2)$	$\frac{1}{4}\theta$	$\frac{1}{8}(1-\theta)$	$\frac{3}{8}(1-\theta)$

and $p(a) = 0$ for all other a. Here $0<\theta<1$ is an unknown parameter, which was estimated in previous homework. We want to find a maximum likelihood estimate of θ.

a. Use the data to find the likelihood $L(\theta)$ and the loglikelihood $\ell(\theta)$.

b. What is the maximum likelihood estimate of θ using the data from the preceding table?

c. Suppose that we have the counts of n different leaves: n_1 of type starchy-green, n_2 of type sugary-white, n_3 of type starchy-white, and n_4 of type sugary-green (so $n = n_1 + n_2 + n_3 + n_4$). Determine the general formula for the maximum likelihood estimate of θ.

Problem 4 Let x_1, x_2,..., x_n be a dataset that is a realization of a random sample from a $U(\alpha, \beta)$ distribution (with α and β unknown, $\alpha < \beta$). Determine the maximum likelihood estimates for α and β.

Problem 5 Let x_1, x_2,..., x_n be a dataset, which is a realization of a random sample from a $Par(\alpha)$ distribution. What is the maximum likelihood estimate for α?