

Homework Assignment for Chapters 17 and 19 (Due by 3pm on Apr. 15)

Reference Exercise Problems: Text Book, 17.6 and 19.6 Exercises.

Homework problems

Problem 1 During the Second World War, London was hit by numerous flying bombs. The following data are from an area in South London of 36 square kilometers. The area was divided into 576 squares with sides of length 1/4 kilometer. For each of the 576 squares the number of hits was recorded. In this way we obtain a dataset x_1, x_2, \dots, x_{576} , where x_i denotes the number of hits in the i th square. The data are summarized in the following table which lists the number of squares with no hits, 1 hit, 2 hits, etc.

Number of hits	0	1	2	3	4	5	6	7
Number of squares	220	211	93	35	7	8	1	1

Source: R.D. Clarke. An application of the Poisson distribution. Journal of the Institute of Actuaries, 72:48, 1946; Table 1 on page 481. Faculty and Institute of Actuaries.

An interesting question is whether London was hit in a completely random manner. In that case a Poisson distribution should fit the data.

1. If we model the dataset as the realization of a random sample from a Poisson distribution with parameter μ , then what would you choose as an estimate for μ ?
2. Check the fit with a Poisson distribution by comparing some of the observed relative frequencies of 0's, 1's, 2's, etc., with the corresponding probabilities for the Poisson distribution with μ estimated as in part a.

Problem 2 Recall Exercise 15.1 from the textbook about the chest circumference of 5732 Scottish soldiers, where we constructed the histogram displayed in Figure 1. The histogram suggests modeling the data as the realization of a random sample from a normal distribution.

- a. Suppose that for the dataset $\sum x_i = 71975$ and $\sum x_i^2 = 9125009$. What would you choose as estimates for the parameters μ and σ of the $N(\mu, \sigma^2)$ distribution?
Hint: you may want to use the relation from textbook Exercise 16.15.
- b. Give an estimate for the probability that a Scottish soldier has a chest circumference between 638.5 and 642.5 inches.

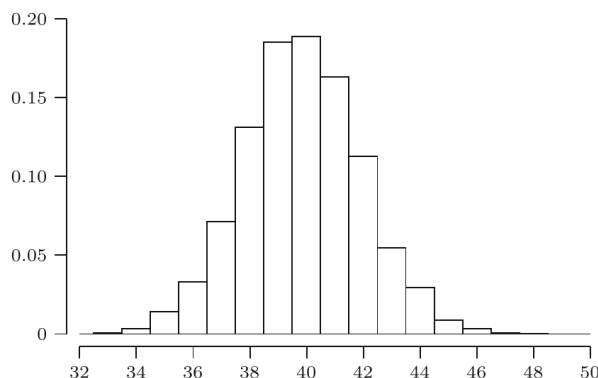


Figure 1: Histogram of chest circumferences. Note: Values on the X-axis should be added to 600.

Problem 3 Suppose the random variables X_1, X_2, \dots, X_n have the same expectation μ .

- Is $S = \frac{5}{8}X_1 + \frac{1}{4}X_2 + \frac{1}{8}X_3$ an unbiased estimator for μ ?
- Under what conditions on constants a_1, a_2, \dots, a_n is $T = a_1X_1 + a_2X_2 + \dots + a_nX_n$ an unbiased estimator for μ ?

Problem 4 Suppose a dataset is modeled as a realization of a random sample X_1, X_2, \dots, X_n from an $Exp(\lambda)$ distribution, where $\lambda > 0$ is unknown. Let μ denote the corresponding expectation and let M_n denote the minimum of X_1, X_2, \dots, X_n . Recall from textbook Exercise 8.18 that M_n has an $Exp(n\lambda)$ distribution. Find out for which constant c the estimator

$$T = \frac{1}{c}M_n$$

is an unbiased estimator for μ .

Problem 5 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 $\frac{1}{4}(\theta + 2)$, $\frac{1}{2}\theta$, $\frac{1}{4}(1 - \theta)$, and $\frac{1}{4}(1 - \theta)$, 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 $p_2 = \frac{1}{8}\theta$. The following table lists the counts for the progeny of self-fertilized heterozygotes 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.

Consider the following two estimators for θ :

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

- Check that both T_1 and T_2 are unbiased estimators for θ .
- Compute the value of both estimators for θ .