Math 10B with Professor Stankova
Worksheet, Discussion #21; Thursday, 4/11/2019
GSI name: Roy Zhao

Statistics and Estimators

Concepts

1. A **statistic** is a function of random variables and we use them to estimate values that we aren't always given. The **estimator of the mean** is

$$\hat{\mu} = \frac{x_1 + \cdots + x_n}{n}.$$

The **biased estimator of the standard deviation** is

$$s_*^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{\mu})^2.$$

The **unbiased estimator of the standard deviation**, also known as the **sample standard deviation** is

$$s^2 = \frac{n}{n-1}s_*^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \hat{\mu})^2.$$

The **95% confidence interval** of the *population* mean is

$$(\hat{\mu} - 2\frac{\hat{\sigma}}{\sqrt{n}}, \hat{\mu} + 2\frac{\hat{\sigma}}{\sqrt{n}}).$$

The PDF of a normal distribution with mean μ and standard deviation σ is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-(x-\mu)^2/(2\sigma^2)}.$$

Example

2. The number of rainy days in Honolulu in a year is Poisson distributed. Suppose that last ten years have had $4, 3, 2, 6, 5, 4, 1, 3, 4, 3$ rainy days. What is the 95% confidence interval for λ?

3. Now suppose that we did not know that the number of rainy days was Poisson distributed. What is the 95% confidence interval for the average number of rainy days per year?

Problems

4. True False If $f(x) = \frac{1}{a}e^{-(x-2019)^2/b}$ is the PDF of a normal distribution. Then $\pi = \frac{a^2}{b}$.
5. True False If we know both the biased estimator s_* and the unbiased estimator s, we can find out the sample size n.
6. True False For a geometric distribution, our estimate for the probability p is $\hat{p} = \frac{1}{\bar{x}+1}$.
7. True False The smaller the 95% confidence interval is, the lower our confidence is that the true parameter is in that interval.
8. True False The smaller the 95% confidence interval is, the higher our confidence is that the true parameter is in that interval.
9. True False The 95% confidence interval means that there is 95% chance that the parameter is in the interval.
10. True False Chebyshev's inequality says that 95% of the sample data much lie within 2 standard deviations of the mean.
11. I flip a biased coin 100 times and get 64 heads. What is the 95% confidence interval for p?
12. For the upcoming ASUC elections, you ask 400 people if they support the basic needs referendum, and 256 of them do. What is the 95% confidence interval for the percentage of all students who support the referendum?
13. Suppose I keep asking students if they are voting in the ASUC elections until I find someone who is. I do this multiple times and suppose that the number of students I have to ask before finding someone who is voting is $2, 2, 1, 8, 3, 5, 6, 3, 3, 7$. What is the 95% confidence interval for the average number of times we need to ask before we ask someone who is voting?
14. Do the previous problem if we use the sample standard deviation to calculate our confidence interval.
15. Every morning for 10 days I try to do the water bottle flip 25 times. I am successful $7, 8, 3, 4, 8, 5, 4, 5, 2, 4$ times. What is the 95% confidence interval for the number of times I am successful tomorrow morning?
16. A Pareto distribution is given by the PDF $f(x) = \frac{p}{x^{p+1}}$ for $x \geq 1$ and 0 for $x < 1$ for some parameter p. Suppose I draw from this distribution 4 times and get the values $1, 2, 1, 1, 1$. What is the 95% confidence interval for μ?
17. An exponential distribution is given by the PDF $f(x) = ce^{-cx}$ for $x \geq 0$ and 0 for $x < 0$. I draw from this distribution 5 times and get the values $\frac{1}{6}, 0, \frac{1}{3}, 1, \frac{1}{6}$. What is the 95% confidence interval for μ?