University of California, Los Angeles
Department of Statistics

Statistics 13 — Instructor: Nicolas Christou

Homework 5

EXERCISE 1
Recently there have been discussions about constructing a subway system that would run from Downtown Los Angeles to Santa Monica through Wilshire Boulevard. Suppose a random sample of 900 voters in Hollywood indicates that 600 support such an idea.

a. Construct a 95% confidence interval for the Hollywood population proportion of residents who would support this idea.

b. Suppose that the City of Los Angeles wants to estimate with 95% confidence the percentage of residents who would support this idea in Hollywood. The city wants the error of estimation to be $\pm 2\%$ of the population proportion. What is the minimum sample size required?

c. Suppose that the City of Los Angeles wants to estimate with 95% confidence the percentage of residents who would support this idea in Westwood. The city wants the error of estimation to be $\pm 2\%$ of the population proportion. What is the minimum sample size required? Assume that there is no prior information about the population proportion.

EXERCISE 2
The following 16 numbers came from a normal random number generator on a computer

5.3299	4.2537	3.1502	3.7032	1.6070	6.3923	3.1181	6.5941
3.5281	4.7433	0.1077	1.5977	5.4920	1.7220	4.1547	2.2799

a. What would you guess the mean and variance (μ and σ^2) of the generating normal distribution were?

b. Give 90%, 95%, and 99% confidence intervals for μ.

c. How much larger a sample do you think you would need to halve the length of the interval for μ?

EXERCISE 3
A sample of 9 instrument readings on the same object using a precision instrument yielded the measurements:
5.2, 4.4, 4.8, 3.6, 3.1, 6.2, 5.8, 5.1, and 3.2.

a. Construct a 99% confidence interval for the population mean.

b. What assumption do you have to make about the population from where the 9 readings were selected.

EXERCISE 4
As part of a study of the development of the thymus gland, researchers weighed the glands of five chicks embryos after 14 days of incubation. The thymus weights (mg) were as follows: 29.6, 21.5, 28.0, 34.6, 44.9.
For these data, $\bar{x} = 31.7, s = 8.7$.

a. Construct a 90% confidence interval for the population mean μ.

b. True or false and why: We are 90% confident that all the thymus weights are included in the interval you constructed in part (a).

EXERCISE 5
The output voltage of a certain electric circuit is specified to be 130 volts. The population standard deviation is known to be $\sigma = 3.0$ volts. A sample of 40 readings on the voltage of this circuit gave a sample mean of 128.6 volts.

a. Construct a 95% confidence interval for the true voltage mean. Based on your confidence interval do you believe that the mean voltage is 130 volts?

b. What is the probability that the confidence interval that you constructed in part (a) contains the true voltage mean? Explain.

EXERCISE 6
A manufacturer of batteries claims that the lifetime of these batteries is normally distributed. The manufacturer provides us with the standard deviation $\sigma = 10$ hours. For a random sample of 50 batteries the sample average was found to be $\overline{x} = 115$ hours.

a. Construct a 95% confidence interval for the population mean μ.

b. How will you interpret the confidence interval in (a) for a friend of yours who has not taken statistics yet?!

EXERCISE 7
The manager of a supermarket would like to know the average time that a person spends at the checkout counter. Using a stopwatch, he observes 100 customers. He computed the sample mean to be $\overline{x} = 15.35$ minutes and the sample standard deviation to be $s = 6.1$ minutes.

a. Construct a 95% confidence interval for the population mean μ.

b. Suppose that the manager wants a smaller error in estimation (smaller than what you found in (a)). He wants his error to be ± 1 minute with 95% confidence. How many customers will he need? For this question assume $\sigma = 6.1$ minutes.

EXERCISE 8
Use `R` to access the data form the Maas river. These data contains the concentration of lead and zinc in ppm at 155 locations at a flooding area of the Maas river in the Netherlands.

```
a <- read.table("http://www.stat.ucla.edu/~nchristo/statistics13/soil.txt", header=TRUE)
```

a. Use `R` to compute the sample mean and sample standard deviation of lead.

b. Construct a 95% confidence interval for the population mean of lead in this area.

c. The level of risk for surface soil based on lead concentration in ppm is given on the table below:

Mean concentration (ppm)	Level of risk
Below 150	Lead-free
Between 150-400	Lead-safe
Above 400	Significant environmental lead hazard

Based on your confidence interval form (b) in which category does the soil of this area fall in terms of the ppm concentration?

d. Do you see any problem in these calculations (meaning by just using the averages)?