Problem Set 8

Statistics - NYU, Summer 2016

Ercan Karadas

Section 1

[1] A company produces breakfast cereal. The true mean weight of the contents of its cereal boxes is 20 ounces, and the standard deviation is 0.6 ounce. The population distribution of weights is normal. Suppose that you purchase four boxes, which can be regarded as a random sample of all those produced.

a) What is the standard error of the sample mean weight, and what is the distribution of sample mean weight?

b) What is the probability that, on average, the contents of these four boxes will weight less than 19.7 ounces?

c) What is the probability that, on average, the contents of these four boxes will weight more than 20.6 ounces?

d) What is the probability that, on average, the contents of these four boxes will weight between 19.7 and 20.5 ounces?

e) Two of the four boxes are chosen at random. What is the probability that the average contents of these two boxes will weight between 19.5 and 20.5 ounces?

[2] The times spent studying by students in the week before final exams follows a normal distribution with standard deviation 8 hours. A random sample of four students was taken in order to estimate the mean study time for the population of all students.

a) What is the probability that the sample mean exceeds the population mean by more than 2 hours?

b) What is the probability that the sample mean is more than 3 hours below the population mean?

c) What is the probability that the sample mean differs from the population mean by more than 4 hours?

d) How large a sample is necessary in order to ensure that the probability that the sample mean differs from the population mean by more than 1.00 is less than 0.10?

[3] A cigarette manufacturer came up with a new brand of cigarettes called Long Life! The nicotine content of the cigarettes follows a normal distribution with a mean of 20 and a standard deviation of 5. A consumer bought a pack of Long Life that contains 25 cigarettes. Consider these 25 cigarettes as a random sample.

a) What is the probability that a cigarette contains over 23 units of nicotine?

b) What is the probability that the average nicotine content for the whole pack of cigarettes is higher than 23?

[4] Suppose that a certain delicate new medical device will fail between 0 and 10 years after it is implanted in the human body. The following table depicts failing time of the device on 1000 patients.

Life of the device	1	2	3	4	5	6	7	8	9	10
Frequency	98	99	102	101	100	97	101	103	99	100

a) Which probability distribution model fits best to the data? In the rest of the problem stick to this probability distribution model. Also, explain if you need to make additional assumptions to work with that probability model.

b) Verify that the mean time to failure is 5 years, and the standard deviation is 2.88 years.

c) Find the probability that an individual device will fail more than 80 months after implantation.

d) Find the probability that in a sample of 36 of these devices, the sample mean time of failure will be 80 months or less.

[5] A random variable B measures the average daily balances in customers' savings accounts which is normally distributed, with a mean of \$108 and a standard deviation of \$15.

a) If a random sample of size 4 is drawn, what is the probability that the B value exceeds \$116?

b) If a random sample of size 16 is drawn, what is the probability that the B value exceeds \$116?

c) What happened to the standard deviation of B when the sample size increased from $n = 4$ to $n = 16$? Explain.

d) What happened to the probability of observing B =\$116 as the sample size increased from $n = 4$ to $n = 16$? Explain.

[6] An advertising executive believes that the length of time that a television viewer can recall a commercial is distributed exponentially with a mean of .25 days.

a) Find the proportion of viewers who will be able to recall the commercial after 7 days.

b) Find how long it will take for 75% of the viewing audience to forget the commercial.

c) Suppose that we have a group of 36 people that are randomly chosen among the television viewers and we are interested in the average recall length. Describe the distribution of average recall length. In particular, discuss whether we can compute the mean and the variance of the average recall length. What about do we the whole distribution, i.e. the whole shape of the distribution?

d) Compute the probability that the average recall length among the group will be at least 5/16 days.

[7] Each member of a random sample of 15 business economists were asked to predict the rate of inflation for the next coming year. Assume that predictions of the whole population of business economists follow a normal distribution with standard deviation 1.8%.

a) The probability is 0.01 that the sample standard deviation is bigger than what number?

b) The probability is 0.025 that the sample standard deviation is smaller than what number?

c) Find a pair of numbers (say a and b) such that the probability that the sample standard deviation lies between these two numbers is 0.90.

Section 2

[8] A college admissions officer for an MBA program has determined that historically applicants have undergraduate grade point averages that are normally distributed with standard deviation 0.45. From a random sample of 25 applications from the current year, the sample mean grade point average is 2.90.

a) Find a 95% confidence interval for the population mean?

b) Without doing calculations, explain whether a 99% confidence interval for the population mean would be wider than, narrower than that of found in part (a).

c) Suppose that population standard deviation is 0.56 (instead of 0.45). Without doing calculations explain whether a 95% confidence interval for the population mean would be wider than, narrower than, or the same width as found in part (a).

d) Suppose that sample mean that is given in the problem is calculated from a sample of size 40 (instead of 25). Without doing calculations explain whether a 95% confidence interval for the population mean would be wider than, narrower than, or the same width as found in part (a).

e) Based on these sample results, a statistician computes for the population mean a confidence interval extending from 2.81 to 2.99. Find the confidence level associated with this interval.

[9] Let $\bar{x}$ be the observed mean of a random sample of size n from a normal distribution with mean μ and variance σ^2.

a) Find n such that from $\bar{x} - \sigma/4$ to $\bar{x} + \sigma/4$ is a 95% confidence interval for μ.

b) Generalize the result in the previous part: find n such that from $\bar{x} - k\sigma$ to $\bar{x} + k\sigma$ is a 95% confidence interval for μ (Note that indeed this question is more general because if we take $k = 1/4$, it reduces to part (a)).

[10] The life of a particular type of light bulb is normally distributed with mean 1000 and standard deviation of 150 hours. The manufacturer is considering a warranty.

a) What proportion of the bulbs have a life of more than 1150 hours?

b) What proportion of the bulbs have a life between 925 and 1150 hours?

c) If the manufacturer wants to replace no more than 5% of the bulbs under warranty, how many hours should the warranty cover?

[11] Sloan and Lorant (1977) studied the relationship between the length of time patients wait in a physicians office and certain demand and cost factors. They obtained data on typical patient waiting times for 4,500 physicians and reported a mean waiting time of 24.7 min and a standard deviation of 19.3 min.

Suppose a pediatrician does not have this set of data and has one of the nurses in the office monitor the waiting times for 64 randomly selected patients during the year.

a) What is the probability that the sample mean falling between 18 and 26 min?

b) Would your answer in part (a) change if Sloan and Lorant (1977) further showed that the date was exponentially distributed? Explain.

[12] A company claims that its accounts receivable follow a normal distribution with a mean of \$500 and a standard deviation of \$75. An auditor will certify the bank's claim only if the mean of a random sample of 50 accounts lies within \$25 of the mean. Assume that the bank has accurately reported its mean accounts receivable. What is the probability that the auditor will certify the bank's claim?

[13] The newly produced 1992-Honda boasts 45 miles per gallon on the highway. Assume that the distribution of the miles per gallon is a normal distribution with a mean of 40 and a standard deviation of 5. The Environmental Protection Agency randomly draws 100 1992-Hondas to test-drive.

a) What is the probability that a certain car can achieve 45 miles per gallon?

b) What is the probability that the average of 100 cars exceeds 45 miles per gallon?

c) What is the probability that of the 100 cars test-driven, more than 35 cars get more than 45 miles per gallon? How many of the 100 cars tested would you expect to get more than 45 miles per gallon?