Exercises: Statistical Inference for One Population

A: Estimation

1 In order to estimate the mean late departure time of commercial airline flights of duration under two hours, between cities in Australia, a survey of a random sample of 77 such flights was conducted over a six-month period last year. It was found that for this sample, the mean late departure time was 2.48 minutes. Based on a previous study, the population standard deviation of the late departure time was assumed to be 12 minutes.

a) What is the point estimate of the mean late departure time of all commercial flights of duration under two hours, between cities in Australia.

b) Compute a 95% confidence interval estimate of the mean late departure time of all commercial flights of duration under three hours, between cities in Australia. Report your result.

2 The marketing director of a large department store wants to estimate the mean number of customers who enter the store every 15 minutes on a Saturday which is usually the busiest shopping day. She randomly selects ten 15-minute intervals on a particular Saturday between 9am and 9pm and counts the number of arrivals at the store. She obtains the following numbers: 58, 32, 41, 47, 56, 80, 45, 29, 32 and 78. The analyst assumes the number of arrivals in 15-minute intervals is normally distributed.

a) What is the point estimate of the mean number of customers who enter the store every 15 minutes?

b) Compute a 95% confidence interval. Report your result.

3 A road authority in a large city wants to estimate the proportion of vehicles on a major highway between the hours of midnight and 5am, that are semi-trailers. The estimate will be used to determine highway repair and construction considerations and in highway patrol planning. Suppose researchers counted vehicles at different locations on the highway for several nights during this time period. Of the 3481 vehicles counted, 927 were semi-trailers.

a) Determine a point estimate for the proportion of vehicles travelling on this major highway during this time period, that are semi-trailers.

b) Compute a 99% confidence interval for the proportion of vehicles travelling on this major highway during this time period, that are semi-trailers. Report your result.

4 A study released showed that *Universal Music Victoria* held a 24.5% share of the online music market. Suppose this figure is actually a point estimate obtained by surveying a random sample of 1,003 people who purchased music online.

a) Use this information to compute a 99% confidence interval for the proportion of the online music market that is held by *Universal Music Victoria.* Report your result.

b) Suppose the figure was obtained from a survey of 10,000 people. Recompute the confidence interval and compare your results with the first confidence interval. How did they differ? What might you conclude about the sample size and confidence intervals?

5 A consumer's organisation is interested in determining the variability in the price of a particular brand of smart phone. A survey of 25 randomly selected retail outlets produced a mean and standard deviation of \$800 and \$20 respectively.

a) Compute a 95% confidence interval estimate of the standard deviation in the price of this brand of smart phone. Report your result.

b) What assumption did you have to make before obtaining the confidence interval?

6 C batteries have a nominal diameter of 26.2mm. Producing batteries with this diameter will ensure they fit properly into all toys and appliances that require C batteries. A tolerance variance of 0.1560 mm^2 is allowable. A random sample of 12 C batteries produced by one of the machines has the following diameters in millimetres.

26.22 29.19 26.24 26.21 26.16 26.18
26.21 26.20 26.22 26.91 26.20 26.18

a) Construct a 95% confidence interval for the variance of these diameters. Report your result.

b) Does this confidence interval estimate of the variance satisfy the tolerance variance?

c) What assumption did you have to make before obtaining the confidence interval?

B: Sample Size Determination

1 A group of investors wants to develop a chain of fast-food restaurants in a large. In determining potential costs for each facility, they must consider, among other expenses, the mean monthly electricity bill. They decide to randomly sample some fast-food restaurants currently operating to estimate the monthly cost of electricity. They want to be 90% confident of their results and want the error of the interval estimate to be no more than \$100. They estimate that such bills range from \$600 to \$2500. How large a sample should they take?

2 Suppose a production facility purchases a particular component in large lots from a supplier. The production manager wants to estimate the proportion of defective parts received from the supplier. She believes the defective proportion is no more than 0.20 and she wants to be within 0.02 of the true proportion of defective parts with a 90% level of confidence. How large a sample should she take?

3 A consumer's organisation is interested in determining the variability in the price of a particular brand of smart phone and a survey of randomly selected retail outlets is going to be conducted. Suppose that an estimate of the standard deviation in the price of this brand of smart phone is required to be within 0.15 of the true standard deviation, with 99% confidence. Determine an approximate number of retail outlets that should be selected.

C: Hypothesis Testing

1) A manufacturing company has been averaging 18.2 orders per week for several years. However, during a recession, orders appeared to slow. Suppose the company's production manager randomly samples 32 weeks and finds a sample mean of 15.6 orders. The population standard deviation is assumed to be 2.3 orders. Using the 10% level of significance, test whether or not, the average number of orders is down.

2) A soft drink company wants to know if the mean fill of its 2-litre bottles is either significantly different from 2 litres. It is assumed that the amount of soft drink is normally distributed. The following data represent the amount of soft drink in a random sample of 50 2-litre bottles.

2.109	2.086	2.066	2.075	2.065	2.057	2.052	2.044	2.036	2.038
2.031	2.029	2.025	2.029	2.023	2.020	2.015	2.014	2.013	2.014
2.012	2.012	2.012	2.010	2.005	2.003	1.999	1.996	1.997	1.992
1.994	1.986	1.984	1.981	1.973	1.975	1.971	1.969	1.966	1.967
1.963	1.957	1.951	1.951	1.947	1.941	1.941	1.938	1.908	1.894

These data are in the file 1hyp_exercise2.csv.

a) Is there evidence that the mean amount of soft drink filled is different from 2.0 litres? Perform the relevant test at the 5% level of significance.

b) Obtain a 95% confidence interval estimate of the amounts of soft drink filled and report the confidence interval.

c) Check for outliers in the data and test whether or not the assumption of normality of the amount of soft drink is valid.

3) A genetic model suggests that 80% of plants grown from a cross between two strains of seeds will be of the dwarf variety. In a random sample of 200 of such plants, 64 were found not to be of the dwarf variety. Is the evidence that the observations contradict the genetic model? Perform the relevant test at the 5% level of significance.

4) A manufacturing company produces bearings. One line of bearings is specified to be 1.64cm in diameter. A major customer requires that the variance of the bearings be no more than 0.001 cm^2. The producer is required to test the bearings before they are shipped, and so the diameters of 16 randomly selected bearings are measured with a precise instrument, resulting in the following values. Assume that the bearing diameters are normally distributed.

1.69	1.62	1.63	1.70	1.66	1.63	1.65	1.71
1.64	1.69	1.57	1.64	1.59	1.66	1.63	1.65

These data are in the file 1hyp_exercise4.csv.

a) Using the 5% level of significance, test if the variance of the bearing diameter is too high.

b) Check for outliers in the data and check if the assumption of normality of the bearing diameters is valid.