STA 2023
Single Sample Inference

5.1 and 5.2 Large Sample Interval Estimation for the Mean

The simplest form of estimation is a point estimate, where a single number (statistic) is calculated from a sample and is used to estimate a population parameter.

e.g. $\bar{X}$ estimates the population mean, μ

s^2 estimates the population variance $\sigma 2$.

In general, T estimates, or is a point estimate of θ.

CONFIDENCE INTERVALS

Note that both and are continuous random variables and for continuous data, we have that P(T = θ) = 0. As an alternative, we can use interval estimation, where we estimate a parameter by an interval generated from the sample data. These intervals can be constructed so that we have a known confidence (probability) that the parameter is "captured" or surrounded by the interval. This probability is called the confidence coefficient and denoted by (1 - α). The resulting interval is called a (1 - α)100% confidence intervals or a (1 - α)100% C.I.

Confidence Interval for μ - large sample or σ known.

Let us start with an example. Suppose you work for a Honda manufacturing plant and buy tires from Michelin. You are interested in the lifetime of a particular brand of tires. To estimate this lifetime, you take a random sample of n = 44 tires and find that $\bar{x}$ = 40,321 miles and s = 2,144 miles. What sort of a conclusion can we reach about the true mean lifetime μ ?

Note that $\bar{x} \sim N(\mu, \sigma^2/n)$ if n is sufficiently large or if the underlying population is normal.

Hence here is what we would do if we wanted a 95% confidence interval for μ = the mean lifetime.

Therefore, a (1-a)100% confidence interval for μ when σ is known is :

$$(\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}})$$

Note that this can be written as

$$\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Note that $z_{\alpha/2}$ is often called the reliability coefficient, and $\frac{\sigma}{\sqrt{n}}$ is the standard error. We can look up $z_{\alpha/2}$ for some α values on page 295 or at the bottom of Table V, page 885

If n is large enough and σ is unknown, substitute with s.

*** Whenever the estimate follows a normal (or approximately normal) distribution, the confidence interval will be of the form:

estimate ± [(reliability coefficient) (standard error)]

Now let us go back to the Honda example and find 90%, 95% and 99% confidence intervals for the mean lifetime of the tires.

90% C.I.:

$1 - \alpha = 0.9$, α = ____________, $\alpha/2$ = ______________.

Thus the interval is:

___________________ ± ___________________

= ___________________ ± ___________________

= (,)

Similarly,

95% C.I.

99% C.I.:

What is happening to the intervals as the confidence level increases?

5.4 Estimating the Parameter of a Binomial Proportion

Let us look at the sampling distribution of . $\hat{p}$ ....

It turns out for large n

$$\hat{p} \sim N(p, p(1-p)/n)$$

Hence using the same reasoning as in the previous sections, a (1-α)100% C.I. for p is given by

*This is a large sample confidence interval *

e.g. A survey of 1,422 shoppers indicated that 727 of these prefer Gain detergent to the rival brand.

a) What is the point estimate of p = true proportion of all shoppers that prefer Gain

b) Construct a 99% C.I. for p.