Statistics 251: Lab 7 Pre-Reading Document 1

Inference (Part 2) - Confidence Intervals and Hypothesis Testing

Objectives:

- **Understand**
  1. **Confidence Intervals** (**Required Reading**)
     - **Known Variance** - Z Test
     - **Unknown Variance** - t Test
  2. Hypothesis Testing (Optional Reading)
     - Types of hypotheses
     - Use of confidence intervals in hypothesis testing
  3. Example (Optional Reading)
     - Use of confidence intervals in hypothesis testing
     - R commands
- **Prepare for lab 7 by trying a pre-lab exercise**

Introduction:

A **confidence interval** (CI) is an estimate of a range of values that is likely to include the true population parameter of interest, as opposed to a single estimate, i.e. a **point estimate**.

Please refer to **section 1 of pre-reading document 2** (pages 1-3) for a detailed description of confidence intervals. Section 1 is required reading to successfully complete the lab exercises. Continue with the rest of this document after you have finished reading **section 1 of pre-reading document 2**.

Interpretation of a 95% confidence interval for μ:

Over the collection of all 95% confidence intervals that could be constructed from repeated random samples of size n obtained from the population, 95% of these CIs will contain the true value of the population mean μ.

1.0 – 3.0: See pre-reading document 2

4.0 Pre-Lab Exercise

4.1 Review of Formulae for Confidence Intervals

<u>Unknown variance</u>

$\bar{x} \pm t_{(n-1)}(\alpha).\frac{s}{\sqrt{n}}$

<u>Known variance</u>

$\bar{x} \pm z(\alpha).\frac{\sigma}{\sqrt{n}}$

For each of these confidence intervals, we require either that the data is Normal to obtain an exact interval, or we can employ the CLT if the sample size is 'large' ($n \geq 20$).

In this lab, you will attempt to generate **confidence intervals for the mean** μ, under different scenarios.

4.2 A Simulation

Suppose you are interested in estimating the mean breaking strength of a material, or the reaction time of a chemical, or lifetime of a battery. You can use a confidence interval, i.e. a range of values that is likely to include the true population parameter of interest.

Let's do a simulation to visualize what percentage of confidence intervals will actually contain the true parameter μ by drawing 10,000 samples of size 16 from a Normal distribution with mean 10 and variance 9.

1. Simulate a sample of <u>size 16</u> from a $N(\mu = 10, \sigma^2 = 9)$ distribution.
   a) Compute your 95% confidence interval for the mean, assuming $\sigma^2 = 9$ is known.
   b) Compute your 95% confidence interval for the mean, assuming σ^2 is unknown, and estimate it with $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$.
   c) For each of the two previous problems, repeat the process 10000 times, and count how many times your confidence intervals do not contain the true mean μ.

4.2.1 Pre-simulation questions:

A) How many (or what percentage) of the CIs do you expect would not contain μ

a. when variance is known?
b. when variance is unknown?

```
N <- 10000 # number of iteration
n <- 16 # sample size
m <- 10 # mean
s <- sqrt(9) # SD
alpha <- 0.05 # (1-confidence level)
```

```
un <- nw <- matrix(NA, nrow =N, ncol =2) # 2 blank matrices
evaluate <- evaluate.true <- rep(FALSE, N) # 2 blank vectors

in.CI <- function(x){ (x[1]<=m & m<=x[2])}
# Define a function called in.CI. The input x is a 2-element vector,
representing an interval. If m is within the interval, in.CI return TRUE,
otherwise returns FALSE.

for (i in 1:N){ # loop starts

 Sample <- rnorm(n, m, s) # generate normal variates with given parameters

 un[i,] <- c(mean(Sample) - (-1)*qt(alpha/2, df = n - 1)*sd(Sample)/sqrt(n),
  mean(Sample) + (-1)*qt(alpha/2, df = n - 1)*sd(Sample)/sqrt(n) )
 # Calculate the i-th confidence interval for estimated SD

 nw[i,] <- c(mean(Sample) - (-1)* qnorm(alpha/2, 0, 1)*s/sqrt(n),
  mean(Sample) + (-1)* qnorm(alpha/2, 0, 1)*s/sqrt(n) )
 # Calculate the i-th confidence interval for known SD

 evaluate[i] <- in.CI(un[i,])
 # m is contained in 1st CI when SD unknown?
 evaluate.true[i] <- in.CI(nw[i,])
 # m is contained in 2nd CI when SD known?

 cat(i, ": Intervals (", round(un[i,], 2), ")", c("*",
"")[as.numeric(evaluate[i]) + 1 ],
  "and (", round(nw[i,],2), ")", c("*", "")[as.numeric(evaluate.true[i]) +
1], "\n") # print the results for i-th CIs
}

sum(evaluate == FALSE)/N # count and make a ratio
sum(evaluate.true == FALSE)/N # count and make a ratio
```

4.2.2 Post-simulation questions:

A) Do the results of the simulation match up with what you expected?

B) Are the results the same for the known variance and the unknown variance?

Note: results of different simulations may vary since we are dealing with random components, but overall conclusion should be the same.

To visualize this, let us plot the first 100 confidence intervals for the unknown variance case.

```
testing.n = 100
plot(5,6, ylim= c(1,testing.n), xlim=c(5,15), type = "n", axes= F, xlab = "",
ylab = "") # making a blank graph
for (i in 1:testing.n){
 lines(c(un[i,1],un[i,2]),c(i,i), col = c("blue", "green")[evaluate[i]+1])
} # putting each confidence intervals one at a time
abline(v = m, col = "red", lwd = 2) # setting parameter
```

This will provide a graph similar to figure 1, where the red line is the true parameter and green lines are the confidence intervals that contain the true parameter. Blue lines are those confidence intervals that do NOT contain the true parameter.

Notice that out of 100 confidence intervals considered here, 6 do not contain the true parameter (blue). How many of your confidence intervals do NOT contain the true parameter?

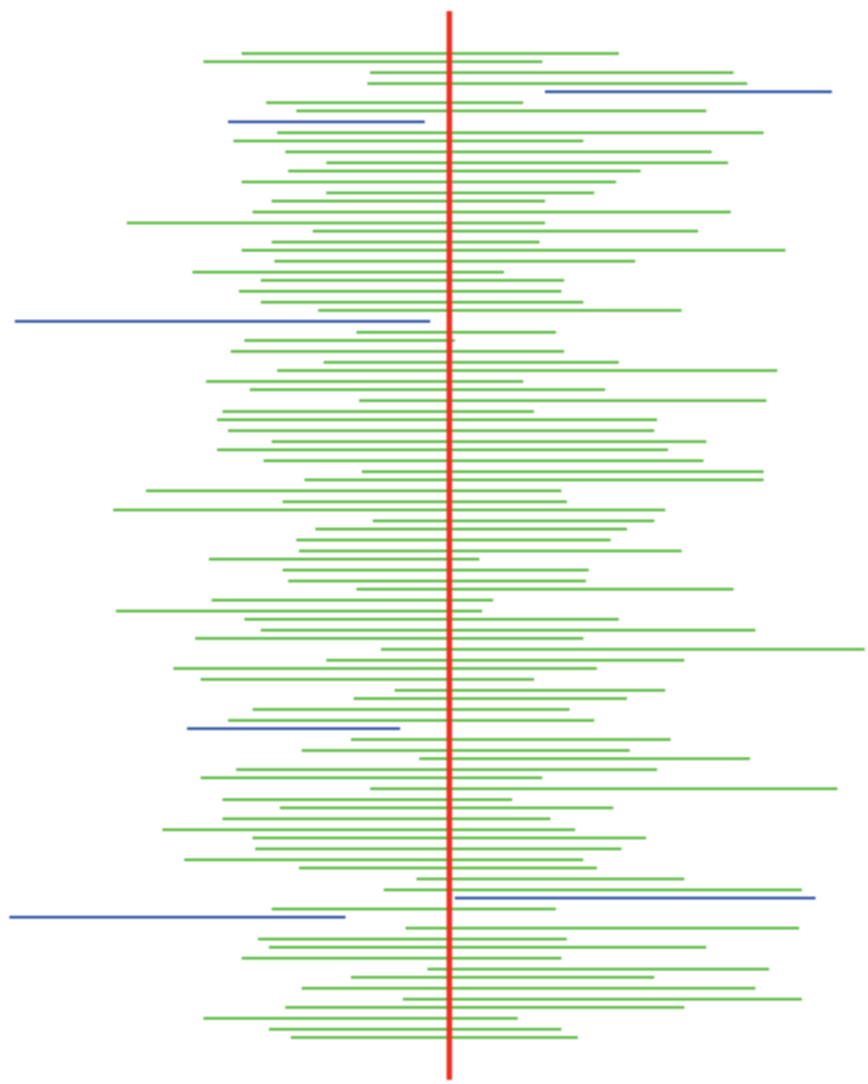

Figure 1: One hundred 95% confidence intervals.
Red line: True parameter
Green lines: Confidence intervals that contain the parameter
Blue lines: Confidence intervals that do NOT contain the parameter

4.3 Food for Thought

1. Do you expect the results of when variance is known to be the same as when variance is unknown if we used samples of size 100?

2. Do you expect the results to be the same if we drew samples from a different distribution, such as a Uniform or Exponential distribution,
    a. when sample size is 16?
    b. when sample size is 100?