**Confidence Intervals (of One Mean) & the Students’ t (t Distribution):**

**M&Ms’ Weight Exercises**

**As a result of completing this exercise you should be able to:**

* Understand the concept of confidence intervals (CI)
* Understand margin of errors (ME) and standard errors (se)
* Construct confidence intervals for one mean
* Distinguish between z *distributions (the Normal Model)* and t *distributions (Students’ T)*
* Understand the degrees of freedom (df)
* Identify the critical value t\* using qt in R

**Instructions**

1. Get the data ready
   1. Load the latest M&M’s bag data (“ISAT251\_MMsBags.csv”) from the Canvas.
   2. Bring the data into R

> myfile <- file.choose()

> mms.bags <- read.csv(myfile,header=TRUE)

*Type or paste the full path to your file here:*

* 1. Check and overview the data. (*Snapshot and paste your codes and the summary outcome below*)

> summary(mms.bags)

A close up of text on a white background

Description automatically generated

**Section I: Constructing CIs for one mean**

1. Find the information (i.e., the mean, sample size, and the standard deviation) you need to construct confidence intervals and critical values using proper R functions (i.e., mean, sd, and dim). Type or write down those statistics with *both the* ***symbols*** *and* ***values****, for example, if the sample size is 10,* ***n*** *=* ***10***, below.

## sample size as the row number in the dimension

> n <- dim(mms.bags)[1]

mean (use R to get the value): = 49.41953

standard deviation (use R to get the value): = 1.615026

sample size: = 64 6

1. Are the assumptions or conditions for this problem true so that the data meet the requirements to construct a confidence interval? Explain why those assumptions can be true and use a plot to check whether the distribution seems nearly normal.

Independent or randomness assumption: The confidence interval of the mean is estimated that the statistic follows a normal distribution. So the variance is independent of the mean.

Sample size condition: Sample size is not large enough to display a normal distribution. The distribution seems nearly normal.

Plot a histogram to display the distribution of the weight of M&M’s, does it look like a normal (in which case, the sample size can be smaller)? (Paste the snapshot of *your codes and plot below.*) 

Chart, histogram

Description automatically generated

1. As all the assumptions are checked, we now start to construct the confidence interval of mean (the average weight of M&M’s). First, what is the standard error of mean of the sample your class got?

standard error (se) = = 0.2018783

1. What is the critical value of t\* for confidence intervals of 90%, 95% and 99% for this problem? (Use R, the qt function to get the actual value for t\*). *Type your answers below and paste* the snapshot of *your R codes.*

# degrees of freedom

> mms.df <- n-1

# t\*

> t\_90 <- qt(p=.95 , df =mms.df)

> t\_95 <- qt(p=.975 , df =mms.df)

> t\_99 <- qt(p=.995 , df =mms.df)

Degrees of freedom (df) = sample size – 1 = n -1 = 63

t\* for 90% CI: 1.669402

t\* for 95% CI: 1.998341

t\* for 99% CI: 2.656145

Text

Description automatically generated

1. Please use the critical values (t\*) you get in question 3 and question 4 to calculate the margin of error (ME) for the three confidence intervals (90%, 95%, 99%):

ME for 90%: t\_90\*mms.se = 0.337016

ME for 95%: t\_95\*mms.se = 0.4034215

ME for 99 %: t\_99\*mms.se = 0.5362179

1. Now, use the values of margin of error (ME) you get in question 5 to construct those confidence intervals (CI):

CI for 90%: (49.082514, 49.756546)

49.41953

CI for 95%: (49.0161085, 49.8229515)

CI for 99 %: (48.8833121, 49.9557479)

1. Interpret the *99% confidence interval* verbally in words.

Hint: ”I am \_\_\_\_\_% confidence that \_\_\_\_\_\_\_\_\_ of \_\_\_\_\_\_\_\_\_ is between \_\_\_\_\_\_ and \_\_\_\_\_\_\_\_\_\_.”

I am 90% confidence that the weight of M&Ms is between 49.082514 g and 49.756546 g.

I am 95% confidence that the weight of M&Ms is between 49.0161085 g and 49.8229515 g.

I am 99% confidence that the weight of M&Ms is between 48.8833121 g and 49.9557479 g.

1. Given that the true value of the population mean is 47.9 gram, does *any of your three confidence intervals* contain the true value (the true value is within the lower and upper bounds)? If yes, which one?

All of the three confidence intervals do not contain the true value.

1. [Bonus] Where can you find the target population parameter, i.e., the average weight of a bag of regular package of M&M’s, in the real world?
2. [Bonus) Can you think of some possible reasons why the confidence intervals do not contain the true value (the population mean, the average weight of M&M’s)? List the reasons and explain.