

STAT 515 hw 6

CIs for mean with σ known, CIs for σ^2 , large-sample CIs for mean/proportion

Attach a sheet with the R plots and R code printed on it. You may write out your other answers by hand if you want. Just try to make it easy for me grade!!

1. A beekeeper wishes to know the proportion of drones in a hive. She scoops a half-cup of bees from the hive and finds 40 drones among the 294 bees in the scoop. Assume each bee has the same probability of being in the scoop.
 - (a) Give an estimate of the proportion of drones in the hive.
 - (b) Give a 95% Wald-type confidence interval for the proportion of drones in the hive.
 - (c) Are the conditions satisfied for using the Wald-type confidence interval?
 - (d) Give a 95% Agresti-Coull-type confidence interval for the proportion of drones in the hive.
 - (e) The beekeeper would like the proportion of drones in the hive to be no greater than 15%. What does her data tell her?
 - (f) If the true proportion of drones in the hive were equal to 0.15, with what probability would the beekeeper obtain 40 or more drones in her scoop of 294 bees?
 - i. Compute this probability exactly, assuming that there are 30,000 bees in the hive.
 - ii. Compute this probability ignoring the fact that she is sampling without replacement.
 - iii. Compute an approximation to this probability using the Normal distribution.
2. Import the abalone data set into R with this command:

```
load(url("https://people.stat.sc.edu/gregorkb/data/abalone.Rdata"))
```

Regard the abalones in this data set as the entire population of abalones in the world. We are interested in the shucked weight of the abalones.

- (a) Make a histogram of the shucked weights of the “world population” of abalones.
- (b) Make a Normal quantile-quantile plot of the shucked weights using the `qqnorm()` function.
- (c) Do the shucked weights of abalones appear to follow a Normal distribution?
- (d) Give the mean μ and the standard deviation σ of the shucked weights of the “world population” of abalones.
- (e) Take the first 40 abalones in the data set and regard them as a random sample of $n = 40$ abalones. Based on the shucked weights of these 40 abalones, give a confidence interval for the mean shucked weight of all abalones at the confidence level
 - i. 80%.
 - ii. 90%.
 - iii. 95%.
- (f) Which of the intervals from the previous part contained the true mean?

- (g) Run a simulation: For the sample sizes $n = 5$ and $n = 40$, draw 1,000 random samples from the “world population” of abalones. With each random sample, construct a 90%, a 95%, and a 99% confidence interval for the mean shucked weight of abalones. Record for each confidence interval whether it contained the true value of the population mean. In the end, record in a table like the one below the proportion of times (out of the 1,000 random samples) the confidence interval contained the true mean:

n	90%	95%	99%
5	0.898	0.956	0.992
40	0.899	0.954	0.993

Here is some sample code to get you started:

```
n <- 5
S <- 1000
cov90 <- numeric(S)
for(s in 1:S){

  X <- sample(population,n, replace = FALSE)
  X.bar <- mean(X)

  lo90 <- X.bar - 1.645 * 0.222 / sqrt(n)
  up90 <- X.bar + 1.645 * 0.222 / sqrt(n)

  cov90[s] <- (lo90 < mu) & (up90 > mu)

}
mean(cov90)
```

- (h) Is there anything surprising about your results?

3. Bring the `trees` dataset (included in R) into the workspace with the command `data(trees)`. We will consider the heights of the trees.
- Make a Normal quantile-quantile plot of the tree heights.
 - State whether you think the tree heights can be regarded as coming from a Normal population.
 - Give a 95% confidence interval for the variance σ^2 of the population of tree heights.

Optional (do not turn in) problems for additional study from McClave, J.T. and Sincich T. (2017) *Statistics*, 13th Edition: 7.8, 7.12, 7.60, 7.62