Sepehr Akbari

March 27, 2024

Confidence intervals and sample size

Problem 1

A fast-food restaurant needs to estimate the mean carbohydrate count in a new sandwich to within 15 grams. How large a sample is needed if the population standard deviation is $\sigma = 25g$? Use 95% confidence.

Answer:

Using the folowing equation we can calculate the sample size (n):

$$n = (\frac{z \times \sigma}{E})^2$$

```
E <- 15
stndDev <- 25
ci <- qnorm(0.975) # 1.960

n <- (ci^2 * stndDev^2) / E^2

cat("The sample size (n) is:",round(n,0))
```

```
The sample size (n) is: 11
```

Problem 2

How many sandwiches would the restaurant need to test to estimate the mean carbohydrate count to within $5g$?

Answer:

```
E <- 5
stndDev <- 25
ci <- qnorm(0.975) # 1.960

n <- (ci^2 * stndDev^2) / E^2

cat("The new sample size (n) is:",round(n,0))
```

```
The new sample size (n) is: 96
```

Problem 3

Suppose the restaurant realizes that they've underestimated σ, the amount of variability in the carbs of their sandwiches. Would the sample sizes in problems 1 and 2 be increased or decreased?

Answer:

If the restaurant underestimates the standard deviation (σ), the actual variability is higher than anticipated. This leads to a wider range of possible carbohydrate values, requiring a larger sample size to achieve the same level of confidence or margin of error.

Therefore, in both problems 1 and 2, the sample sizes would need to be increased to compensate for the higher variability cause by the underestimated standard deviation.

We can actually also test this real quick:

```r
# CONSTANTS
E <- 10
ci <- qnorm(0.975) # 1.960

# VARIABLES
stndDev <- 20 # Overestimated
stndDev2 <- 25 # Original
stndDev3 <- 30 # Underestimated

n <- (ci^2 * stndDev^2) / E^2
n2 <- (ci^2 * stndDev2^2) / E^2
n3 <- (ci^2 * stndDev3^2) / E^2

cat("The new sample size (n), with the overestimated deviation is:"
    ,round(n,3),"\n")
```

```
The new sample size (n), with the overestimated deviation is: 15.366
```

```r
cat("The new sample size (n), with the original deviation is:"
    ,round(n2,3),"\n")
```

```
The new sample size (n), with the original deviation is: 24.009
```

```r
cat("The new sample size (n), with the underestimated deviation is:"
    ,round(n3,3))
```

```
The new sample size (n), with the underestimated deviation is: 34.573
```