A_M6

October 19, 2024

1 Module 6 Autograded Assignment

[ ]:
```
# Load Packages
library(testthat)
```

2 Problem 1

You are a quality assurance tester for hardware equipment and been given a lot of $10,000$ nails to test. You decide to randomly sample 100 nails from the lot and test if they are defective. The whole lot would be considered defective if the average of the sample is less than 11.97cm or greater than 12.04cm. Given that the true population mean μ is 12cm with a true standard deviation σ of 0.2cm, what is the probability that the lot is **not** defective? Round your answer to two decimal places. Save your answer as `p1`.

[ ]:
```
p1 = NA
# your code here
mu <- 12
sigma <- 0.2
n <- 100
lower_bound <- 11.97
upper_bound <- 12.04

SE <- sigma / sqrt(n)

z_lower <- (lower_bound - mu) / SE
z_upper <- (upper_bound - mu) / SE

p_lower <- pnorm(z_lower)
p_upper <- pnorm(z_upper)

p1 <- p_upper - p_lower
round(p1, 2)
p1
```

[ ]:
```
# Hidden Test Cell
```

3 Problem 2

Suppose you are out fishing and you want to predict the amount of time (in minutes) it takes to catch 3 fish. Let $T_F \sim N(15, 10)$ be the time it takes to reel in a fish. Let $T_T \sim N(10, 8)$ be the amount of time you spend waiting for the fish to take the bait. Assume T_F and T_T are independent. Also assume you start timing as soon as you start reeling in the first fish, and the time ends when you have finished catching the third fish.

a) Let T be the random variable for the total amount of time spent fishing. What is $E[T]$ and $Var[T]$? Save your variables as `exp.T` and `var.T`. (Hint: Note that a variance of 122 is incorrect for this problem. Why? If Z_1 and Z_2 are indepedent N(0,1) random variables, the distribution of $Z_1 + Z_2 $ is not the same as $2Z_1$. If you want to explore this, you can run some simulations to see for yourself.)

[11]:
```
exp.T = NA
var.T = NA
# your code here
# Given parameters
# Given values
E_X <- 15
Var_X <- 10
E_Y <- 10
Var_Y <- 8

exp.T <- 3 * E_X + 2 * E_Y

var.T <- 3 * Var_X + 2 * Var_Y

exp.T
var.T
```

65

46

[ ]:
```
# Hidden Test Cell
```

b) What is the probability that you finish catching the three fish in 60 minutes or less? Round your answer to three decimal places. Save your answer as `p2.2`.

[10]:
```
p2.2 = NA

# your code here
mean_T <- 65
var_T <- 46
sd_T <- sqrt(var_T)

X <- 60
Z <- (X - mean_T) / sd_T
```

```
p2.2 <- pnorm(Z)

p2.2 <- round(p2.2, 3)

p2.2
```

0.325

[ ]:
```
# Hidden Test Cell
```

4 Problem 3

Exams given by a particular instructor in a freshman-level statistics course have a mean of 75 and a standard deviation of 17. The instructor gives an exam to two classes, one with 30 students and the other with 90 students.

a) Let $\bar{X}$ represent the average exam score in the class of 30. Approximate the probability that the average test score $\bar{X}$ exceeds 80, that is, estimate $P(\bar{X} \geq 80)$. Round your answer to four decimal places. Save your answer as `p3.a`.

[ ]:
```
p3.a = NA
# your code here
mu <- 75
sigma <- 17
n <- 30
X_bar <- 80

SE <- sigma / sqrt(n)

Z <- (X_bar - mu) / SE

p3.a <- 1 - pnorm(Z)

p3.a <- round(p3.a, 4)

p3.a
```

[ ]:
```
# Hidden Test Cell
```

b) Let $\bar{Y}$ represent the average exam score in the class of 90. Approximate the probability that the average test score $\bar{Y}$ exceeds 80, that is, estimate $P(\bar{Y} \geq 80)$. Round your answer to four decimal places. Save your answer as `p3.b`.

[ ]:
```
p3.b = NA
# your code here
mu <- 75
```

```
sigma <- 17
n <- 90
Y_bar <- 80

SE <- sigma / sqrt(n)

Z <- (Y_bar - mu) / SE

p3.b <- 1 - pnorm(Z)

p3.b <- round(p3.b, 4)

p3.b
```

[ ]:
```
# Hidden Test Cell
```

Thought Question (Not Graded) Would you be surpised to learn that the 90 person class had an exam average of 80 or higher?

c) Estimate the probability that the averages of the two classes are within one point of each other, that is, estimate $P(-1 \leq \bar{Y} - \bar{X} \leq 1)$. Hint: What is the distribution of $\bar{Y} - \bar{X}$? Round your answer to two decimal places. Save your answer as `p3.c`.

[ ]:
```
p3.c = NA
# your code here
sigma <- 17
n_X <- 30
n_Y <- 90

var_diff <- (sigma^2 / n_Y) + (sigma^2 / n_X)
sigma_diff <- sqrt(var_diff)

Z_neg1 <- (-1 - 0) / sigma_diff
Z_pos1 <- (1 - 0) / sigma_diff

p3.c <- pnorm(Z_pos1) - pnorm(Z_neg1)

p3.c <- round(p3.c, 2)

p3.c
```

[ ]:
```
# Hidden Test Cell
```

5 Problem 4

The Central Limit Theorem states: Let $X_1, \ldots, X_n$ be a random sample with $E[X_i] = \mu$ and $Var(X_i) = \sigma^2$. If n is sufficiently large, then $\bar{X}$ has approximately a normal distribution with mean $\mu_{\bar{X}} = \mu$ and variance $\sigma^2_{\bar{X}} = \sigma^2/n$. It's also true that $\sum_{i=1}^{n} X_i \sim N(n\mu, n\sigma^2)$.

Suppose a machine requires a specific type of battery that lasts an exponential amount of time with mean 25 hours. As soon as the battery fails, you replace it immediately. If you have 50 such batteries, estimate the probability that the machine is still operating after 1300 hours. Round your answer to three decimal places. Save your answer as `p4`.

[ ]:
```r
p4 = NA
# your code here
n <- 50
mu <- 25
total_hours <- 1300

mean_total <- n * mu
var_total <- n * mu^2
sigma_total <- sqrt(var_total)

Z <- (total_hours - mean_total) / sigma_total

p4 <- 1 - pnorm(Z)

p4 <- round(p4, 3)

p4
```

[ ]:
```r
# Hidden Test Cell
```