```
#problem 1
#given N(31.5, 0.8)
pnorm(1,lower.tail = TRUE) - pnorm(-1, lower.tail = TRUE)

#problem 2
pnorm(33.9, mean = 31.5, sd = 0.8, lower.tail = TRUE) - pnorm(29.1, mean = 31.5, sd =
0.8, lower.tail = TRUE)

#problem 3, part A
data <- c(99, 123, 75, 138, 105, 65, 116)
sample_mean <- mean(data)
print(sample_mean)

#part B
sample_var <- var(data)
sample_sd <- sd(data)
print(paste("This is sample variance:", sample_var))
print(paste("This is sample SD:", sample_sd))

#part C
degrees_of_freedom <- length(data) - 1
bound <- qt(0.025, degrees_of_freedom, lower.tail=TRUE) * sample_sd /
sqrt(degrees_of_freedom)
abs_bound <- abs(bound)
print(paste("The 95% interval is: (",(sample_mean - abs_bound), "; ", (sample_mean +
abs_bound), ")"))

#part D
null_value <- 90
confidence_level <- 0.9
result <- t.test(data, mu = null_value, alternative = "two.sided",conf.level =
confidence_level)
print(result)
#Since p-value > alpha (0.23>0.1), we do not have sufficient evidence
#to reject the H0.

#problem 4
#Given: n=81, sample mean = 57.8 feet, s = 6.02 feet
#Competitors claim 60 feet

#Part A: 90% and 95% C-Intervals for Mu
degrees_of_freedom <- 80
sample_mean <- 57.8
sample_sd <- 6.02
bound <- qt(0.05, degrees_of_freedom, lower.tail=TRUE) * sample_sd /
sqrt(degrees_of_freedom)
abs_bound <- abs(bound)
print(paste("The 90% interval is: (",(sample_mean - abs_bound), "; ", (sample_mean +
abs_bound), ")"))
bound <- qt(0.025, degrees_of_freedom, lower.tail=TRUE) * sample_sd /
sqrt(degrees_of_freedom)
abs_bound <- abs(bound)
print(paste("The 95% interval is: (",(sample_mean - abs_bound), "; ", (sample_mean +
abs_bound), ")"))

#Part B
#H0: Mu = 60; Ha: Mu < 60
#Since n>30, by CLT we can use Normal distribution
n <- 81
sample_mean <- 57.8
sample_sd <- 6.02
null_value <- 60
confidence_level <- 0.95
z_score <- (sample_mean - null_value) / (sample_sd / sqrt(n))
p_value <- pnorm(z_score, lower.tail = TRUE)
cat("Z-score:", z_score, "\n")
cat("p-value:", p_value, "\n")
```

HW 1

Since independent, $E(XY) = E(X) \cdot E(Y)$

$E(X^2) - (E(X))^2 = 1$
$E(X^2) = 1 + 2^2 = 5$

$E(Y^2) - (E(Y))^2 = 1$
$E(Y^2) = 1 + 3^2 = 10$

1. Consider random variables X and Y which are **independent** to each other. Suppose $E[X] = 2$ and $E[Y] = 3$, $Var(X) = 1$ and $Var(Y) = 1$. Calculate the followings:

   (a) $E[3X(1+Y)]$ (2 pts) $= 3 \cdot E(X + XY) = 3E(X) + 3E(XY) = 6 + 3 \cdot 2 \cdot 3 = 24$

   (b) $Var(2X - Y + 1)$ (2 pts) $= 4Var(X) + Var(Y) + Var(1) = 4 + 1 + 0 = 5$

   (c) $E[(X+Y)^2]$ (Hint: use $Var(X) = E[X^2] - E[X]^2$) (2 pts)
   $E(X^2 + 2XY + Y^2) = 5 + 10 + 2 \cdot E(XY) = 15 + 2 \cdot 3 \cdot 2 = 27$

2. Suppose that the population of all gasoline mileages for the GSX-50 is normally distributed with mean $\mu = 31.5$mpg and standard deviation $\sigma = 0.8$mpg. Let y denote a mileage randomly selected from this population. Find the following probabilities.

   $N(31.5, 0.8)$

   (a) $P(30.7 \leq y \leq 32.3)$ (3 pts) $= P\left(\frac{30.7 - 31.5}{0.8} \leq \frac{y - \mu}{\sigma} \leq \frac{32.3 - 31.5}{0.8}\right) = P(-1 \leq Z \leq 1) =$ pnorm(1, l.tail=T) − pnorm(−1, l.tail=T) $= 0.6826895$

   (b) $P(29.1 \leq y \leq 33.9)$ (3 pts) $= P\left(\frac{29.1 - 31.5}{0.8} \leq \frac{y - \mu}{\sigma} \leq \frac{33.9 - 31.5}{0.8}\right) = 0.9973002$

3. Consider the following sample of seven dollar amounts owed by customers with delinquent charge accounts:

   \$99 \$123 \$75 \$138 \$105 \$65 \$116

   (a) Calculate the sample mean $\bar{y}$ (2 pts). $= 103.$

   (b) Calculate the sample variance s^2 and sample standard deviation s (2 pts). $S^2 = 673.\bar{6}$; $S = 25.955$

   (c) Find 95% confidence intervals for the mean. (2 pts). $(77.07 ; 128.93)$

   (d) Now, a store owner wants to check that if there is an evidence that the mean of delinquent charge is **different from** 90. Answer the following questions. (3 pts)

   - State your hypothesis (H_0 and H_a) $H_0: \mu = 90$ $H_A: \mu \neq 90$
   - Find Test statistic $t = 1.3252$
   - Find p-value $p = 0.233$
   - Write your conclusion. Use $\alpha = 0.1$. Since p-value is bigger than α $(0.23 > 0.1)$, we do not have sufficient evidence to reject H_0.

4. National Motors has equipped the ZX-900 with a new disk brake system. We define the stopping distance for a ZX-900 to be the distance (in feet) required to bring the automobile to a complete stop from a speed of 35mph under normal driving conditions using this new brake system. In addition, we define μ to be a mean stopping distance of all ZX-900s. One of the ZX-900's major competitors is advertised as achieving a mean stopping distance of 60 feet. National Motors would like to claim in a new advertising campaign that the ZX-900 achieves a shorter mean stopping distance.
   Suppose that National Motors randomly selects a sample of $n = 81$ ZX-900s. The company records the stopping distance of each of these automobiles and calculates the mean and standard deviation of the sample of $n = 81$ stopping distances to be $\bar{y} = 57.8$feet and $s = 6.02$feet.

(a) Calculate 90% and 95% confidence intervals for μ. Also, using the 95% confidence can National Motors be at least 95% confident that μ is less than 60 feet? Explain.(3 pts)

(b) The company wants to check that if there is an evidence that the mean of stopping distance is **less than** 60. Answer the following questions. (3 pts)

- State your hypothesis (H_0 and H_a)
- Find Test statistic
- Find p-value
- Write your conclusion. Use $\alpha = 0.05$.

(c) If the conclusion in (b) was in error, what type of error is it, Type I or Type II? Explain the reason. (2 point)

a) 90% int: (56.6799; 58.9200)

95% int: (56.4606; 59.1394)

Yes, b/c the interval does not include the value of 60 feet.

b) $H_0: \mu = 60$ $\quad$ $Z = -3.289037$

$H_A: \mu < 60$ $\quad$ p-value $= 0.000562648$

Since p-value $< \alpha$ $(0.0005 < 0.05)$, we do have sufficient evidence to reject the H_0.

c) Since Type I error happens if H_0 is true, but it is rejected — that fits our situation and therefore, is possible.