Homwork2-Exam 1 Preparation HW

Group_6#

2025-02-24

NOTE: only those who contributes and fully participates in the work will get credit

**Scribe:*group 6*

**Moderator:*Faisal Rehman*

**All contributors: * Ankit , Jay and Faisal Equal*

This homework is accumulative, with the intention of assisting you to prepare for Exam1. It is due Wednesday 11:59pm, you will get extra 10pts for this HW if you finish it by Tuesday 11:59pm. Yes, I know it has lots of problems, but I believe that you want to perform well for the coming exam, so this HW will be very beneficial!

Attention: Please finish all 14 problems step by step, rather than using the advance functions `t.test`, `prop.test`. However, you are encouraged to use more advance functions to check your solutions.

##Q1 Devore 2nd edition 7.1.5 : Assume that the helium porosity (in percentage) of coal samples taken from any particular seam is normally distributed with true standard deviation .75. a. Compute a 95% CI for the true average porosity of a certain seam if the average porosity for 20 specimens from the seam was 4.85. b. Compute a 98% CI for true average porosity of another seam based on 16 specimens with a sample average porosity of 4.56. c. How large a sample size is necessary if the width of the 95% interval is to be .40? d. What sample size is necessary to estimate true average porosity to within .2 with 99% confidence?

```
sd <- 0.75
n  <- 20
sample_mean <- 4.85
alpha <- 0.05
z_alpha <- qnorm(1 - alpha/2)
# part A :
l <- sample_mean - z_alpha * sd / sqrt(n)
u <- sample_mean + z_alpha * sd / sqrt(n)
cat("CI:", round(l, 4), ",", round(u, 4), "\n")
```

```
## CI: 4.5213 , 5.1787
```

```
# part B :
z_alpha <- qnorm(1 - 0.02/2)
n <- 16
sample_mean <- 4.56
lower_bound <- sample_mean - z_alpha * sd / sqrt(n)
upper_bound <- sample_mean + z_alpha * sd / sqrt(n)
cat("CI:", round(lower_bound, 4), ",", round(upper_bound, 4), "\n")
```

```
## CI: 4.1238 , 4.9962
```

```r
# Part c:
Sdc <- 0.75
alpha <- 0.05
width <- 0.40
z_alpha <- qnorm(1 - alpha/2)
n_req <- ((z_alpha * Sdc) / (width / 2))^2
cat("Sample Size =", ceiling(n_req), "\n")
```

```
## Sample Size = 55
```

```r
# Part d:
s <- 0.75
alpha <- 0.01
MOE <- 0.2
z_alpha <- qnorm(1 - alpha/2)
n_required <- ((z_alpha * s) / MOE)^2
cat("Sample Size =", n_required, "\n")
```

```
## Sample Size = 93.30323
```

Q2 7.2.13:

The article "Gas Cooking, Kitchen Ventilation, and Exposure to Combustion Products" (Indoor Air, 2006: 65–73) reported that for a sample of 50 kitchens with gas cooking appliances monitored during a one week period, the sample mean CO2 level (ppm) was 654.16, and the sample standard deviation was 164.43. a. Calculate and interpret a 95% (two-sided) confidence interval for true average CO2 level in the population of all homes from which the sample was
selected. b. Suppose the investigators had made a rough guess of 175 for the value of s before collecting data. What sample size would be necessary to obtain an
interval width of 50 ppm for a confidence level of 95%?.

```r
mean <- 654.16
sd <- 164.43
n <- 50
# Part a:
alpha <- 1 - 0.95
t_critical <- 1.96
MOE <- t_critical * (sd / sqrt(n))
CI_l <- mean - MOE
CI_u <- mean + MOE
cat("CI: ", round(CI_l, 4), ",", round(CI_u, 4), "\n")
```

```
## CI:  608.5823 , 699.7377
```

```r
# Part b:
w <- 50
s <- 175
z_critical <- qnorm(1 - alpha/2)
n <- ((2 * z_critical * s) / w) ^ 2
cat("Sample Size:", ceiling(n), "\n")
```

```
## Sample Size: 189
```

Q3 7.2.20:

TV advertising agencies face increasing challenges in reaching audience members because viewing TV programs via digital streaming is gaining in popularity. The Harris poll reported on November 13, 2012, that 53% of 2343 American adults surveyed said they have watched digitally streamed TV programming on some type of device. a. Calculate and interpret a confidence interval at the 99% confidence level for the proportion of all adult Americans who watched streamed programming up to that point in time. b. What sample size would be required for the width of a 99% CI to be at most .05 irrespective of the value of phat?

```
n <- 2343
phat <- 0.53
alpha <- 1 - 0.99

# Part a:
z_cr <- qnorm(1 - alpha/2)
se <- sqrt((phat * (1 - phat)) / n)
margin_error <- z_cr * se
C_lower <- phat - margin_error
C_upper <- phat + margin_error
cat("CI: ", round(C_lower, 4), ",", round(C_upper, 4), "\n")
```

```
## CI:  0.5034 , 0.5566
```

```
# Part b:
width <- 0.05
phatmax <- 0.5
qhatmax <- 1 - phatmax
alpha <- 1 - 0.99
z_crp <- qnorm(1 - alpha/2)
n_Sample <- (( z_crp * sqrt(phatmax * qhatmax)) / width) ^ 2
cat("Sample Size:", ceiling(n_Sample), "\n")
```

```
## Sample Size: 664
```

Q4 7.3.36:

A normal probability plot of the n 5 26 observations on escape time given in Exercise 36 of Chapter 1 shows a substantial linear pattern; the sample mean and sample standard deviation are 370.69 and 24.36, respectively. a. Calculate an upper confidence bound for population mean escape time using a confidence level of 95%. b. Calculate an upper prediction bound for the escape time of a single additional worker using a prediction level of 95%. How does this bound compare with
the confidence bound of part (a)? c. Suppose that two additional workers will be chosen to participate in the simulated escape exercise. Denote their escape times by X27 and X28, and let Xnew denote the average of these two values. Modify the formula for a PI for a single x value to obtain a PI for Xnew, and calculate a 95% two-sided interval
based on the given escape data.

```
n <- 26
sample_mean <- 370.69
sample_sd <- 24.36
alpha <- 1 -0.95
```

```r
# Part a:
df <- n - 1
t_f_tailed <- qt(1 - alpha, df)
error_mean <- t_f_tailed * (sample_sd / sqrt(n))
CI_level <- sample_mean + error_mean
cat("CI:", round(CI_level, 4), "\n")
```

```
## CI: 378.8505
```

```r
# Part b:
error_pr <- t_f_tailed * sample_sd * sqrt(1 + (1 / n))
pred_obs <- sample_mean + error_pr
cat("First Observation:", round(pred_obs, 2), "\n")
```

```
## First Observation: 413.09
```

```r
# Part c:
se_avg_two <- sample_sd * sqrt((1 / 2) + (1 / n))
t_s_tailed <- qt(1 - alpha, df)
lower <- sample_mean - t_s_tailed * se_avg_two
upper <- sample_mean + t_s_tailed * se_avg_two
cat("Second Observations: [",  round(lower, 4), ",", round(upper, 4), "]\n")
```

```
## Second Observations: [ 340.1564 , 401.2236 ]
```

Q5 8.1.13:

The calibration of a scale is to be checked by weighing a 10-kg test specimen 25 times. Suppose that the results of different weighing are independent of one another and that the weight on each trial is normally distributed with sd = 0.200 kg. Let m denote the true average weight reading on the scale. a. What hypotheses should be tested?. b. With the sample mean itself as the test statistic, what is the P-value when xbar = 9.85, and what would you conclude at significance level .01? c. For a test with alpha =0.01, what is the probability that re calibration is judged unnecessary when in fact m_1= 10.1? When m_2=9.8?.

```r
mu <- 10
sd <- 0.200
n <- 25
alpha <- 0.01

# Part a:
cat("Hypotheses\n")
```

```
## Hypotheses
```

```r
cat("H0: mu = 10\n")
```

```
## H0: mu = 10
```

```r
cat("H1: mu != 10\n\n")
```

```
## H1: mu != 10
```

```r
# Part b:
mean_b <- 9.85
z_score <- (mean_b - mu) / (sd / sqrt(n))
p_value <- 2 * pnorm(-abs(z_score))
cat("P-value =", (p_value), "\n")
```

```
## P-value = 0.0001768346
```

```r
cat("We Reject the null hypothesis at alpha =", alpha, "\n\n")
```

```
## We Reject the null hypothesis at alpha = 0.01
```

```r
# Part c:
se <- sd / sqrt(n)
z_critical <- qnorm(1 - alpha / 2)
lower <- mu - z_critical * se
upper <- mu + z_critical * se
cat("Z_Critical: ", round(lower, 3), ",", round(upper, 3), "\n")
```

```
## Z_Critical:  9.897 , 10.103
```

```r
type_II_error <- function(true_mean) {
  lower_zCrt <- (lower - true_mean) / se
  upper_zzCrt <- (upper - true_mean) / se
  beta <- pnorm(upper_zzCrt) - pnorm(lower_zCrt)
  return(beta)
}
beta_1 <- type_II_error(10.1)
beta_98 <- type_II_error(9.8)
cat("Beta :", round(beta_1, 4), "\n")
```

```
## Beta : 0.5302
```

```r
cat("Beta:", round(beta_98, 4), "\n")
```

```
## Beta: 0.0077
```

Q6 8.2.21:

The desired percentage of SiO2 in a certain type of aluminous cement is 5.5. To test whether the true average percentage is 5.5 for a particular production facility, 16 independently obtained samples are analyzed. Suppose that the percentage of SiO2 in a sample is normally distributed with sd = 0.3 and that xbar = 5.25. a. Does this indicate conclusively that the true average percentage differs from 5.5? b. If the true average percentage is mu=5.6 and a level alpha= .01 test based on n = 16 is used, what is the probability of detecting this departure from H0? c. What value of n is required to satisfy alpha= .01 and beta(5.6) = .01?.

```r
mu <- 5.5
xbar <- 5.25
sd1 <- 0.3
n <- 16
alpha <- 0.01

# Part a:
z_score <- (xbar - mu) / (sd1 / sqrt(n))
p_value <- 2 * (1 - pnorm(abs(z_score)))
cat("Z-score =", round(z_score, 4), "\n")
```

```
## Z-score = -3.3333
```

```r
cat("P-value =", round(p_value, 5), "\n")
```

```
## P-value = 0.00086
```

```r
cat("We Reject the null hypothesis at alpha =", alpha, "\n\n")
```

```
## We Reject the null hypothesis at alpha = 0.01
```

```r
# Part b:
mu_true <- 5.6
z_crit <- qnorm(1 - alpha / 2)
l_bound <- mu - z_crit * (sd1 / sqrt(n))
u_bound <- mu + z_crit * (sd1 / sqrt(n))
lower_z <- (l_bound - mu_true) / (sd1 / sqrt(n))
upper_z <- (u_bound - mu_true) / (sd1 / sqrt(n))
beta <- pnorm(upper_z) - pnorm(lower_z)
power <- 1 - beta
cat("Beta =", round(beta, 3), "\n")
```

```
## Beta = 0.893
```

```r
cat("Power =", round(power, 3), "\n\n")
```

```
## Power = 0.107
```

```r
# Part c:
z_alpha <- qnorm(1 - alpha / 2)
z_beta <- qnorm(1 - 0.01)
N_values <- ((sd1 * (z_alpha + z_beta)) / abs(mu - mu_true))^2
cat("Sample Size =", ceiling(N_values), "\n")
```

```
## Sample Size = 217
```

Q7 8.3.34:

The following observations are on stopping distance (ft) of a particular truck at 20 mph under specified experimental conditions ("Experimental Measurement of the Stopping Performance of a Tractor-Semitrailer from Multiple Speeds," NHTSA, DOT HS 811 488, June 2011): 32.1 30.6 31.4 30.4 31.0 31.9 The cited report states that under these conditions, the maximum allowable stopping distance is 30. A normal probability plot validates the assumption that stopping distance is normally distributed. a. Does the data suggest that true average stopping distance exceeds this maximum value? Test the appropriate hypotheses using alpha= .01. b. Determine the probability of a type II error when alpha= .01, sd= .65, and the actual value of mu is 31. Repeat this for mu = 32 (use either statistical software or Table A.17). c. Repeat (b) using sd= .80 and compare to the results of (b). d. What sample size would be necessary to have alpha=.01 and beta= .10 when mu= 31 and sd= .65?

```r
#type codes here
data <- c(32.1, 30.6, 31.4, 30.4, 31.0, 31.9)
mu_0 <- 30
n <- 6
xbar <- mean(data)
sdd <- sd(data)
alpha <- 0.01

# Part a:
t_stat <- (xbar - mu_0) / (sdd / sqrt(n))
p_value <- 1 - pt(t_stat, df = n - 1)
cat("T-statistic =", round(t_stat, 4), "\n")
```

```
## T-statistic = 4.3849
```

```r
cat("P-value =", round(p_value, 4), "\n")
```

```
## P-value = 0.0036
```

```r
cat("We Reject the null hypothesis at alpha = ", alpha, "\n\n")
```

```
## We Reject the null hypothesis at alpha =  0.01
```

```r
# Part b:
s65 <- 0.65
t_critical <- qt(1 - alpha, df = n - 1)
x_crit <- mu_0 + t_critical * s65 / sqrt(n)
z_31 <- (x_crit - 31) / (s65 / sqrt(n))
z_32 <- (x_crit - 32) / (s65 / sqrt(n))
beta_31 <- pnorm(z_31)
beta_32 <- pnorm(z_32)
cat("Beta :", round(beta_31, 4), "\n")
```

```
## Beta : 0.3433
```

```r
cat("Beta :", round(beta_32, 5), "\n\n")
```

```
## Beta : 2e-05
```

```r
# Part c:
sigma_80 <- 0.80
se_80 <- sigma_80 / sqrt(n)
z_31_80 <- (x_crit - 31) / se_80
beta_31_80 <- pnorm(z_31_80)
z_32_80 <- (x_crit - 32) / se_80
beta_32_80 <- pnorm(z_32_80)
cat("Beta :", round(beta_31_80, 4), "\n")
```

```
## Beta : 0.3715
```

```r
cat("Beta :", round(beta_32_80, 6), "\n\n")
```

```
## Beta : 0.00035
```

```r
# Part d: Sample Size Calculation
sigma <- 0.65
mu_alt <- 31
beta <- 0.10
z_alpha <- qnorm(1 - alpha/2)
z_beta <- qnorm(1 - beta/2)
n_required <- ((z_alpha + z_beta) * sigma / (mu_alt - mu_0))^2
cat("Sample Size =", ceiling(n_required), "\n")
```

```
## Sample Size = 8
```

Q8 8.4.48:

With domestic sources of building supplies running low several years ago, roughly 60,000 homes were built with imported Chinese drywall. According to the article "Report Links Chinese Drywall to Home Problems" (New York Times, Nov. 24, 2009), federal investigators identified a strong association between chemicals in the drywall and electrical problems, and there is also strong evidence of respiratory difficulties due to the emission of hydrogen sulfide gas. An extensive examination of 51 homes found that 41 had such problems. Suppose these 51 were randomly sampled from the population of all homes having Chinese drywall. a. Does the data provide strong evidence for concluding that more than 50% of all homes with Chinese drywall have electrical/environmental problems? Carry out a test of hypotheses using a 5 .01. b. Calculate a lower confidence bound using a confidence level of 99% for the percentage of all such homes that have electrical/environmental problems. c. If it is actually the case that 80% of all such homes have problems, how likely is it that the test of (a) would not conclude that more than 50% do?

```r
n <- 51
x <- 41
p0 <- 0.5
alpha <- 0.01
phat <- x / n

# Part a:
z_stat <- (phat - p0) / sqrt(p0 * (1 - p0) / n)
p_value <- 1 - pnorm(z_stat)
cat("Z-Statistic: ", round(z_stat, 5), "\n")
```

```
## Z-Statistic:  4.34087
```

```r
cat("P-Value:", round(p_value, 3), "\n")
```

```
## P-Value: 0
```

```r
cat(" we reject the null hypothesis at alpha =", alpha, "\n")
```

```
##  we reject the null hypothesis at alpha = 0.01
```

```r
# Part b:
z_alpha <- qnorm(1 - 0.01)
lower_bound <- phat - z_alpha * sqrt(phat * (1 - phat) / n)
cat("CI  :", round(lower_bound, 4), "\n")
```

```
## CI  : 0.6746
```

```r
# Part c:
p_alt <- 0.8
numerator <- p0 - p_alt + z_alpha * sqrt(p0 * (1 - p0) / n)
denominator <- sqrt(p_alt * (1 - p_alt) / n)
z_beta <- numerator / denominator
beta <- pnorm(z_beta)
cat("Beta : ", round(beta, 4), "\n")
```

```
## Beta :  0.0072
```

Q9 9.1.6:

An experiment to compare the tension bond strength of polymer latex modified mortar (Portland cement mortar to which polymer latex emulsions have been added during mixing) to that of unmodified mortar resulted in xbar 18.12 kgf/cm2 for the modified mortar (mu = 40) and ybar 5 16.87 kgf/cm2 for the unmodified mortar (mu2= 32). Let m1 and m2 be the true average tension bond strengths for the modified and unmodified mortars, respectively. Assume that the bond strength distributions are both normal. a. Assuming that s1 = 1.6 and s2= 1.4, test H0: m1 - m2 = 0 versus Ha: m1 - m2 greater then 0 at level .01. b. Compute the probability of a type II error for the test of part (a) when m1 - m2 = 1. c. Suppose the investigator decided to use a level .05 test and wished beta= .10 when m1 - m2 = 1. If m u= 40, what value of n is necessary? d. How would the analysis and conclusion of part (a) change if s1 and s2 were unknown but s1 = 1.6 and s2 = 1.4?

```r
xbar1 <- 18.12
ybar <- 16.87
s1 <- 1.6
s2 <- 1.4
n1 <- 40
n2 <- 32
alpha_a <- 0.01

# Part a:
z_stat <- (xbar1 - ybar) / (sqrt((s1^2)/n1 + (s2^2)/n2))
p_value <- 1 - pnorm(z_stat)
cat("Z-Statistic =", round(z_stat, 4), "\n")
```

```
## Z-Statistic = 3.532
```

```r
cat("P-Value =", round(p_value, 5), "\n")
```

```
## P-Value = 0.00021
```

```r
cat(" So we Reject the null hypothesis at alpha =", alpha_a, "\n")
```

```
##  So we Reject the null hypothesis at alpha = 0.01
```

```r
# Part b:
delta <- 1
z_alpha <- qnorm(1 - alpha_a)
beta <- pnorm(z_alpha - (delta / se))
cat("Beta =", round(beta, 4), "\n")
```

```
## Beta = 0
```

```r
# Part c:
alpha_c <- 0.05
beta_c <- 0.10
z_alpha_c <- qnorm(1 - alpha_c)
z_beta_c <- qnorm(1 - beta_c)
pooled_var <- s1^2 + s2^2
n_required <- ((z_alpha_c + z_beta_c)^2 * pooled_var) / delta^2
cat("Sample Size :", n_required, "\n")
```

```
## Sample Size : 38.70859
```

```r
# Part d:
cat(" The n = 32 and by using two-sample t-test if s1 and s2 are unknown, but result remains similar du
```

```
##  The n = 32 and by using two-sample t-test if s1 and s2 are unknown, but result remains similar due
```

Q10 9.2.24:

Damage to grapes from bird predation is a serious problem for grape growers. The article "Experimental Method to Investigate and Monitor Bird Behavior and Damage to Vineyards" (Amer. J. of Enology and Viticulture, 2004: 288–291) reported on an experiment involving a bird-feeder table, time-lapse video, and artificial foods. Information was collected for two different bird species at both the experimental location and at a natural vineyard setting. Consider the following data on time (sec) spent on a single visit to the location. Table is attached : a. Calculate an upper confidence bound for the true average time that blackbirds spend on a single visit at the experimental location. b. Does it appear that true average time spent by blackbirds at the experimental location exceeds the true average time birds of this type spend at the natural location? Carry out a test of appropriate hypotheses. c. Estimate the difference between the true average time blackbirds spend at the natural location and true average time that silvereyes spend at the natural location, and do so in a way that conveys information about reliability and precision. [Note: The sample medians reported in the article all seemed significantly smaller than the means, suggesting substantial population distribution skewness. The authors actually used the distribution-free test procedure presented in Section 2 of Chapter 15.]

```r
x_bar <- 13.4
SE <- 2.05
n <- 65
df <- n - 1
alpha <- 0.05

#  part a :
t_critical <- qt(1 - alpha, df)
upp <- x_bar + t_critical * SE
cat("T-critical =", round(t_critical, 4), "\n")
```

```
## T-critical = 1.669
```

```r
cat("CI : =", round(upp, 4), "seconds\n")
```

```
## CI : = 16.8215 seconds
```

```r
#part b:
x_bar2 <- 9.7
SE2 <- 1.76
n2 <- 50
t_stat <- (x_bar - x_bar2) / sqrt(SE^2 + SE2^2)
df_num <- (SE^2 + SE2^2)^2
df_denom <- ((SE^2)^2 / (n1 - 1)) + ((SE2^2)^2 / (n2 - 1))
df <- df_num / df_denom
p_value <- 1 - pt(t_stat, df)
cat("T-Statistic =",t_stat, "\n")
```

```
## T-Statistic = 1.369422
```

```r
cat("P-Value =",p_value, "\n\n")
```

```
## P-Value = 0.0872996
```

```r
cat("We reject the null hypothesis at alpha =", alpha,"There is significant evidence that the true avera
```

```
## We reject the null hypothesis at alpha = 0.05 There is significant evidence that the true average ti
```

```r
#part c:
xbar_silvereye <- 38.4
xbar_blackbird <- 9.7
se_silvereye <- 5.06
se_blackbird <- 1.76
n_silvereye <- 46
n_blackbird <- 50
df <- 55
t_critical <- qt(1 - 0.05/2, df)
mean_diff <- xbar_silvereye - xbar_blackbird
se_diff <- sqrt(se_silvereye^2 + se_blackbird^2)
lower_bound <- mean_diff - t_critical * se_diff
upper_bound <- mean_diff + t_critical * se_diff
cat("CI : (", round(lower_bound, 2), ",", round(upper_bound, 2), ") seconds\n")
```

```
## CI : ( 17.96 , 39.44 ) seconds
```

Q11 9.3.40:

Lactation promotes a temporary loss of bone mass to provide adequate amounts of calcium for milk production. The paper "Bone Mass Is Recovered from Lactation to Postweaning in Adolescent Mothers with Low Calcium Intakes" (Amer. J. of Clinical Nutr., 2004: 1322–1326) gave the following data on total body bone mineral content (TBBMC) (g) for a sample both during lactation (L) and in the post-weaning period (P). a. Does the data suggest that true average total body bone mineral content during postweaning exceeds that during lactation by more than 25 g? State and test the appropriate hypotheses using a significance level of .05. [Note: The appropriate normal probability plot shows some curvature but not enough to cast substantial doubt on a normality assumption.] b. Calculate an upper confidence bound using a 95% confidence level for the true average difference between TBBMC during postweaning and during lactation. c. Does the (incorrect) use of the two-sample t test to test the hypotheses suggested in (a) lead to the same conclusion that you obtained there? Explain.

```r
n <- 10
mean_diff <- 105.7
sd_diff <- 103.845
mu_0 <- 25
# part a:
alpha <- 0.05
t_stat <- (mean_diff - mu_0) / (sd_diff / sqrt(n))
df <- n - 1
p_value <- 1 - pt(t_stat, df)
cat("T-statistic =", round(t_stat, 4), "\n")
```

```
## T-statistic = 2.4575
```

```r
cat("P-value =", round(p_value, 4), "\n")
```

```
## P-value = 0.0182
```

```r
cat("We reject the null hypothesis: True average TBBMC exceeds 25 g.\n\n")
```

```
## We reject the null hypothesis: True average TBBMC exceeds 25 g.
```

```r
# part b:
d_bar <- 105.7
alpha <- 0.05
df <- n - 1
t_critical <- qt(1 - alpha, df)
upper_bound <- d_bar + t_critical * (sd_diff  / sqrt(n))
cat("Critical T-value =", round(t_critical, 4), "\n")
```

```
## Critical T-value = 1.8331
```

```r
cat("CI =", round(upper_bound, 2), "g\n\n")
```

```
## CI = 165.9 g
```

```r
# part c:
new_se <- 235
df_independent <- 17
t_stat_incorrect <- d_bar / new_se
p_value_incorrect <- 1 - pt(t_stat_incorrect, df = df_independent)
cat("T-Statistic =", round(t_stat_incorrect, 4), "\n")
```

```
## T-Statistic = 0.4498
```

```r
cat("P-Value =", round(p_value_incorrect, 4), "\n")
```

```
## P-Value = 0.3293
```

```r
cat("We reject the null hypothesis at alpha =", alpha, "\n")
```

```
## We reject the null hypothesis at alpha = 0.05
```

Q12 : 9.4.50 :

Recent incidents of food contamination have caused great concern among consumers. The article "How Safe Is That Chicken?" (Consumer Reports, Jan. 2010: 19–23) reported that 35 of 80 randomly selected Perdue brand broilers tested positively for either campylobacter or salmonella (or both), the leading bacterial causes of food-borne disease, whereas 66 of 80 Tyson brand broilers tested positive. a. Does it appear that the true proportion of noncontaminated Perdue broilers differs from that for the Tyson brand? Carry out a test of hypotheses using a significance level .01. b. If the true proportions of non-contaminated chickens for the Perdue and Tyson brands are .50 and .25, respectively, how likely is it that the null
hypothesis of equal proportions will be rejected when a .01 significance level is used and the sample sizes are both 80?

```r
n1 <- 300
x1 <- 63
n2 <- 180
x2 <- 75
# part A:
p1 <- x1 / n1
p2 <- x2 / n2
p_pool <- (x1 + x2) / (n1 + n2)
z_stat <- (p1 - p2) / (sqrt(p_pool * (1 - p_pool) * (1/n1 + 1/n2)))
p_value <- 2 * pnorm(-abs(z_stat))
cat("phat:", round(p_pool, 4), "\n")
```

```
## phat: 0.2875
```

```r
cat("Z-Statistic:", round(z_stat, 4), "\n")
```

```
## Z-Statistic: -4.8432
```

```r
cat("we Reject the null hypothesis: Significant difference exists \n")
```

```
## we Reject the null hypothesis: Significant difference exists
```

```r
# PART B:
sigma <- 0.0432
sigma_alt <- 0.0421  # Adjusted sigma from the formula
z_alpha <- 1.96      # Critical Z-value for alpha = 0.05
margin <- 0.2
z1 <- ((z_alpha * sigma_alt) + margin) / sigma
z2 <- ((-z_alpha * sigma_alt) + margin) / sigma
phi_z1 <- pnorm(z1)
phi_z2 <- pnorm(z2)
power <- 1 - (phi_z1 - phi_z2)
cat("Power of the Test =", round(power, 4), "\n")
```

```
## Power of the Test = 0.9967
```

Q13 9.5.63:

Toxaphene is an insecticide that has been identified as a pollutant in the Great Lakes ecosystem. To investigate the effect of toxaphene exposure on animals, groups of rats were given toxaphene in their diet. The article "Reproduction Study of Toxaphene in the Rat" (J. of Environ. Sci. Health, 1988: 101–126) reports weight gains (in grams) for rats given a low dose (4 ppm) and for control rats whose diet did not include the insecticide. The sample standard deviation for 23 female control rats was 32 g and for 20 female low-dose rats was 54 g. Does this data suggest that there is more variability in low-dose weight gains than in control weight gains? Assuming normality, carry out a test of hypotheses at significance level .05.

```r
s1 <- 54
s2 <- 32
n1 <- 20
n2 <- 23
alpha <- 0.05
F_stat <- (s1^2) / (s2^2)
df1 <- n1 - 1
df2 <- n2 - 1
p_value <- 1 - pf(F_stat, df1, df2)
cat("F-Statistic:", round(F_stat, 4), "\n")
```

```
## F-Statistic: 2.8477
```

```r
cat("P-value:", round(p_value, 4), "\n")
```

```
## P-value: 0.0101
```

```r
cat("We Reject the null hypothesis. There is more variability in low-dose weight gains.\n")
```

```
## We Reject the null hypothesis. There is more variability in low-dose weight gains.
```

Q14 9.5.66:

Reconsider the data of Example 9.6, and calculate a 95% upper confidence bound for the ratio of the standard deviation of the triacetate porosity distribution to that of the cotton porosity distribution.

Example 9.6 :

The void volume within a textile fabric affects comfort, flammability, and insulation properties. Permeability of a fabric refers to the accessibility of void space to the flow of a gas or liquid. The article "The Relationship Between Porosity and Air Permeability of Woven Textile Fabrics" (J. of Testing and Eval., 1997: 108–114) gave summary information on air permeability (cm3/cm2/sec) for a number of different fabric types. Consider the following data on two different types of plain-weave fabric:

```r
alpha <- 0.05
s_tri <- 3.59
s_cotton <- 0.79
n <- 10
df<- n - 1
n_cotton <- 10
df_cotton <- n_cotton - 1
F_critical <- qf(1 - alpha, df, df_cotton)
upper <- sqrt((s_tri^2 * F_critical) / (s_cotton^2))
cat("The CI :", round(upper, 4), "\n")
```

```
## The CI : 8.1022
```