Assignment 2

Tim But, Mehmet Bedirhan Gursoy, Vincentas Ryliskis, group 035

20 November 2023

Setting a Seed

In order to ensure that the results in this report are reproducible, seed (123) was used when generating random numbers.

```r
set.seed(123)
```

Task 1

a)

In order to achieve a 95% confidence interval, we will use a Z-value of $Z = 1.96$, and in order to achieve a margin of error of 2%, we will use a value of $E = 0.02$. Due to the fact that we are working with a proportion, our calculation for the number (n) of surveyed flights should look as follows:

$n = (Z\frac{\sqrt{p(1-p)}}{E})^2$

Because we do not have a value for p, we can assume that p is equal to 0.5 - the value at which $p(1-p)$ will have it's highest value. Therefore, after accounting for all values, we get:

$n = (1.96\frac{\sqrt{0.5(1-0.5)}}{0.02})^2 = (1.96\frac{\sqrt{0.25}}{0.02})^2 = 2401$

Therefore, we need 2401 samples in order to get the desired margin of error over the desired confidence interval

b)

As the desired confidence interval and Z score remain the same as in part a), we can keep the previously used values of $E = 0.02$ and $Z = 1.96$. Additionally, the formula for calculating n stays the same:

$n = (Z\frac{\sqrt{p(1-p)}}{E})^2$

However, we are given a value for p in this part: 0.9. For this value of p, the calculation for n looks as follows:

$n = (1.96\frac{\sqrt{0.9(1-0.9)}}{0.02})^2 = (1.96\frac{\sqrt{0.09}}{0.02})^2 = 864.36$

Therefore, we need 865 samples in order to get the desired margin of error over the desired confidence interval

Task 2

In task 2, we are given the following values:

$\bar{x}_1 = 1124.3$

$\bar{x}_2 = 1118.1$

$s_d = 57.8$

$n = 30$

In order to find a confidence interval for $\bar{x}_1 - \bar{x}_2$, we will use the formula for two dependent samples, as the two sets are of first and second-born twins:

$[\bar{d} - t_{n-1,\alpha/2}\frac{s_d}{\sqrt{n}}, \bar{d} + t_{n-1,\alpha/2}\frac{s_d}{\sqrt{n}}]$

Where $\bar{d} = \bar{x}_1 - \bar{x}_2 = 6.2$. Therefore, when we substitute the numbers, we get:

$[6.2 - t_{29,0.05}\frac{57.8}{\sqrt{30}}, 6.2 + t_{29,0.05}\frac{57.8}{\sqrt{30}}]$

Further evaluating the numbers (and critical t values) gives us:

$[6.2 - 1.699\frac{57.8}{\sqrt{30}}, 6.2 + 1.699\frac{57.8}{\sqrt{30}}] = [6.2 - 17.929, 6.2 + 17.929]$

Resulting in a confidence interval of:

$[-11.729, 6.2 + 24.129]$

Task 3

Test Claim: Whether there is a difference in the population mean volumes of first- and second-born babies.

$H_0 : \mu 1 = \mu 2, H_1 : \mu 1 \neq \mu 2, \alpha = 0.05$

In task 3, we are given the following values:

$\bar{x}_1 = 1131.3$

$\bar{x}_2 = 1123.8$

$s_1 = 129.0$

$s_2 = 127.2$

$n_1 = 25$

$n_2 = 20$

We will be using the following test statistic:

$t_2 = \frac{x_1 - x_2}{\sqrt{\frac{s_1^2}{n1} + \frac{s_2^2}{n2}}}$

When all values are accounted for, it looks as follows:

$t_2 = \frac{1131.3-1123.8}{\sqrt{\frac{129^2}{25} + \frac{127.2^2}{20}}} = \frac{7.5}{\sqrt{\frac{16641}{25} + \frac{16179.84}{20}}} = \frac{7.5}{\sqrt{1474.632}} = 0.195$

Under the null hypothesis, the test statistic has a t-distribution with $\tilde{n}$ degrees of freedom. Without a calculator, we can make the assumption that $\tilde{n} = min(n_1 - 1, n_2 - 1) = min(24, 19) = 19$.

$t_{19,0.05/2} = 2.093$

Since $0.195 > -2.093$ and $0.195 < 2.093$, we fail to reject H_0. There is not sufficient evidence to reject the claim that first- and second-born children have the same mean brain volume.

Assumptions Made:
- $\tilde{n} = min(n_1 - 1, n_2 - 1)$ - This is an estimate of the true value of $\tilde{n}$
- $H_0 : \mu 1 = \mu 2$ - This assumption was made in order to evaluate the data.
- Additionally, a t-distribution is assumed for the collected data.

Task 4

a)

```
# Load the data
birthweights <- read.table(file="/Users/bedirhangursoy/stat_assignments/Assignment 2/birthweig

# Check normality using the Shapiro-Wilk test
shapiro_test_result <- shapiro.test(birthweights$birthweight)

cat("Since the p-value is", shapiro_test_result$p.value, "(way bigger than 0.05) and W value is
```

Since the p-value is 0.8995395 (way bigger than 0.05) and W value is 0.9959499 (really close to 1), we fail to reject the null hypothesis, and that suggests that the data is consistent with a normal distribution

```
# Print the test result
print(shapiro_test_result)
```

```
Shapiro-Wilk normality test
```

data: birthweights$birthweight W = 0.99595, p-value = 0.8995

```
# Construct a bounded 96%-CI for mu
t_test_result <- t.test(birthweights, conf.level = 0.96)

# Print the confidence interval
cat("Bounded 96%-CI for mu:", t_test_result$conf.int[1], "to", t_test_result$conf.int[2], "\n")
```

Bounded 96%-CI for mu: 2808.084 to 3018.501

```
# Evaluate the required sample size to ensure CI length is at most 100
desired_ci_length <- 100
percentage <- 0.98

# Formula for sample size for a given CI length
required_sample_size <- (((qnorm(percentage) * sd(birthweights$birthweight)) / (desired_ci_leng
# Print the required sample size
cat("Required sample size for a CI length of", desired_ci_length, "is:", required_sample_size,
```

Required sample size for a CI length of 100 is: 205.2029 ## b)

```
t_test_result_claim <- t.test(birthweights$birthweight, mu = 2800, alternative = "greater")

cat("While lower endpoint of the confidence interval is a proper number", t_test_result_claim$c
In other words this is a one sided interval.")
```

While lower endpoint of the confidence interval is a proper number 2829.202 the higher endpoint of the confiednce interval is Inf . This occurs because we were trying to find a single number x which indicates that the mean is likely to be greater than or equal to 2829.202 . In other words this is a one sided interval. ## c)

```
p_L_hat <- 0.25
p_hat <- sum(birthweights$birthweight < 2600) / length(birthweights$birthweight)
E <- (p_hat - p_L_hat)
p_R_hat <- p_hat + E

confidence_interval_reconstructed <- c(p_L_hat, p_R_hat)

z <- (E / sqrt((p_hat * (1 - p_hat)) / length(birthweights$birthweight)))
confidence_level <- pnorm(z)
cat("Confidence Interval:", confidence_interval_reconstructed, ". Confidence level:", confidenc
```

Confidence Interval: 0.25 0.4095745 . Confidence level: 0.9900163

Task 5

a)

In order to give the an estimate for the difference of mean working time, and a 90% confidence interval, we used the code below. We assume that the samples are paired, as each sample for Alice has a corresponding sample for Bob for the same night (and vice versa). Therefore, if one of the two were to experience a change in working hours that day (eg. due to a late shipment, or leaving early due to a holiday), then the other would also likely experience a similar change. This has to be accounted for in the calculation. The findings of this code will be printed below.

```r
df = read.table('alice.txt',sep=' ')
alice = df[,1]
bob = df[,2]

t_test_result = t.test(alice, bob, conf.level = 0.9, paired = TRUE)

mean_difference = t_test_result$estimate

confidence_interval = t_test_result$conf.int

cat("The estimated mean difference is:", mean_difference, "\\newline The 90%",
    "confidence interval for the mean difference is: [",
    confidence_interval,"]")
```

The estimated mean difference is: 0.1544989
The 90% confidence interval for the mean difference is: [0.0194331 0.2895647]

b)

Because the test that would be appropriate for this section is the same as has been carried out in part a) (two tailed t-test of paired samples with a 90% confidence interval), all we need to do is get the p value, and see if it is smaller than $\alpha = 0.1$. The findings of this code will be printed below.

```r
alpha = 0.1

p_value = t_test_result$p.value

if (p_value < alpha) {
  cat("The calculated p-value is:", p_value,
      "\\newline Reject the  manager's (null) hypothesis. There is enough",
      "evidence to suggest that the mean working hours for Alice and Bob",
      " are different.")
} else {
  cat("The calculated p-value is:", p_value,
      "\\newline Fail to reject the manager's (null) hypothesis. There is",
      "not enough evidence to suggest a difference in the mean working hours",
      "for Alice and Bob.")
}
```

The calculated p-value is: 0.06097823
Reject the manager's (null) hypothesis. There is enough evidence to suggest that the mean working hours for Alice and Bob are different.

c)

In order to test Alice's hypothesis, we need to make a few adjustments to the code written in parts a) and b). Namely, we need to redefine $\alpha = 0.01$, and we also need to perform a one-tailed t-test on the data - with the alternative hypothesis that the true mean difference is greater than 0. Since the true mean difference is Alice's mean - Bob's mean, this ensures that we test specifically for Alice's hours. Then, as before, we must compare the new p-value with the α. The findings of this code will be printed below.

```
alpha = 0.01

# Conf.level doesn't need changing, but it's nice to see it reflect alpha
alice_t_test_result = t.test(alice, bob, conf.level = 0.99, paired = TRUE,
                             alternative = "greater")

alice_p_value = alice_t_test_result$p.value

if (alice_p_value < alpha) {
  cat("The calculated p-value is:", alice_p_value,
      "\\newline Reject the  manager's (null) hypothesis. There is enough",
      "evidence to suggest that Alice's mean working hours are greater than",
      "those of Bob")
} else {
  cat("The calculated p-value is:", alice_p_value, "\\newline Fail to reject",
      "the manager's (null) hypothesis. There is not enough evidence to",
      "suggest that Alice's mean working hours are greater than those of Bob.")
}
```

The calculated p-value is: 0.03048911
Fail to reject the manager's (null) hypothesis. There is not enough evidence to suggest that Alice's mean working hours are greater than those of Bob.

d)

More changes need to be enacted to the code from before. We will now be using the prop.test function, since we are testing the proportion of hours per evening that are above 3.8. As a result, now all that's needed is to calculate the proportion for Alice and Bob in a one-tailed test (similar to before), with the alternative hypothesis that the true mean difference is greater than 0. Additionally, we will need to set "paired" to false, as the hours are now assumed to be randomly sampled from different days. This will give us a p value, which we can compare to $\alpha = 0.01$, in order to see if the results are statistically significant. The findings of this code will be printed below.

```
alpha = 0.01
threshold = 3.8

alice_above = sum(alice > threshold)
bob_above = sum(bob > threshold)
total = length(alice)

prop_test = prop.test(c(alice_above, bob_above), c(total, total),
                      conf.level = 0.99, alternative = "greater")
```

```r
prop_p_value = prop_test$p.value


if (prop_p_value < alpha) {
  cat("The calculated p-value is:", prop_p_value,
      "\\newline Reject the null hypothesis. There is enough evidence to",
      "suggest thatthe proportion of evenings which Alice worked > 3.8 hours",
      "is greater than that of Bob")
} else {
  cat("The calculated p-value is:", prop_p_value, "\\newline Fail to reject",
      "the null hypothesis. There is not enough evidence to suggest a",
      "difference in proportions.")
}
```

The calculated p-value is: 0.01291652
Fail to reject the null hypothesis. There is not enough evidence to suggest a difference in proportions.