557 Homework 1 Solutions

Shrusti Ghela

January 20, 2022

Question 1

Suppose that you flip a coin 40 times and count the number of heads.

1.1. What is the distribution of the number of heads assuming the coin is fair? The number of heads is distributed Binomial($n = 40$, $p = 0.5$).

1.2. The sample proportion of heads has an approximately normal distribution. What are the mean and standard deviation of this distribution assuming the coin is fair? The mean of the sample proportion is

$$\mu = 0.5$$

The standard deviation is

$$\sigma = \sqrt{\frac{0.5(1-0.5)}{40}} = \frac{0.5}{2\sqrt{10}} = \frac{0.25}{\sqrt{10}} =$$

```
## [1] "sd = 0.079"
```

1.3. Define the Z-statistic for conducting a test of the null hypothesis that the coin is fair (i.e., has probability of a head equal to 0.5).

$$Z = \frac{\hat{p} - 0.5}{\sqrt{\frac{0.5 \cdot 0.5}{40}}} = \frac{\hat{p} - 0.5}{\frac{0.5}{\sqrt{40}}} = \frac{\hat{p} - 0.5}{0.079}$$

OR

$$Z = \frac{X - 20}{\sqrt{40 \cdot 0.5 \cdot 0.5}} = \frac{X - 20}{\sqrt{10}} = \frac{X - 20}{3.162}$$

1.4. Suppose the experiment results in 15 heads and 25 tails. Conduct a test of the null hypothesis with type I error probability 0.05 using the normal approximation. State the Z statistic, the p-value, and the conclusion of the test (do you reject the null hypothesis or not). Method 1, using the proportion of heads:

$$H_0 : p = 0.5$$

$$H_A : p \neq 0.5$$

Will conduct test with 95% significance, therefore $\alpha = 0.05$.

$$Z = \frac{\hat{p} - 0.5}{\sqrt{\frac{0.5 \cdot 0.5}{40}}} = \frac{0.375 - 0.5}{0.079} = -1.581$$

```
p_hat = 15 / n
z = (p_hat - 0.5) / sqrt(0.5*0.5/n)
p = 2 * pnorm(z)
print(paste0('Z = ', round(z, 3)))
```

```
## [1] "Z = -1.581"
```

```
print(paste0('p = ', round(p, 3)))
```

```
## [1] "p = 0.114"
```

Method 2, using the number of heads (yields the exact same Z-statistic and p-value):

$$H_0 : \mu = 20$$

$$H_A : \mu \neq 20$$

$$Z = \frac{X - 20}{\sqrt{10}} = \frac{15 - 20}{\sqrt{10}} = -1.581$$

```
n = 15 + 25
z = (15 - (n * 0.5)) / sqrt(n * 0.5 * 0.5)
p = 2 * pnorm(z)
print(paste0('Z = ', round(z, 3)))
```

```
## [1] "Z = -1.581"
```

```
print(paste0('p = ', round(p, 3)))
```

```
## [1] "p = 0.114"
```

Since $p = 0.114 > \alpha$ and $Z =$ -1.581 is within the acceptance region of $[-1.96, 1.96]$, we fail to reject the null hypothesis that the coin is unbiased.

It is **incorrect** to calculate the standard deviation for the Z statistic with $\hat{p}$, such as:

$$Z = \frac{\hat{p} - 0.5}{\sqrt{\frac{\hat{p}(1-\hat{p})}{40}}} = Z = \frac{0.375 - 0.5}{\sqrt{\frac{0.375 \cdot 0.625}{40}}} = \frac{-0.125}{\sqrt{0.0059}} = \frac{-0.125}{0.0765} = -1.633$$

This is because for a proportion test, we know that the underlying distribution (Bernoulli/Binomial) has a mean p and a standard deviation that is directly determined by p: $\sigma = \sqrt{p(1-p)}$. Therefore, under the null hypothesis that $p = 0.5$, we know the standard deviation under the null hypothesis. For tests about means, we do not have that certainty since the standard deviation is a separate parameter.

It is also **incorrect** to answer this question using `prop.test`. According to R's `prop.test` documentation, the default for this function is to apply continuity correction to the results, which will change your p-values. This is demonstrated below.

```r
# Incorrect p-value using continuity correction
print(
  prop.test(x = 15, n = 40, p = 0.5, alternative = 't', conf.level = 0.95)
)
```

```
##
##  1-sample proportions test with continuity correction
##
## data:  15 out of 40, null probability 0.5
## X-squared = 2.025, df = 1, p-value = 0.1547
## alternative hypothesis: true p is not equal to 0.5
## 95 percent confidence interval:
##  0.2317406 0.5419036
## sample estimates:
##     p
## 0.375
```

```r
# Correct p-value using continuity correction
print(
  prop.test(x = 15, n = 40, p = 0.5, alternative = 't', conf.level = 0.95, correct = FALSE)
)
```

```
##
##  1-sample proportions test without continuity correction
##
## data:  15 out of 40, null probability 0.5
## X-squared = 2.5, df = 1, p-value = 0.1138
## alternative hypothesis: true p is not equal to 0.5
## 95 percent confidence interval:
##  0.2422298 0.5296756
## sample estimates:
##     p
## 0.375
```

1.5. If you had decided to use a type I error probability of 0.1 instead of 0.05 would your conclusion be different? Explain. The conclusion would not be different because the p-value from the z-statistic earlier exceeds $\alpha = 0.1$, and the Z-statistic still falls within the acceptance region of [-1.645, 1.645]. Hence, we will still fail to reject the null hypothesis.

1.6. Calculate the p-value using the binomial distribution. Do you reach the same conclusion with the binomial distribution as with the normal approximation? Using the binomial distribution, we do not achieve a p-value great enough to reject the null hypothesis under $\alpha = 0.05$ nor $\alpha = 0.1$

```r
delta = 20 - 15 # 5
lower = 20 - delta # 15
upper = 20 + delta # 25
print(paste0(
  "p = ",
  round(sum(dbinom(c(0:lower, upper:40), 40, 0.5)), 3)
))
```

```
## [1] "p = 0.154"
```

1.7. Calculate a 95% confidence interval for the probability of a head using the normal approximation. Does the confidence interval include the value 0.5? Use $Z^* = 1.96$ to get the 95% confidence interval: $\hat{p} \pm Z^* SE$

```
z_star = qnorm(0.975)
n = 15 + 25
p_hat = 15 / n

se = sqrt(p_hat * (1 - p_hat) / n) # 0.0765

lower = p_hat - (z_star * se)
upper = p_hat + (z_star * se)
print(round(c(lower, upper), 3))
```

```
## [1] 0.225 0.525
```

This interval contains 0.5.

Note that in this case we do not use p in calculating the standard error and instead use $\hat{p}$.

1.8. Calculate a 90% confidence interval for the probability of a head using the normal approximation. How does it compare to the 95% confidence interval? Use $Z^* = 1.645$ to get the 90% confidence interval: $\hat{p} \pm Z^* SE$

```
z_star = qnorm(0.95)
lower = p_hat - (z_star * se)
upper = p_hat + (z_star * se)
print(round(c(lower, upper), 3))
```

```
## [1] 0.249 0.501
```

This interval also contains 0.5, though it is narrower than the 95% confidence interval.

Question 2

A study is done to determine if enhanced seatbelt enforcement has an effect on the proportion of drivers wearing seatbelts. Prior to the intervention (enhanced enforcement) the proportion of drivers wearing their seatbelt was 0.7. The researcher wishes to test the null hypothesis that the proportion of drivers wearing their seatbelt after the intervention is equal to 0.7 (i.e., unchanged from before). The alternative hypothesis is that the proportion of drivers wearing their seatbelt is not equal to 0.7 (either $<$ 0.7 or $>$ 0.7). After the intervention, a random sample of 400 drivers was selected and the number of drivers wearing their seatbelt was found to be 305.

2.1. Calculate the estimated standard error of the proportion of drivers wearing seatbelts after the intervention.

$$SE = \sqrt{\frac{\hat{p} \cdot (1 - \hat{p})}{n}} = \sqrt{\frac{0.7625 \cdot (0.2375)}{400}} = 0.021$$

```r
n = 400
p_hat = 305 / n

se = sqrt(p_hat * (1 - p_hat) / n)
print(paste0('SE = ', round(se, 3)))
```

```
## [1] "SE = 0.021"
```

2.2. Calculate a 95% confidence interval for the proportion of drivers wearing seatbelts after the intervention. What conclusion would you draw based on the confidence interval? Use $Z^* = 1.96$ to get the 95% confidence interval: $\hat{p} \pm Z^* SE$

```r
z_star = qnorm(0.975)
lower = p_hat - (z_star * se)
upper = p_hat + (z_star * se)
print(round(c(lower, upper), 3))
```

```
## [1] 0.721 0.804
```

Tthe lower bound of the confidence interval 0.721 is above the null hypothesis of 0.7 so we can conclude that the intervention did have an effect on proportion of drivers wearing seatbelts pre-intervention with 95% confidence.

2.3. Conduct a test of the null hypothesis with type I error probability 0.05 using the normal approximation. Should the null hypothesis be rejected? How does your conclusion compare to the conclusion from the confidence interval?

$$H_0 : p = 0.7$$

$$H_A : p \neq 0.7$$

$$Z = \frac{\hat{p} - 0.7}{\sqrt{\frac{0.7 \cdot 0.3}{400}}} = \frac{0.7625 - 0.7}{0.023} =$$

```r
z_stat = (305 - (n * 0.7)) / sqrt(n * 0.7 * (1 - 0.7))
print(paste0("z = ", round(z_stat, 3)))
```

```
## [1] "z = 2.728"
```

```r
p_val = 2 * (1 - pnorm(abs(z_stat)))
print(paste0("P-value = ", round(p_val, 4)))
```

```
## [1] "P-value = 0.0064"
```

Given that the p-value is less than $\alpha = 0.05$ and that $Z = 2.728$ is outside of the rejection region $[-1.96, 1.96]$, we reject the null hypothesis that the intervention did not have an effect on the proportion of drivers wearing seatbelts with 95% confidence.

Here, as well as in question 1, we know the underlying data is distributed Bernoulli/Binomial, hence we use p instead of $\hat{p}$ in calculating the denominator of the Z-statistic.

INCORRECT SOLUTION

```r
# R's prop.test automaticall defaults to performing continuity correction.
# The p-values it produces are incorrect for this assignment, as are its confidence intervals.
print(prop.test(x = 305, n = 400, p = 0.7, alternative = 't', conf.level = 0.95))
```

```
##
##  1-sample proportions test with continuity correction
##
## data:  305 out of 400, null probability 0.7
## X-squared = 7.1458, df = 1, p-value = 0.007514
## alternative hypothesis: true p is not equal to 0.7
## 95 percent confidence interval:
##  0.7171113 0.8027460
## sample estimates:
##      p
## 0.7625
```

```r
# Without correction
print(prop.test(x = 305, n = 400, p = 0.7, alternative = 't', conf.level = 0.95, correct = FALSE))
```

```
##
##  1-sample proportions test without continuity correction
##
## data:  305 out of 400, null probability 0.7
## X-squared = 7.4405, df = 1, p-value = 0.006377
## alternative hypothesis: true p is not equal to 0.7
## 95 percent confidence interval:
##  0.7184236 0.8015825
## sample estimates:
##      p
## 0.7625
```

```r
# Neither produces the correct confidence interval, but without correction
# produces correct p-val
```

2.4. Calculate the approximate p-value using the normal approximation and the exact p-value using the binomial distribution. Are the two p-values very different? We know the under the null hypothesis that $\mu = n \cdot p = 400 \cdot 0.7 = 280$.

The result of the experiment yields $\hat{\mu} = 305$.

```r
print(paste0(
  'Normal approximation p-val = ',
  round(p_val, 4)
  )
)
```

```
## [1] "Normal approximation p-val = 0.0064"
```

```r
delta = abs(n * 0.7 - 305) # 25
lower = n * 0.7 - delta # 255
upper = n * 0.7 + delta # 305
```

```r
print(paste0(
  'Exact p-val using Binomial dist = ',
  round(sum(dbinom(c(0:lower, upper:400), 400, 0.7)), 4)
  )
)
```

```
## [1] "Exact p-val using Binomial dist = 0.0074"
```

The p-value from using the binomial distribution is very similar the one resulting from normal approximation.

NOTE Using `binom.test` was acceptable for this answer, but it calculates the probability in a different way, but one that is also considered acceptable in the world of statistics. Instead of calculating the area under the pdf of the binomial distribution, which is *asymmetrical* in the case of $p = 0.7$, it takes the cumulative probability from $[305, 400]$ and multiplies by two. This is demonstrated below.

```r
# Returns p=0.00644
print(
  binom.test(x = 305, n = 400, p = 0.7, alternative = 't', conf.level = 0.95)
)
```

```
##
##  Exact binomial test
##
## data:  305 and 400
## number of successes = 305, number of trials = 400, p-value = 0.0063
## alternative hypothesis: true probability of success is not equal to 0.7
## 95 percent confidence interval:
##  0.7176897 0.8033774
## sample estimates:
## probability of success
##                 0.7625
```

```r
# This is equivalent to the below code
print(round(
  sum(dbinom(c(305:400), 400, 0.7)) * 2, 3
))
```

```
## [1] 0.006
```

This is a reasonable approximation though not 100% accurate because the binomial distribution is *asymmetric* under $p = 0.7$.
Interestingly, if you take a two-sided binomial test from `binom.test` using the lower range, 255, you get the correct answer. In this `binom.test` is taking the actual sum of the pdf under [0, 255] and [305, 400] and *not* just the sum under [0, 255] and doubling it.

```r
# Correct p-value
print(
  binom.test(x = 255, n = 400, p = 0.7, alternative = 't', conf.level = 0.95)
)
```

```
##
##  Exact binomial test
```

```
##
## data:  255 and 400
## number of successes = 255, number of trials = 400, p-value = 0.007444
## alternative hypothesis: true probability of success is not equal to 0.7
## 95 percent confidence interval:
##  0.5882595 0.6846892
## sample estimates:
## probability of success
##                 0.6375
```

```
# This is NOT the same as the below code
print(round(
  sum(dbinom(c(0:255), 400, 0.7)) * 2, 3
))
```

```
## [1] 0.008
```

2.5. Calculate the power of the test to detect the alternative hypothesis that the proportion of drivers wearing their seatbelt after the intervention is equal to 0.8. We are trying to find $P(Reject H_0 | p = 0.8) = P(X < 262 \cup X > 298 | p = 0.8)$.
How did I get those regions? By determining the rejection region under the null hypothesis.

```
# Find the rejection region under the null hypothesis with alpha = 0.05
lower = qbinom(0.025, 400, 0.7) # 262
# Get the difference from the null hypothesis average
upper = qbinom(0.975, 400, 0.7) #298
# Rejection region for two-sided test with alpha = .05 is between 0:262, 298:400
print(
  round(sum(dbinom(c(0:lower, upper:400), 400, 0.8)), 3)
)
```

```
## [1] 0.997
```

The power of the test is 99.7%, which is very high.

Question 3

The data come from a study of lead exposure and IQ in children. IQ scores were measured on a sample of children living in a community near a source of lead. The IQ scores were age-standardized using established normal values for the US population. Such age-standardized scores have a mean of 100 and a standard deviation of 15 in the US population.

3.1. Create a histogram of the IQ variable. Is the distribution approximately normal?

HW-1_Solutions_files/figure-latex/unnamed-chunk-17-1.pdf

This histogram shows that IQ scores are not normally distributed.
ALSO ACCEPTABLE: Any thoughtful justification

3.2. Calculate the sample mean and sample SD of IQ. How do they compare numerically to the US population values? $\bar{x} = 91.081$
$s = 14.404$

The mean IQ is lower and the standard deviation in IQ scores is higher than the US population values.

3.3. Test the null hypothesis that the mean IQ score in the community is equal to 100 using the 2-sided 1-sample t-test with a significance level of 0.05. State the value of the test statistic and whether or not you reject the null hypothesis at significance level 0.05.

$$H_0 : \mu = 100$$

$$H_A : \mu \neq 100$$

$$\alpha = .05$$

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

```
z = (mean_iq - 100) / (sd_iq / sqrt(nrow(iq)))
t_crit = qt(.025, df = nrow(iq) - 1)
print(paste0('Test statistic = ', round(z, 3)))
```

```
## [1] "Test statistic = -6.895"
```

```
print(paste0('T-critical = ', round(t_crit, 3)))
```

```
## [1] "T-critical = -1.979"
```

$t =$ -6.895 $< t_{.05,123} =$ -1.979

Since our calculated t is less than our critical T statistic with $\alpha = .05$ and 123 degrees of freedom, we reject the null hypothesis that the mean IQ of this sample is 100.

3.4. Give the p-value for the test in the previous question. State the interpretation of the p-value. Calculate this result using the the T-distribution with $n - 1$ degrees of freedom.

```
p_val = 2 * pt(z, df = nrow(iq) - 1)
print(paste0('p = ', round(p_val, 4)))
```

```
## [1] "p = 0"
```

The p-value for this test is $< .001$ **which means** there is < 0.001 probability of seeing a test statistic as extreme or more extreme than our sample mean under the null hypothesis.

3.5. Compute a 95% confidence interval for the mean IQ. Do the confidence interval and hypothesis test give results that agree or conflict with each other? Explain. Use $t^* = 1.979$ to find $\hat{IQ} \pm t^*_{df=123} SE$

```
t_crit = abs(qt(0.025, df = nrow(iq) - 1))
se = sd_iq / sqrt(nrow(iq))
lower = mean_iq - t_crit * se
upper = mean_iq + t_crit * se
print(round(c(lower, upper), 3))
```

```
## [1] 88.520 93.641
```

These results agree with the previous questions' results. Since the null hypothesis IQ is not contained within the 95% confidence interval, we can reject the null hypothesis that $\bar{IQ} = 100$.

Since we have a small sample size, it is incorrect to use 1.96 as your critical value to calculate the confidence interval.

3.6. Repeat the hypothesis test and confidence interval using a significance level of 0.01 and a 99% confidence interval. Use $t^* = 2.616$ to find $\hat{IQ} \pm t^*_{df=123} SE$

```
t_crit = abs(qt(0.005, df = nrow(iq) - 1))
se = sd_iq / sqrt(nrow(iq))
lower = mean_iq - t_crit * se
upper = mean_iq + t_crit * se
print(round(c(lower, upper), 3))
```

```
## [1] 87.696 94.465
```

Here, too, we reject the null hypothesis because it is not contained within the confidence interval.