

DSC 152 HW1 Answer Key Guide

Question 1

We are testing Rosa Smith's true 3-point percentage p with

$$H_0 : p = 0.30 \quad \text{vs.} \quad H_A : p > 0.30.$$

This is an exact binomial setting: each attempt is a make/miss outcome, the number of attempts is fixed, and the alternative is one-sided upper-tail.

1(a)

Full solution

After 4 games, Smith has attempted $n = 9$ threes. Let X be the number she made.

Under the null,

$$X \sim \text{Binomial}(n = 9, p = 0.30).$$

Because $H_A : p > 0.30$, larger values of X are more supportive of the alternative. So any valid rejection region must be of the form

$$\{c, c + 1, \dots, 9\}.$$

The question asks for the significance level closest to 0.05. So we compare exact upper-tail probabilities under H_0 :

$$P(X \geq 5) = 0.09880866$$

$$P(X \geq 6) = 0.02529484$$

$$P(X \geq 7) = 0.004290894.$$

Now compare distances from 0.05:

- $|0.09880866 - 0.05| = 0.04880866$
- $|0.02529484 - 0.05| = 0.02470516$
- $|0.004290894 - 0.05| = 0.045709106$

So the closest exact level is $P(X \geq 6)$. Therefore the rejection region is

$$\boxed{\{6, 7, 8, 9\}}.$$

The actual significance level is

$$\alpha = P(X \geq 6 \mid p = 0.30) = 0.02529484.$$

Interpretation in words

If Smith's true 3-point percentage were really 30%, then using the rule "reject H_0 when she makes at least 6 of 9" would incorrectly reject the null about 2.53% of the time. We cannot make the exact binomial significance level equal to 0.05 here because the binomial distribution is discrete, so the attainable tail probabilities jump.

R code

```
1 # Number of attempts and null shooting percentage
2 n <- 9
3 p0 <- 0.30
4
5 # For a one-sided upper-tail test, rejection regions are of the form {c, c+1, ...,
6   9}.
7 # Because pbinom(q, ..., lower.tail = FALSE) gives P(X > q),
8 # to compute P(X >= k) we must use q = k - 1.
9
10 # Exact significance level if we reject for X >= 5
11 alpha_ge_5 <- pbinom(4, size = n, prob = p0, lower.tail = FALSE)
12
13 # Exact significance level if we reject for X >= 6
14 alpha_ge_6 <- pbinom(5, size = n, prob = p0, lower.tail = FALSE)
15
16 # Exact significance level if we reject for X >= 7
17 alpha_ge_7 <- pbinom(6, size = n, prob = p0, lower.tail = FALSE)
18
19 alpha_ge_5
20 alpha_ge_6
21 alpha_ge_7
```

1(b)

Full solution

Now suppose Smith's true 3-point percentage is actually $p = 0.375$. The power is

$$P(\text{reject } H_0 \mid p = 0.375).$$

From part (a), rejection means $X \geq 6$. So

$$\text{Power} = P(X \geq 6 \mid X \sim \text{Binomial}(9, 0.375)).$$

Numerically,

$$\text{Power} = 0.07404184.$$

Interpretation in words

This is only about 7.4% power, which is extremely low. So even if Smith really had improved to a 37.5% 3-point shooter, a test based on only 9 attempts would detect evidence for improvement only about 7 times out of 100. This is bad power, and the reason is the tiny sample size.

R code

```
1 # Power is P(reject H0 | p = 0.375).
2 # From part (a), reject when X >= 6.
3 # Again, P(X >= 6) = pbinom(5, ..., lower.tail = FALSE)
4
5 power_q1b <- pbinom(5, size = 9, prob = 0.375, lower.tail = FALSE)
6 power_q1b
```

1(c)

Full solution

If Smith has made 5 out of 9 3-point attempts, then her success percentage is approximately 55.56%, which is well into the alternative hypothesis. However, as we can see from part (a), 5 successes is not in the rejection region. Specifically, the p-value would be obtained with:

```
1 pbinom(4, size=9, prob=0.3, lower.tail=FALSE)
```

which gives a p-value of 0.09880866 (the same as the value of $P(X \geq 5)$ that was shown in part (a)).

1(d)

Full solution

The HW provides the rejection cutoffs whose exact significance levels are closest to 0.05 for sample sizes 10, 20, ..., 150:

```
1 rej_regions_cutoff <- c(6, 10, 14, 17, 21, 24, 28, 31, 35, 38, 42, 45, 48, 52, 55)
```

For a given sample size n , if the cutoff is c_n , then power at a true percentage p is

$$P(X \geq c_n \mid X \sim \text{Binomial}(n, p)).$$

The three effect sizes shown in the HW figure are $p = 0.325$, 0.35 , and 0.375 . A correct recreation should show:

- for any fixed n , the curve for $p = 0.375$ is highest, then $p = 0.35$, then $p = 0.325$
- power generally increases with sample size
- the curves are not perfectly smooth because the exact binomial test is discrete, so the nominal significance level changes slightly as n changes.

A few benchmark values that a correct implementation should be close to are:

- at $n = 10$: powers are about 0.0683, 0.0949, 0.1275 for $p = 0.325, 0.35, 0.375$
- at $n = 50$: powers are about 0.1013, 0.1861, 0.3017
- at $n = 100$: powers are about 0.1432, 0.2976, 0.4966
- at $n = 150$: powers are about 0.1581, 0.3633, 0.6135.

Interpretation in words

The main takeaway is that power rises with both sample size and effect size. Small effects, like 0.325 versus 0.30, remain difficult to detect even at $n = 150$. Larger effects, like 0.375 versus 0.30, become much easier to detect as n grows.

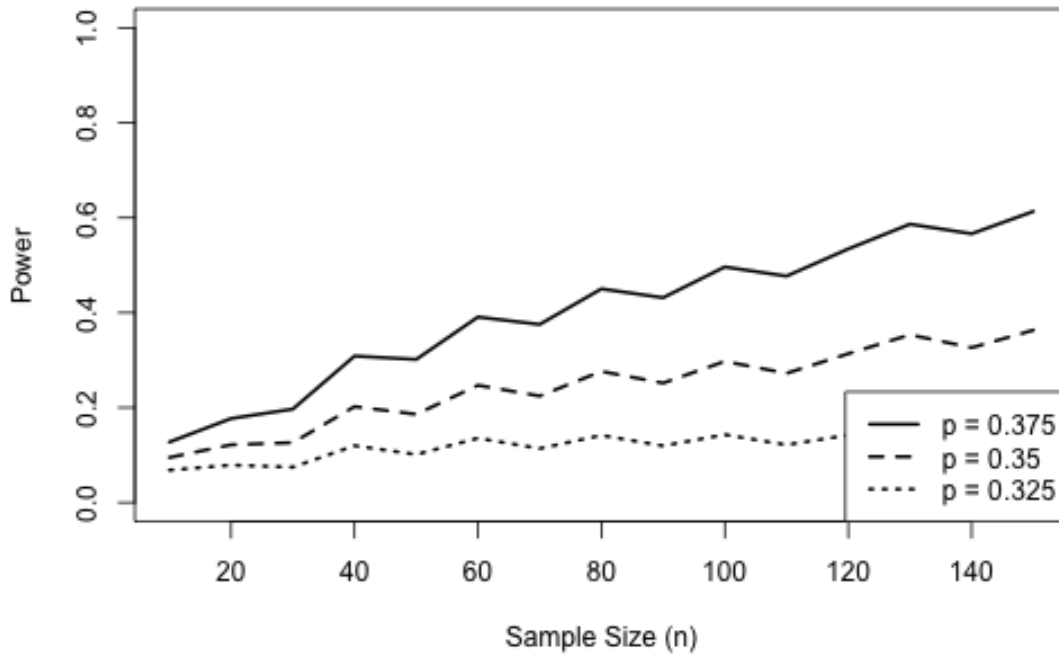
R code

```

1 # Sample sizes used in the HW figure
2 sample_sizes <- seq(10, 150, by = 10)
3
4 # Provided rejection cutoffs whose exact alpha levels are closest to 0.05
5 rej_regions_cutoff <- c(6,10,14,17,21,24,28,31,35,38,42,45,48,52,55)
6
7 # Compute power for each sample size at true p = 0.325
8 power_0325 <- mapply(function(n, cutoff){
9   # Reject when X >= cutoff
10  pbinom(cutoff - 1, size = n, prob = 0.325, lower.tail = FALSE)
11 }, sample_sizes, rej_regions_cutoff)
12
13 # Compute power for each sample size at true p = 0.35
14 power_035 <- mapply(function(n, cutoff){
15   pbinom(cutoff - 1, size = n, prob = 0.35, lower.tail = FALSE)
16 }, sample_sizes, rej_regions_cutoff)
17
18 # Compute power for each sample size at true p = 0.375
19 power_0375 <- mapply(function(n, cutoff){
20   pbinom(cutoff - 1, size = n, prob = 0.375, lower.tail = FALSE)
21 }, sample_sizes, rej_regions_cutoff)
22
23 # Plot the three power curves
24 plot(sample_sizes, power_0375, type = "l", lwd = 2,
25       ylim = c(0, 1),
26       xlab = "n",
27       ylab = "power",
28       main = "Power Curves over various Effect Sizes")
29
30 lines(sample_sizes, power_035, lwd = 2, lty = 2)
31 lines(sample_sizes, power_0325, lwd = 2, lty = 3)
32
33 legend("bottomright",
34       legend = c("p = 0.375", "p = 0.35", "p = 0.325"),
35       lwd = 2,
36       lty = c(1, 2, 3))

```

Power Curves over Various Effect Sizes



1(e)

Full solution

The 3FG/A column stores strings like 1-2, meaning “1 made out of 2 attempted.” So we must split that column into two numeric pieces, one for makes and one for attempts, then total across the whole season. The homework explicitly asks for that parsing step.

From the smith.txt data:

- total 3-point makes = 57
- total 3-point attempts = 162

So the sample proportion is

$$\hat{p} = \frac{57}{162} = 0.3518519.$$

Now use the same exact test as in Question 1:

$$H_0 : p = 0.30 \quad \text{vs.} \quad H_A : p > 0.30.$$

Under H_0 , the total number of makes X across all 162 attempts satisfies

$$X \sim \text{Binomial}(162, 0.30).$$

The exact one-sided p-value is

$$P(X \geq 57 \mid X \sim \text{Binomial}(162, 0.30)) = 0.08914555.$$

So the required answers are:

$$\hat{p} = 0.3518519$$

$$p\text{-value} = 0.08914555$$

Interpretation in words

Smith's observed season 3-point percentage was about 35.19%, which is higher than 30%, but the exact one-sided p-value is about 0.0891. At the usual $\alpha = 0.05$ level, we fail to reject H_0 . So there is not enough statistical evidence to conclude that her true 3-point percentage was greater than 30%.

R code

```
1 # Read the tab-delimited data file
2 smith <- read.delim("smith.txt", sep = "\t")
3
4 # Split the 3FG/A string column at the "-" character
5 three_split <- strsplit(as.character(smith$'3FG/A'), "-")
6
7 # Extract the number made and number attempted as numeric vectors
8 smith$made_3 <- as.numeric(sapply(three_split, '[', 1))
9 smith$att_3 <- as.numeric(sapply(three_split, '[', 2))
10
11 # Total makes and attempts across the full season
12 total_made <- sum(smith$made_3)
13 total_att <- sum(smith$att_3)
14
15 # Sample proportion made
16 phat <- total_made / total_att
17 phat
18
19 # Exact one-sided binomial p-value
20 # We observed X = total_made makes, so p-value = P(X >= total_made)
21 # = pbinom(total_made - 1, ..., lower.tail = FALSE)
22 pval_q1d <- pbinom(total_made - 1,
23                   size = total_att,
24                   prob = 0.30,
25                   lower.tail = FALSE)
26
27 pval_q1d
```

Question 2

Question 2 uses the AST column and explicitly says the distribution looks fairly skewed, so a t-test may not be

valid; instead students should use a Sign Test for the median. The HW also explicitly says the only condition required for the sign test is independence.

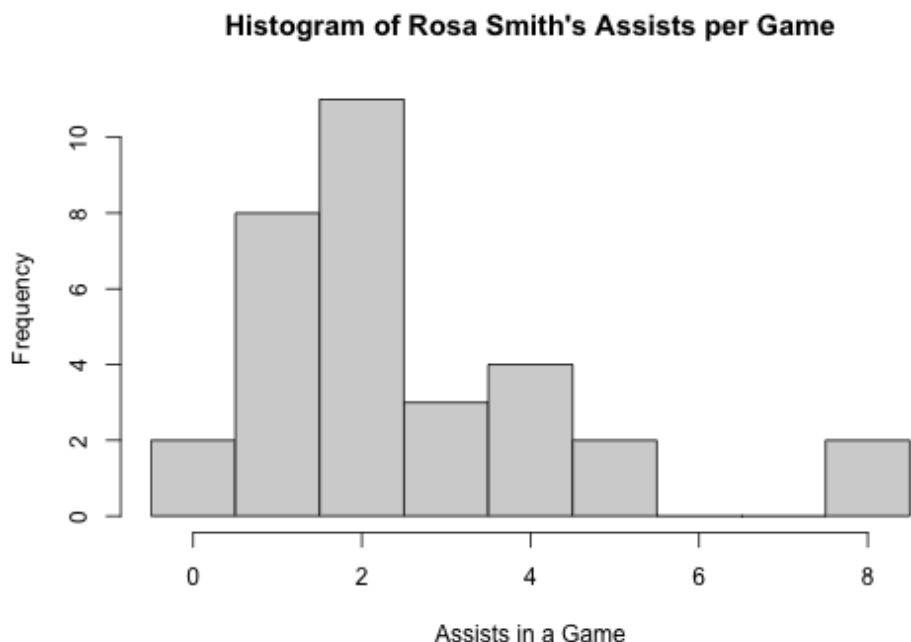
2(a)

Full solution

A correct histogram of AST should show most games clustered at low assist totals, especially around 1 and 2, with a smaller number of larger values such as 5 and 8. So the shape is right-skewed, which is consistent with the HW's instruction not to rely on a t-test here.

R code

```
1 # Plot a histogram of assists per game
2 # Using breaks centered on integers makes the histogram easier to read
3 hist(smith$AST,
4       breaks = seq(-0.5, max(smith$AST) + 0.5, by = 1),
5       main = "Histogram of Rosa Smith's Assists per Game",
6       xlab = "Assists in a Game",
7       ylab = "Frequency")
```



2(b)

Full solution

We are testing

$$H_0 : \tilde{\mu} = 1.5 \quad \text{vs.} \quad H_A : \tilde{\mu} > 1.5,$$

where $\tilde{\mu}$ is Smith's true median assists per game. This is exactly a one-sample sign test situation. Under H_0 ,

if the true median is 1.5 and the observations are independent, values above and below 1.5 should be equally likely.

Let Y be the number of games with assists above 1.5. Because AST is integer-valued, there are no ties at 1.5. In the uploaded data:

- $n = 32$ games
- $Y = 22$ games with $AST > 1.5$
- 10 games with $AST < 1.5$.

Under H_0 ,

$$Y \sim \text{Binomial}(32, 0.5).$$

To find the rejection region whose significance level is closest to 0.05, compare:

$$P(Y \geq 20) = 0.10766357$$

$$P(Y \geq 21) = 0.05509208$$

$$P(Y \geq 22) = 0.02505123.$$

The one closest to 0.05 is $P(Y \geq 21)$. So the rejection region is

$$\boxed{Y \geq 21}$$

and the actual significance level is

$$\boxed{\alpha = 0.05509208.}$$

For the observed data, the one-sided p-value is

$$P(Y \geq 22) = 0.02505123.$$

Thus,

$$\boxed{p\text{-value} = 0.02505123.}$$

Interpretation in words

At the 5% level, we reject H_0 and conclude there is evidence that Smith's true median number of assists is greater than 1.5.

R code

```

1 # Extract the assists column
2 ast <- smith$AST
3
4 # Number of observations strictly above 1.5
5 # Since AST is integer-valued, there are no ties at 1.5
6 y <- sum(ast > 1.5)
7
8 # Effective sample size for the sign test
9 n_eff <- sum(ast != 1.5)
10
11 y
12 n_eff
13
14 # Check nearby exact one-sided significance levels
15 alpha_ge_20 <- pbinom(19, size = n_eff, prob = 0.5, lower.tail = FALSE) # P(Y >=
    20)
16 alpha_ge_21 <- pbinom(20, size = n_eff, prob = 0.5, lower.tail = FALSE) # P(Y >=
    21)
17 alpha_ge_22 <- pbinom(21, size = n_eff, prob = 0.5, lower.tail = FALSE) # P(Y >=
    22)
18
19 alpha_ge_20
20 alpha_ge_21
21 alpha_ge_22
22
23 # One-sided sign-test p-value for the observed count y = 22
24 pval_q2b <- pbinom(y - 1, size = n_eff, prob = 0.5, lower.tail = FALSE)
25 pval_q2b

```

2(c)

Full solution

The sign test requires the observations to be independent. That is not guaranteed here, because these are repeated games from the same player, and one game may be related to another through opponent strength, playing time, health, fatigue, team strategy, or momentum over the season. On the other hand, the games are separate events played on different dates, so treating them as approximately independent may be a reasonable simplification. Overall, independence is plausible but not certain.

Question 3

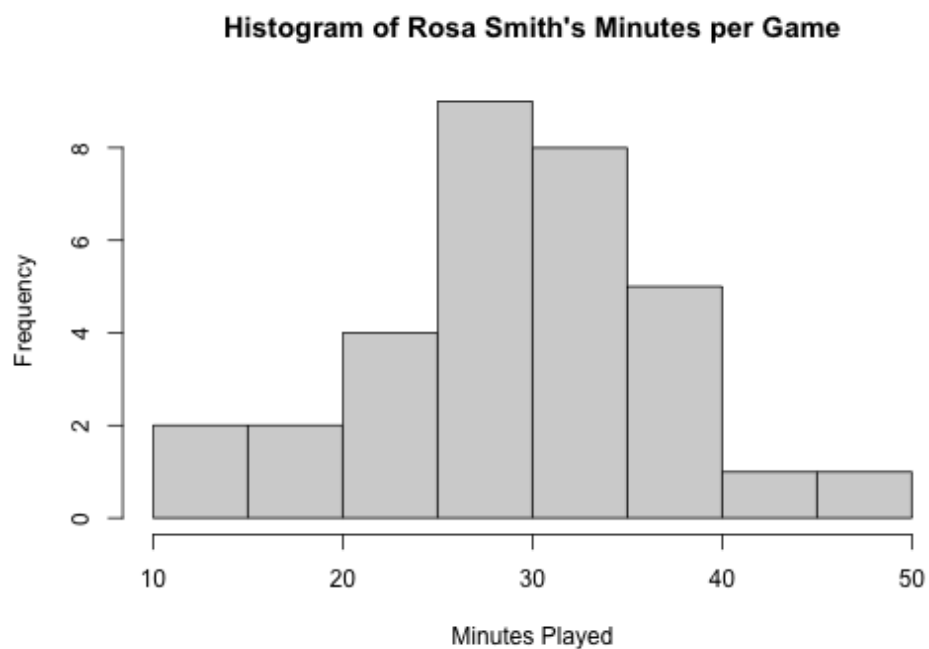
3(a)

Full solution

A correct histogram of minutes should show a distribution centered around the low 30s, much closer to symmetric and bell-shaped than the assists histogram. So a t-test is much more defensible here than it was in Question 2.

R code

```
1 # Plot a histogram of minutes played per game
2 hist(smith$MIN,
3     breaks = 10,
4     main = "Histogram of Rosa Smith's Minutes per Game",
5     xlab = "Minutes Played",
6     ylab = "Frequency")
```



3(b)

Part 1: one-sample t-test

We test

$$H_0 : \mu = 33.5 \quad \text{vs.} \quad H_A : \mu \neq 33.5.$$

From the uploaded data:

- $n = 32$
- $\bar{x} = 30.125$
- $s = 8.6351$.

Using the one-sample t statistic,

$$t_s = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{30.125 - 33.5}{8.6351/\sqrt{32}} \approx -2.211.$$

With $df = 31$, the two-sided p-value is

$$0.0345455.$$

So at $\alpha = 0.05$, we reject H_0 and conclude there is evidence that Smith's true mean minutes differed from 33.5. Since the sample mean is below 33.5, the data point toward the average being lower than that target.

Part 2: sign test

We also test

$$H_0 : \tilde{\mu} = 33.5 \quad \text{vs.} \quad H_A : \tilde{\mu} \neq 33.5.$$

Let:

- N_+ = number of games with MIN > 33.5
- N_- = number of games with MIN < 33.5.

Because minutes are integer-valued in the file, there are no ties at 33.5. In the data:

- $N_+ = 13$
- $N_- = 19$
- $n_{\text{eff}} = 32$.

Under H_0 , the number above 33.5 follows

$$Y \sim \text{Binomial}(32, 0.5).$$

For the two-sided sign test, use the smaller tail and double it:

$$p\text{-value} = 2P(Y \leq 13) = 0.37708559.$$

So

$$p\text{-value} = 0.3770856.$$

Thus we fail to reject H_0 with the sign test. There is not enough evidence that Smith's true median minutes differed from 33.5.

Comparison/Comment

These results are not contradictory, because the tests target different parameters:

- the t-test tests the mean
- the sign test tests the median.

Also, the sign test uses only whether each value is above or below 33.5, while the t-test uses the actual magnitudes of the deviations. So the sign test throws away information and is usually less powerful when the t-test conditions are satisfied.

R code

```
1 # Extract the minutes column
2 mins <- smith$MIN
3
4 # -----
5 # One-sample t-test for the mean
6 # -----
7 ttest_result <- t.test(mins, mu = 33.5, alternative = "two.sided")
8 ttest_result
9
10 # -----
11 # Two-sided sign test for the median
12 # -----
13 # Count how many games are above and below 33.5
14 num_above <- sum(mins > 33.5)
15 num_below <- sum(mins < 33.5)
16
17 # Effective sample size excludes ties
18 n_eff <- num_above + num_below
19
20 num_above
21 num_below
22 n_eff
23
24 # Exact two-sided sign-test p-value:
25 # double the smaller binomial tail
26 pval_sign_q3b <- 2 * pbinom(min(num_above, num_below),
27                             size = n_eff,
28                             prob = 0.5)
29
30 # In general, cap at 1
31 pval_sign_q3b <- min(1, pval_sign_q3b)
32 pval_sign_q3b
```

3(c)

Full solution

This is the only part with many possible correct answers. The HW asks for one scenario in which the t-test has at least 20 percentage points higher power than the sign test, using the same null value. It also says students may use a p-value cutoff of 0.05 for both.

A clean, fully acceptable scenario is:

- generate data from a normal distribution
- true mean = 0.5
- true standard deviation = 1

- sample size $n = 30$
- null value = 0
- use two-sided tests.

This is a good choice because under normality the t-test assumptions are satisfied, and for a symmetric normal distribution the mean and median shift together. So it is a fair comparison that should favor the t-test.

With a reproducible seed and many simulations, you should get power estimates roughly around:

- t-test power ≈ 0.75
- sign-test power ≈ 0.54 to 0.55 .

That is comfortably more than a 20 percentage point gap, so it satisfies the prompt. Because this is simulation, exact values vary a little by seed and number of repetitions. The logic is what matters.

R code

```

1 # Set a seed so the simulation is reproducible
2 set.seed(152)
3
4 # Number of simulation repetitions
5 B <- 20000
6
7 # Choose a scenario where the t-test assumptions are met
8 # True data distribution: Normal(mean = 0.5, sd = 1)
9 # Null value for both tests: 0
10 n <- 30
11 mu_true <- 0.5
12 sigma_true <- 1
13 m0 <- 0
14
15 # Helper function: exact two-sided sign-test p-value
16 sign_test_pval <- function(values, m0){
17   num_above <- sum(values > m0)
18   num_below <- sum(values < m0)
19   n_eff <- num_above + num_below
20
21   # If every value ties, return p-value 1
22   if(n_eff == 0){
23     return(1)
24   }
25
26   # Two-sided exact sign-test p-value
27   p_val <- 2 * pbinom(min(num_above, num_below),
28                       size = n_eff,
29                       prob = 0.5)
30
31   return(min(1, p_val))
32 }
33
34 # Counters for rejections

```

```

35 t_reject_count <- 0
36 sign_reject_count <- 0
37
38 for(i in 1:B){
39   # Simulate one sample from the chosen alternative distribution
40   x <- rnorm(n, mean = mu_true, sd = sigma_true)
41
42   # Two-sided one-sample t-test
43   p_t <- t.test(x, mu = m0, alternative = "two.sided")$p.value
44   if(p_t < 0.05){
45     t_reject_count <- t_reject_count + 1
46   }
47
48   # Two-sided sign test
49   p_sign <- sign_test_pval(x, m0)
50   if(p_sign < 0.05){
51     sign_reject_count <- sign_reject_count + 1
52   }
53 }
54
55 # Estimated powers
56 power_t <- t_reject_count / B
57 power_sign <- sign_reject_count / B
58
59 power_t
60 power_sign

```

Model write-up

I simulated samples of size 30 from a normal distribution with true mean 0.5 and true standard deviation 1, and I tested against null value 0 using two-sided tests. In this setting the t-test assumptions are met. My estimated power was about 0.75 for the t-test and about 0.54 to 0.55 for the sign test, so the t-test had over 20 percentage points more power. This makes sense because the t-test uses the full numerical information in the sample, while the sign test only uses whether observations are above or below the null value.