Midterm 1 Preparation

DSC 152: Week 3 Discussion

Zida Jin

University of California, San Diego

2026.10.12

All data in this class are simulated for practice.

Scenario: e-scooters on campus

From Autumn 2026, UC San Diego requires every e-scooter rider to register, including pass a safety tutorial and show they own a helmet. The university cited rising injuries and speed-related incidents, among other reasons.

- Campus records: as comparison, long-run **e-bike** incident rate $= 2.0$ incidents per 1,000 rides (treat as fixed)
- The safety office samples days at random and records each day's **e-scooter** incident rate in `scooter_days$rate`
- Daily rates are right-skewed

Definitions

μ = true mean daily e-scooter incident rate. $H_0: \mu = 2$

Simulated data. Rule details: transportation.ucsd.edu/micromobility/scooter-registration.html

Q1(a): t-statistic and p-value

3 minutes

A sample of $n = 9$ days gives $\bar{x} = 2.6$ and $s = 0.9$. The alternative is H_A: $\mu \neq 2$.

1. Compute the test statistic t.
2. Write **one line of R** that gives the two-sided p-value without using `t.test()`.

Reference: `pt(-2, df = 8)` $= 0.0403$

Q1(a): Solution

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{2.6 - 2}{0.9/\sqrt{9}} = \frac{0.6}{0.3} = 2, \qquad \text{df} = n - 1 = 8$$

```
2 * pt(-2, df = 8)        # 0.0806
```

Also accepted:

```
2 * pt(2, df = 8, lower.tail = FALSE)
```

Why double?

Two-sided: values at least as extreme in *either* direction, $t \leq -2$ or $t \geq 2$.

Review: Lec02 pp. 6–10

Q1(b): Interpreting the p-value

3 minutes

The two-sided p-value is about **0.081**. ☐ **Select all that apply.**

A. We reject H_0 at $\alpha = 0.10$ but not at $\alpha = 0.05$.

B. There is a 0.081 probability that H_0 is true.

C. If H_0 is true and the conditions are met, the probability of getting $t \geq 2$ or $t \leq -2$ is about 0.081.

D. For H_A: $\mu > 2$, the p-value would be about 0.040.

E. For H_A: $\mu < 2$, the p-value would be about 0.040.

Q1(b): Solution

✓ **A** $0.081 < 0.10$, but $0.081 > 0.05$.

✗ **B** A p-value is computed *assuming* H_0 is true, so it cannot be the probability that H_0 is true.

✓ **C** This is the definition of the p-value.

✓ **D** $\bar{x} = 2.6$ lies in the direction of H_A, so $p = P(T \geq 2) = \frac{1}{2}P(|T| \geq 2) = 0.0403$.

✗ **E** $\bar{x}$ lies in the opposite direction, so $p = P(T \leq 2) = 1 - 0.0403 = 0.960$.

Review: Lec01 pp. 17, 27; Lec05 p. 31

Q2(a): Simulating the Type I error rate

3 minutes

Suppose daily rates truly follow a Gamma distribution with parameters `shape = 2`, `scale = 1`. Its mean is 2, so H_0 is true.

Fill in the blanks to estimate the t-test's Type I error rate for $n = 9$ at $\alpha = 0.10$.

```
count <- 0
for(i in 1:10000){
  sim_rates <- __(a)__
  p.val <- __(b)__
  if(__(c)__){ count <- count + 1 }
}
TypeI <- count / 10000
```

Q2(a): Solution

```
count <- 0
for(i in 1:10000){
  sim_rates <- rgamma(n = 9, shape = 2, scale = 1)       # (a)
  p.val <- t.test(sim_rates, mu = 2)$p.value             # (b)
  if(p.val < 0.10){ count <- count + 1 }                 # (c)
}
TypeI <- count / 10000
```

Logic

Generate data where H_0 is *true*, run the test, and count how often it (wrongly) rejects.

Review: Lec02 pp. 21–27

Q2(b): t-test against bootstrap

2 minutes

The code returns about **0.127**. A colleague repeats the simulation with a **90% bootstrap percentile CI test** (reject when 2 lies outside the interval) and gets about **0.18**.

◯ **Select one.**

A. Both tests are valid, because their rates are close to 0.10.

B. Only the bootstrap test is valid, because it is non-parametric.

C. Both tests have an inflated Type I error rate.

D. Both tests have a deflated Type I error rate.

CI = confidence interval.

Q2(b): Solution

✗ **A** A valid test has Type I error $= \alpha$. 0.127 and 0.18 are clearly above 0.10.

✗ **B** Non-parametric does *not* mean valid; the bootstrap is the *worse* of the two here.

✓ **C** Both exceed 0.10. With $n = 9$ from a skewed population, the t-test's normality condition fails, and the sample is too small to represent the population for the bootstrap.

✗ **D** Neither rate is below 0.10.

Tips

Increasing n brings the t-test's rate towards 0.10, because $\bar{x}$ becomes approximately normal (Central Limit Theorem).

Review: Lec02 pp. 14–15; Lec03 pp. 18–23

Q3(a): Sign sequence and p-value

3 minutes

Nine daily rates are recorded:

$$3.1,\ 2.4,\ 4.6,\ 1.2,\ 2.0,\ 2.7,\ 3.8,\ 2.2,\ 5.9$$

Use **2.0** as the null value.

1. Write the sign sequence.
2. State the effective sample size.
3. Write R code for the two-sided p-value.

Reference ($n = 8$, $p = 0.5$):

`pbinom(0, 8, 0.5)` $= 0.0039$ `pbinom(1, 8, 0.5)` $= 0.0352$ `pbinom(2, 8, 0.5)` $= 0.1445$

Q3(a): Solution

Rate	3.1	2.4	4.6	1.2	2.0	2.7	3.8	2.2	5.9
Sign	+	+	+	−	0	+	+	+	+

7 "+" and 1 "−" (the tie is dropped), so the effective sample size is $n = 8$.

```
2 * pbinom(1, size = 8, prob = 0.5)     # 0.0703
```

Review: Lec03 pp. 6–10

Q3(b): Choosing the rejection region

3 minutes

Before collecting data, the office chose the two-sided rejection region whose Type I error is closest to 0.05.

1. Name that region.
2. State its α.
3. Give the decision for this sample.

Reference ($n = 8$, $p = 0.5$):

`pbinom(0, 8, 0.5)` $= 0.0039$ `pbinom(1, 8, 0.5)` $= 0.0352$ `pbinom(2, 8, 0.5)` $= 0.1445$

Q3(b): Solution

Rejection region for X	Type I error	Distance from 0.05
$\{0, 8\}$	$2 \times 0.0039 = 0.0078$	0.042
$\{0, 1, 7, 8\}$	$2 \times 0.0352 = 0.0703$	0.020
$\{0, 1, 2, 6, 7, 8\}$	$2 \times 0.1445 = 0.289$	0.239

- Chosen region: $\{0, 1, 7, 8\}$, with $\alpha = 0.0703$
- Observed $X = 7$ lies in the region, so reject H_0

Takeäway

The binomial is discrete, so only some α values are attainable.

Q4(a): Simulating power

2 minutes

A study of $n = 16$ days is planned. Assume daily rates follow $N(\mu = 2.75,\ \sigma = 1.5)$. We do a two-sided test at $\alpha = 0.05$.

Fill in the blanks to estimate the power of the t-test.

```
count <- 0
for(i in 1:10000){
  sim_rates <- __(a)__
  pval <- t.test(sim_rates, mu = 2)$p.value
  if(__(b)__){ count <- count + 1 }
}
power <- count / 10000
```

Q4(a): Solution

```
sim_rates <- rnorm(n = 16, mean = 2.75, sd = 1.5)    # (a)
...
if(pval < 0.05){ count <- count + 1 }                # (b)
```

Type I error (Q2)

Simulate under H_0: `mean = 2`
Count rejections $\Rightarrow$ *wrong* decisions

Power (Q4)

Simulate under H_A: `mean = 2.75`
Count rejections $\Rightarrow$ *correct* decisions

Review: Lec04 pp. 20–23

Q4(b): Interpreting power

2 points · 1 minute

The code returns about **0.47**. Which one is a correct interpretation? ◯ **Select one.**

A. There is a 47% probability that H_0 is false.

B. If the true mean is 2.75 and $\sigma = 1.5$, about 47% of 16-day studies would reject H_0 at $\alpha = 0.05$.

C. There is a 47% probability of making a Type I error.

D. About 47% of days have an incident rate above 2.

Q4(b): Solution

✗ **A** Power is the probability of the test's *decision* under an assumed truth, not a probability about H_0 itself.

✓ **B** Power = probability of correctly rejecting H_0 when this particular alternative is true.

✗ **C** A Type I error requires H_0 to be true, but the data were simulated with $\mu = 2.75$.

✗ **D** Power describes repeated *studies* of 16 days, not individual days.

Review: Lec04 pp. 3–4

Q4(c): What increases power?

2 minutes

Which changes, applied **one at a time**, would increase power? ☐ **Select all that apply.**

A. Sample 25 days instead of 16.

B. Halve σ from 1.5 to 0.75.

C. The true mean is 2.25 instead of 2.75.

D. Lower α from 0.05 to 0.01.

Q4(c): Solution

✓ **A** A larger n gives a smaller standard error, so the same difference is detected more often.

✓ **B** Less day-to-day noise makes the same difference easier to detect.

✗ **C** The effect shrinks from 0.75 to 0.25, so power *falls*.

✗ **D** A smaller α means a smaller rejection region, so rejecting is harder and power falls.

Review: Lec02 p. 3; Lec04 pp. 5, 27

Q5(a): A very large sample

3 minutes

A statewide database pools $n = 40{,}000$ campus-days: $\bar{x} = 2.03$, $s = 1.6$, $p \approx 0.0002$, 95% CI $= (2.014,\ 2.046)$. At $\alpha = 0.05$: ☐ **Select all that apply.**

A. We reject H_0.

B. The tiny p-value shows that e-scooters are substantially more dangerous than e-bikes.

C. The CI suggests the excess is at most about 0.05 incidents per 1,000 rides.

D. Because daily rates are skewed, the t-test is invalid even with $n = 40{,}000$.

E. Even if the true mean were 2.01, a large enough n would make rejection almost certain.

Q5(a): Solution

✓ **A** $0.0002 < 0.05$.

✗ **B** A p-value measures the strength of evidence, not the size of the effect. The estimated excess is only 0.03.

✓ **C** The upper CI bound of 2.046 limits the plausible excess to about 0.046.

✗ **D** With $n = 40{,}000$, $\bar{x}$ is approximately normal (the same lesson as Q2).

✓ **E** For any non-zero difference, power approaches 1 as n grows.

Review: Lec05 pp. 3–8

Q5(b): A small pilot study

3 minutes

A pilot study at one busy location covers $n = 10$ days: $\bar{x} = 2.9$, $s = 1.6$, $p = 0.109$. ☐ **Select all that apply.**

A. We accept H_0: e-scooters there are as safe as e-bikes.

B. Because $p > 0.05$, the difference must be small.

C. The data are insufficient at $\alpha = 0.05$, but a rate 45% above the benchmark could matter, so a larger study is warranted.

D. `power.t.test(delta = 0.9, sd = 1.6, power = 0.80, type = "one.sample")` estimates the n needed for 80% power, treating the pilot estimates as the truth.

E. There is a 0.109 probability that H_0 is true.

Q5(b): Solution

✗ **A** Failing to reject H_0 is not evidence that H_0 is true.

✗ **B** The estimated difference (0.9 per 1,000 rides) is large; p is large because $n = 10$ is small.

✓ **C** Same situation as the poker example in lecture: promising, but not yet enough data.

✓ **D** Pilot estimates are a standard way to set power-calculation inputs before the main study.

✗ **E** The same misreading of the p-value as Q1, option B.

Review: Lec04 p. 7; Lec05 pp. 31–36

Three things to remember

1. **p-value $\neq P(H_0$ is true$)$.** It is computed *assuming* H_0 is true.
2. **Same code can show different truth.** Simulate under H_0 to get the Type I error rate; simulate under H_A to get power.
3. **Significance and importance are separate questions.**
   - Significant $\neq$ important: look at the confidence interval
   - Not significant $\neq$ no effect: check the sample size