Classical Hypothesis Tests

Stat 250

Click here for PDF version

Review: Logic of testing

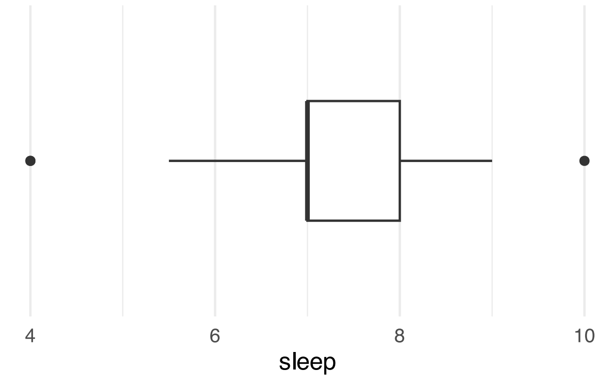
1. Formulate two competing **hypotheses** about the population
2. Calculate a **test statistic** summarizing the relevant information to the claims
3. Look at the **behavior of the test statistic** assuming that the initial claim is true
4. **Compare** the observed test statistic to the expected behavior—i.e., determine the strength of evidence against the null
5. State a **conclusion in context**

One-sample test

Do Carls sleep less than 8 hours per night?

- Surveyed Stat 120 students
- "On average, how many hours of sleep do you get on a weeknight?"

min	Q1	median	Q3	max	mean	sd	n	missing
4.00	7.00	7.00	8.00	10.00	7.07	1.23	30	0



One-sample problem

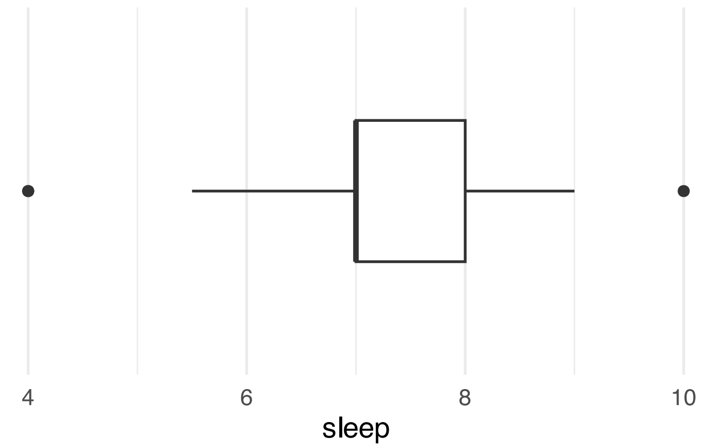
Assume	We have a random sample from a $N(\mu, \sigma^2)$ population
Hypothesize	$H_0 : \mu = \mu_0$ vs. $H_a : \mu \begin{matrix} < \\ \neq \\ > \end{matrix} \mu_0$
Test statistic	$T = \dfrac{\bar{X} - \mu_0}{s/\sqrt{n}}$
Reference distribution	$T \sim t_{n-1}$ is H_0 is true

Your turn

Do Carls sleep less than 8 hours per night?

Work with a neighbor to complete this hypothesis test and state a conclusion.

min	Q1	median	Q3	max	mean	sd	n	missing
4.00	7.00	7.00	8.00	10.00	7.07	1.23	30	0



04:00

t-tests in R

```r
t.test(~sleep, data = survey, mu = 8, alternative = "les
```

```
     One Sample t-test

data:  sleep
t = -4.1565, df = 29, p-value = 0.0001306
alternative hypothesis: true mean is less than 8
95 percent confidence interval:
     -Inf 7.448201
sample estimates:
mean of x
 7.066667
```

Extracting CIs

```
t.test(~sleep, data = survey)$conf
```

```
[1] 6.607416 7.525917
attr(,"conf.level")
[1] 0.95
```

Set `alternative` to `"greater"` or `"less"` for a one-sided confidence bound

Example

Do Americans support a national health plan?

- A Kaiser Family Foundation poll for a random sample of US adults in 2019 found that 79% of Democrats, 55% of Independents, and 24% of Republicans supported a generic “National Health Plan.”
- There were 347 Democrats, 298 Republicans, and 617 Independents surveyed.
- A political pundit on TV claims that a majority of Independents support a National Health Plan.
- Do these data provide strong evidence to support this type of statement?

One-proportion problem

Assume	$X \sim \mathrm{Binom}(n, p)$
Hypothesize	$H_0 : p = p_0$ vs. $H_a : p \begin{smallmatrix} < \\ \neq \\ > \end{smallmatrix} p_0$
Test statistic	The observed count
Reference distribution	$X \sim \mathrm{Binom}(n, p_0)$ if H_0 is true

Your turn

- 339 out of 617 independents supported a National Health Plan
- Do these data provide strong evidence to support the claim that a majority of Independents support a National Health Plan?

04:00

Exact test for proportion in R

```
binom.test(x = 339, n = 617, p = 0.5, alternative = "grea
```

```
data:  339 out of 617
number of successes = 339, number of trials = 617, p-
value = 0.007823
alternative hypothesis: true probability of success is
greater than 0.5
95 percent confidence interval:
 0.5155507 1.0000000
sample estimates:
probability of success
             0.5494327
```

Approximate test for *p*

CLT

For large enough n, $\hat{p} \mathrel{\dot\sim} N\left(p, \dfrac{p(1-p)}{n}\right)$

Approx. reference distribution

If $p = p_0$, then $\hat{p} \mathrel{\dot\sim} N\left(p_0, \dfrac{p_0(1-p_0)}{n}\right)$

Large-sample test statistic

$$Z = \frac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$$

Approximate test for proportion in R

```
prop.test(x = 339, n = 617, p = 0.5, alternative = "grea
```

```
	1-sample proportions test with continuity correction

data:  339 out of 617
X-squared = 5.8347, df = 1, p-value = 0.007857
alternative hypothesis: true p is greater than 0.5
95 percent confidence interval:
 0.5155287 1.0000000
sample estimates:
        p 
0.5494327 
```

Is n large enough?

- Many textbooks suggest $np_0 \geq 10$ and $n(1 - p_0) \geq 10$
- Our textbook suggests $np_0 \geq 384$ and $n(1 - p_0) \geq 384$
- Use the binomial test otherwise

Two-sample t-test

Is it better to hand write or type notes?

- Student researchers randomly assigned 20 college students to the paper-based note-taking group and 20 students to the computer-based note taking group
- All subjects showed a 12-minute video about the sun and allowed to take notes using the assigned method
- After video, notes collected, then subjects were given a 10-question quiz
- Does the note taking method impact the average score?

Two-sample t-test

Assume

- Both samples are iid draws from $N(\mu_i, \sigma_i^2)$ populations
- Independent groups

Hypothesize

$$H_0 : \mu_1 - \mu_2 = \delta_0 \text{ vs. } H_a : \mu_1 - \mu_2 \ \begin{smallmatrix} < \\ \neq \\ > \end{smallmatrix} \ \delta_0$$

Test statistic

$$T = \frac{\bar{X}_n - \bar{Y}_m - \delta_0}{\sqrt{\frac{s_1^2}{n} + \frac{s_1^2}{m}}}$$

Reference distribution

$T \sim t_{df}$ if H_0 is true, where df are the Welch-Sattherthwaite approx. d.f.

Two-sample t-test in R

```
t.test(Score ~ Method, data = notes, alternative = "two.s
```

```
	Welch Two Sample t-test

data:  Score by Method
t = -2.769, df = 28.682, p-value = 0.00975
alternative hypothesis: true difference in means between
group Computer and group Paper is not equal to 0
95 percent confidence interval:
 -2.4780273 -0.3719727
sample estimates:
mean in group Computer    mean in group Paper
                 5.500                  6.925
```