STAT 515 fa 2023 Lec 14 slides

Hypothesis testing

Karl B. Gregory

University of South Carolina

These slides are an instructional aid; their sole purpose is to display, during the lecture, definitions, plots, results, etc. which take too much time to write by hand on the blackboard. They are not intended to explain or expound on any material.

A *statistical inference* is a conclusion about a pop. parameter based on a rs.

Specifically, a decision concerning contradictory statements about the parameter:

- The *null hypothesis* H_0.
- The *alternate hypothesis* H_1.

The decision is whether to

1. Reject H_0, thereby concluding that H_1 is true.
2. Not reject H_0, thereby not concluding anything.

A *test of hypotheses* is a rule for when to reject H_0 based on the data.

Exercise: We want to know whether a coin is unfair. Let p be the prob. of heads.

We want to test

$$H_0\colon p = 1/2 \text{ versus } H_1\colon p \neq 1/2.$$

Suppose we toss the coin 100 times. Discuss the following:

1. Reject or fail to reject H_0 if 51 heads observed?
2. Reject or fail to reject H_0 if 60 heads observed?
3. Reject or fail to reject H_0 if 90 heads observed?
4. Reject or fail to reject H_0 if 50 heads observed?
5. What possible evidence could convince us that $p = 1/2$?
6. If the coin is fair, find prob. of observing a # of heads ≥ 60 or ≤ 40.

Exercise: Is a treatment effective in lowering cholesterol levels? Let μ represent the average difference (after-minus-before treatment) in cholesterol levels.

We want to test

$$H_0\colon \mu \geq 0 \text{ versus } H_1\colon \mu < 0.$$

Suppose we obtain $\bar{X}_n$ from $n = 100$ subjects. Discuss the following:

1. Reject or fail to reject H_0 if $\bar{X}_n = 10$?
2. Reject or fail to reject H_0 if $\bar{X}_n = -10$?
3. If the changes in chol. level are $\text{N}(\mu = 0, \sigma^2 = (25)^2)$, find $P(\bar{X}_n < -10)$.

Our data may lead us to an incorrect decision about H_0 and H_1:

- A *Type I error* is rejecting H_0 when H_0 is true.
- A *Type II error* is failing to reject H_0 when H_0 is false.

Make table summarizing possible outcomes of inference.

We like to calibrate our tests of hypotheses such that $P(\text{Type I error}) \leq \alpha$.

Then we call α the *significance level* of the test.

Suppose $X_1, \dots, X_n \overset{\text{ind}}{\sim} \text{Normal}(\mu, \sigma^2)$, with σ known.

We will consider null and alternate hypotheses of the form

$$\begin{array}{lcl}H_0\colon \mu \geq \mu_0 & & H_0\colon \mu = \mu_0 & & H_0\colon \mu \leq \mu_0 \\ H_1\colon \mu < \mu_0 & \text{or} & H_1\colon \mu \neq \mu_0 & \text{or} & H_1\colon \mu > \mu_0.\end{array}$$

Here μ_0 is a value specified by the researcher called the *null value*.

Exercise: For each set of hypotheses, find a test based on the *test statistic*

$$\frac{\bar{X}_n - \mu_0}{\sigma/\sqrt{n}}$$

with $P(\text{Type I error}) \leq \alpha$.

Let $X_1, \dots, X_n \overset{\text{ind}}{\sim} \text{Normal}(\mu, \sigma^2)$, σ^2 known.

Tests about μ when σ is known

For some null value μ_0, define the test statistic

$$Z_{\text{test}} = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}.$$

Then the following tests have $P(\text{Type I error}) \leq \alpha$.

H_0: $\mu \geq \mu_0$ H_1: $\mu < \mu_0$	H_0: $\mu = \mu_0$ H_1: $\mu \neq \mu_0$	H_0: $\mu \leq \mu_0$ H_1: $\mu > \mu_0$
Reject H_0 if $Z_{\text{test}} < -z_\alpha$	Reject H_0 if $\lvert Z_{\text{test}} \rvert > z_{\alpha/2}$	Reject H_0 if $Z_{\text{test}} > z_\alpha$

The *test statistic* is the quantity based on which we decide to reject/not reject H_0.

The $-z_\alpha$, $z_{\alpha/2}$, z_α (to which we compare the test stat) are called *critical values*.

Exercise: Suppose a bottler of soft-drinks claims that its bottling process results in an internal pressure of 157 psi with standard deviation $\sigma = 3$ psi. You want to know whether the mean pressure is less than 157 (Ex 6.92 in [1]).

1. What are the relevant hypotheses?
2. Based on a sample of size $n = 40$ you get $\bar{X} = 155.7$. Suppose $\sigma = 3$. What is your inference at the $\alpha = 0.05$ significance level?
3. Identify the following as a correct decision, a Type I error, or a Type II error:
   a. Suppose $\mu = 157.5$ and your data leads you to reject H_0.
   b. Suppose $\mu = 157.5$ and your data leads you to not reject H_0.
   c. Suppose $\mu = 156.5$ and your data leads you to reject H_0.
   d. Suppose $\mu = 156.5$ and your data leads you to not reject H_0.

Exercise: A machine should produce ball bearings with diameters having mean 0.5 in. and std. dev. $\sigma = 0.001$ in. Is the mean truly 0.5 in.? (Ex 6.84 in [1]).

1. What are the relevant hypotheses?
2. Suppose the diameters are Normal and you get $\bar{X}_n = 0.499$ with $n = 5$. Decide whether to reject or not reject H_0 at $\alpha = 0.05$.
3. Compute a 95% CI for μ based on $\bar{X}_n = 0.499$ with $n = 5$.
4. Identify the following as a correct decision, a Type I error, or a Type II error:
   a. Suppose $\mu = 0.52$ and your data leads you to reject H_0.
   b. Suppose $\mu = 0.52$ and your data leads you to not reject H_0.
   c. Suppose $\mu = 0.50$ and your data leads you to reject H_0.
   d. Suppose $\mu = 0.50$ and your data leads you to not reject H_0.

Let $X_1, \dots, X_n \overset{\text{ind}}{\sim} \text{Normal}(\mu, \sigma^2)$, σ^2 unknown.

Tests about μ when σ is unknown

For some null value μ_0, define the test statistic

$$T_{\text{test}} = \frac{\bar{X} - \mu_0}{S_n/\sqrt{n}}.$$

Then the following tests have $P(\text{Type I error}) \leq \alpha$.

H_0: $\mu \geq \mu_0$ H_1: $\mu < \mu_0$	H_0: $\mu = \mu_0$ H_1: $\mu \neq \mu_0$	H_0: $\mu \leq \mu_0$ H_1: $\mu > \mu_0$
Reject H_0 if $T_{\text{test}} < -t_{n-1,\alpha}$	Reject H_0 if $\lvert T_{\text{test}} \rvert > t_{n-1,\alpha/2}$	Reject H_0 if $T_{\text{test}} > t_{n-1,\alpha}$

Exercise: The average height of 14 randomly selected ten-yr-old Loblolly pine trees was $\bar{X}_n = 27.44$ and the sample standard deviation was $S_n = 1.54$. Assume that the heights of ten-yr-old Loblolly pine trees are Normally distributed.

1. Test the hypotheses H_0: $\mu \leq 26$ versus H_1: $\mu > 26$ at $\alpha = 0.05$.
2. Test the hypotheses H_0: $\mu \geq 26$ versus H_1: $\mu < 26$ at $\alpha = 0.05$.
3. Test the hypotheses H_0: $\mu = 26$ versus H_1: $\mu \neq 26$ at $\alpha = 0.05$.
4. Build a 95% CI for μ.

For two-sided tests at α, just see if $(1-\alpha)100\%$ CI contains the null value!

For H_0: $\mu = \mu_0$ versus H_1: $\mu \neq \mu_0$ we have:

- For σ unknown

$$|T_{\text{test}}| > t_{n-1,\alpha/2} \iff \mu_0 \notin \left(\bar{X}_n - t_{n-1,\alpha/2}\frac{S_n}{\sqrt{n}}, \bar{X}_n + t_{n-1,\alpha/2}\frac{S_n}{\sqrt{n}}\right).$$

- For σ known

$$|Z_{\text{test}}| > z_{\alpha/2} \iff \mu_0 \notin \left(\bar{X}_n - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{X}_n + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right).$$

Since $\sqrt{n}(\bar{X}_n - \mu)/S_n$ behaves like $Z \sim \text{Normal}(0,1)$ for large n...

Tests about μ when data non-Normal and $n \geq 30$

For some null value μ_0, define the test statistic

$$T_{\text{test}} = \frac{\bar{X} - \mu_0}{S_n/\sqrt{n}}.$$

Then the following tests have $P(\text{Type I error}) \leq \alpha$.

$H_0\colon \mu \geq \mu_0$ $H_1\colon \mu < \mu_0$	$H_0\colon \mu = \mu_0$ $H_1\colon \mu \neq \mu_0$	$H_0\colon \mu \leq \mu_0$ $H_1\colon \mu > \mu_0$
Reject H_0 if $T_{\text{test}} < -z_\alpha$	Reject H_0 if $\lvert T_{\text{test}} \rvert > z_{\alpha/2}$	Reject H_0 if $T_{\text{test}} > z_\alpha$

Exercise:

1. Draw a random sample of size $n = 35$ from the 2009 Boston Marathon `women's finishing times` and test the hypotheses

$$H_0\colon \mu \leq 4 \text{ versus } H_1\colon \mu > 4$$

at the $\alpha = 0.05$ significance level.

2. Repeat this 1000 times and record the proportion of times you reject H_0.

Since $\sqrt{n}(\hat{p}_n - p)/\sqrt{p(1-p)}$ behaves like $Z \sim \text{Normal}(0,1)$ for large n...

Tests about p (for $np_0 \geq 15$ and $n(1-p_0) \geq 15$)

For some null value μ_0, define the test statistic

$$Z_{\text{test}} = \frac{\hat{p}_n - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}.$$

Then the following tests have $P(\text{Type I error}) \leq \alpha$.

$H_0\colon p \geq p_0$ $H_1\colon p < p_0$	$H_0\colon p = p_0$ $H_1\colon p \neq p_0$	$H_0\colon p \leq p_0$ $H_1\colon p > p_0$
Reject H_0 if $Z_{\text{test}} < -z_\alpha$	Reject H_0 if $\lvert Z_{\text{test}}\rvert > z_{\alpha/2}$	Reject H_0 if $Z_{\text{test}} > z_\alpha$

Exercise: Does a female-inhabiting parasite tip the sex ratio of its hosts' offspring in favor of females? A sample of size $n = 500$ offspring from parasite-infected females is collected, among which there are 287 females.

1. What are the relevant hypotheses?
2. Carry out a test of the hypotheses at the $\alpha = 0.05$ significance level.
3. Identify the following as a correct decision, a Type I error, or a Type II error:
   a. Suppose $p = 0.60$ and your data leads you to reject H_0.
   b. Suppose $p = 0.60$ and your data leads you to not reject H_0.
   c. Suppose $p = 0.50$ and your data leads you to reject H_0.
   d. Suppose $p = 0.50$ and your data leads you to not reject H_0.

Exercise: In a tasting experiment, each o 121 blindfolded students was fed either a red or green gummy bear, (each with probability $1/2$) and asked to identify the color from the taste. Of the 121, 97 correctly identified the color(Ex. 8.82 of [1]).

1. If the students guessed "red" or "green" based on flipping a coin, with what probability would they guess the color correctly?
2. Suppose you wish to know if the students are doing better or worse than guessing. What are the relevant hypotheses?
3. Test the hypotheses at the $\alpha = 0.01$ significance level.

J.T. McClave and T.T. Sincich.
Statistics.
Pearson Education, 2016.