Chapter 10: Hypothesis Testing

10.1 Fundamentals of Hypothesis Testing

A hypothesis is a claim involving the numerical value of a population parameter.

- found in journalism / research studies
- e.g. "average person consumes 156 lbs of sugar per year"

Hypothesis testing is statistical process for determining whether a hypothesis is true

Procedure for performing a hypothesis test

Step 1: State the null and alternative hypothesis

Null hypothesis, H_0, expresses value currently believed to be true. Always statement of equality.

Alternative hypothesis, H_a, is hypothesis that researcher is trying to gather evidence to support.

e.g. syllabi in a college claim students spend 10 hrs / week on coursework. Director of eLearning concerned that students actually required to spend more

$H_0: \mu = 10$

$H_a: \mu > 10$

e.g. news authority claims governor's approval rating has dropped. Previous polls put rating at 56%.

$H_0: p = 0.56$

$H_a: p < 0.56$

Step 2: Determine which statistic to use for test statistic and state level of significance

$$z = \frac{\bar{x} - \mu}{\left(\frac{\sigma}{\sqrt{n}}\right)}, \quad t = \frac{\bar{x} - \mu}{\left(\frac{s}{\sqrt{n}}\right)}$$

↳ Mean, μ

↳ Proportion, p

↳ Variance, σ^2

$$z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}, \quad \chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

Sample statistic is statistically significant if far enough from pop. parameter to conclude statistic unlikely to be "by chance" if H_0 true

Level of significance, α, is probability of making the error of rejecting a true null hypothesis

Step 3 Gather data & calculate the sample statistics

Step 4 Draw a conclusion and interpret decision

Conclusions
- reject H_0
- fail to reject H_0

Types of error in hypothesis testing

What if conclusion of hypothesis test is wrong?

e.g. Say H_0 : student passes Math 101
H_a : student does not pass Math 101

	H_0 true	H_0 false
Reject H_0	Type I error	correct
Fail to reject H_0	correct	Type II error

The level of significance α is the willingness to accept a Type I error, i.e. α is the probability of making a Type I error. Similarly, β is the probability of making a Type II error.

e.g. Television executive believes 99% of households have at least one television. Intern has task of determining if actually less than 99%. The intern decides to fail to reject the null hypothesis. If in reality 96.7% was an error made and what type?

10.2 Hypothesis testing for population means (σ unknown)

we follow steps of a hypothesis test for μ and σ is known.

Step 1. State H_0 and H_a. They will be written using μ

Step 2 Determin which distribution to use for test statistic and level of significance

when (1) all samples = prob. of being chosen, (2) σ is known, (3) either $n \geq 30$ or approx. normal, use the following test statistic

$$z = \frac{\bar{x} - \mu}{\left(\frac{\sigma}{\sqrt{n}}\right)}$$

$\bar{x}$, sample mean; μ is presumed value in H_0

σ is known std.; n is sample size

Then determine level of significance,

e.g. $\alpha = 0.01, 0.05, 0.10$

Step 3 gather data and calculate the test statistic

Step 4. Either reject or accept H_0.

Here are the methods:

Method 1: using rejection region, i.e. area of tail, based on H_a, equal to level of significance.

We reject H_0 is test statistic falls within rejection region.

Types of hypothesis tests

H_a	Type
$<$ Value	left-tailed
$>$ Value	right-tailed
$\neq$ Value	2-tailed

Rejection region determined by type of test and α (level of significance)

Boundary of rejection region marked by critical z-score

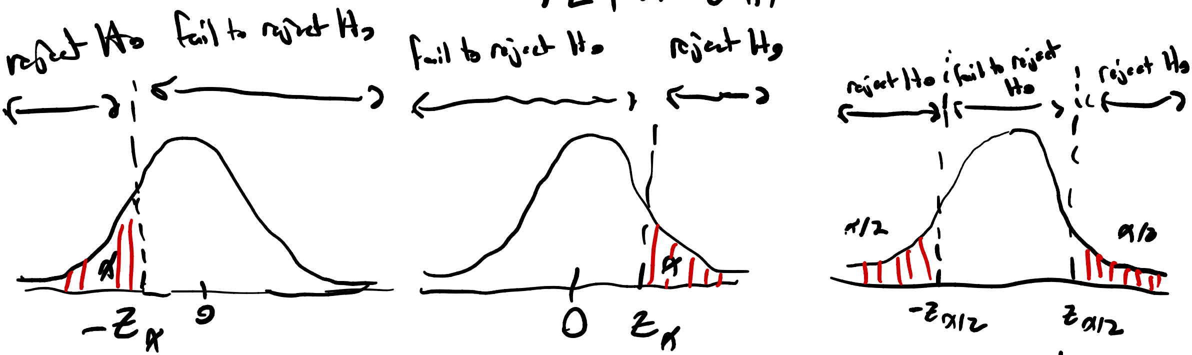
⇒ rejection regions for tests for μ (σ known)

Reject H_0 if

$z \le -z_\alpha$ for left-tail

$z \ge z_\alpha$ for right-tail

$|z| \ge z_{\alpha/2}$ for two-tail



<u>Method 2</u> drawing a conclusion using p-values

<u>Def'n</u>: A <u>p-value</u> is the probability of obtaining a sample statistic as extreme or more extreme than the one observed from the data, when H_0 is assumed true

- if p-value $\le \alpha$, reject the null hypothesis
- if p-value $> \alpha$, fail to reject null hypothesis

e.g. mean reading level of 4th graders over last 5 years is 850L. The developers of a new program want to raise reading level to above 850L. Sample of $n = 1000$ 4th graders had $\bar{x} = 856L$. Assuming population has $\sigma = 98L$. Using a significance level $\alpha = 0.05$, should the findings of the study convince the dept. of education of the positive effect of the new program.

step. 1 $\quad H_0: \mu = 850$

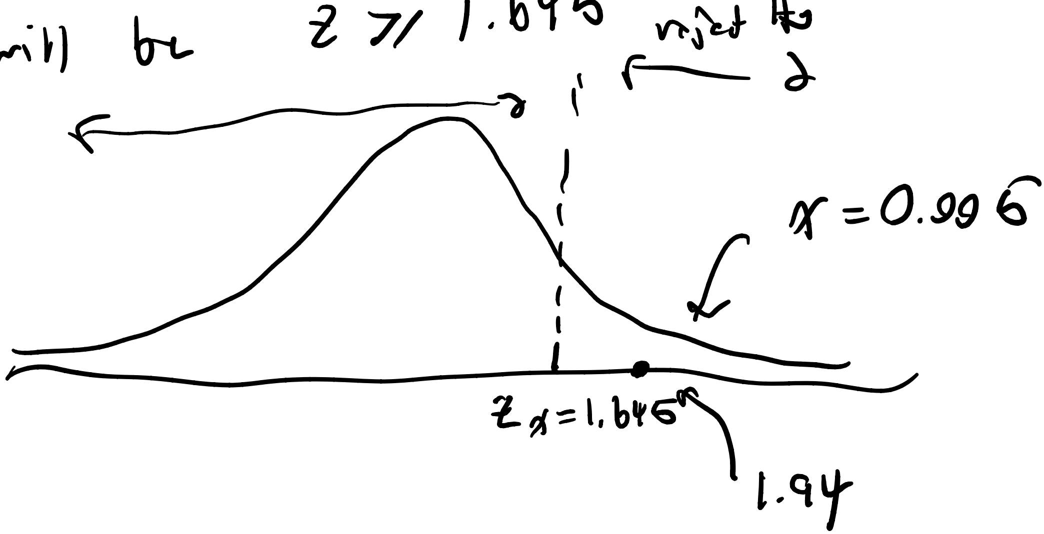
$\quad\quad\quad H_a: \mu > 850$

step 2 $\quad z = \dfrac{\bar{x} - \mu}{\left(\frac{\sigma}{\sqrt{n}}\right)} = \dfrac{856 - 850}{\left(\frac{98}{\sqrt{1000}}\right)}$

≈ 1.94

Step 3 : right-tailed test!

look up $c = 0.95$ in critical z-values table which will be $z \geq 1.645$



Since inside rejection region, we reject H_0

How can I use p-value method?

Conclusion: can be 95% sure that $\mu > 850L$.

e.g. average systolic blood pressure for patients taking a drug is $\mu = 120$

You conduct a study on 30 patients taking drug.

$\sigma = 10$

$\bar{x} = 115$

$H_0 : \mu = 120$

$H_a : \mu < 120$

left tailed test!

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \approx -2.74$$

[using p-values] look up -2.74

$\Rightarrow$ p-value $= 0.0031$

If significance level is $\alpha = 0.05$,

then since p-value $< \alpha$ we reject H_0

$\Rightarrow$ significant evidence that drug reduces blood pressure.

Hypothesis tests for μ with σ known (recap)

Step 1) $H_0: \mu = \mu_0$

$H_a: \mu \neq \mu_0$ (2 tailed test)

$H_a: \mu < \mu_0$ (left-tailed)

$H_a: \mu > \mu_0$ (right-tailed)

Step 2: choose significance level α

Step 3: calculate test statistic

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

Step 4: either calculate critical value

or p-value

10.3 Hypothesis Testing for μ [σ unknown]

Same test as before, but now

test statistic is

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

where

warning: $n \geq 30$
or distribution of x
is approx. normal

$\bar{x}$, sample mean
s, sample std.
n, sample size

e.g. blood pressure

$n = 30$
$df = 30 - 1 = 29$

$H_0: \mu = 120$
$H_a: \mu < 120$

$\bar{x} = 115$
$s = 10$

but now we have

$$t = \frac{115 - 120}{10/\sqrt{30}} = -2.74$$

if $\alpha = 0.05$, $t_\alpha = -1.699$

$t < t_\alpha \Rightarrow$ reject H_0

e.g. $\mu_0 = 2.0$ lb beef portions

found $s = 0.33$ lb and $\bar{x} = 2.10$ from random sample of $n = 20$ packages of beef. Is there evidence at 0.01 level of significance that machine not working properly.

$H_0 : \mu = 2.0$

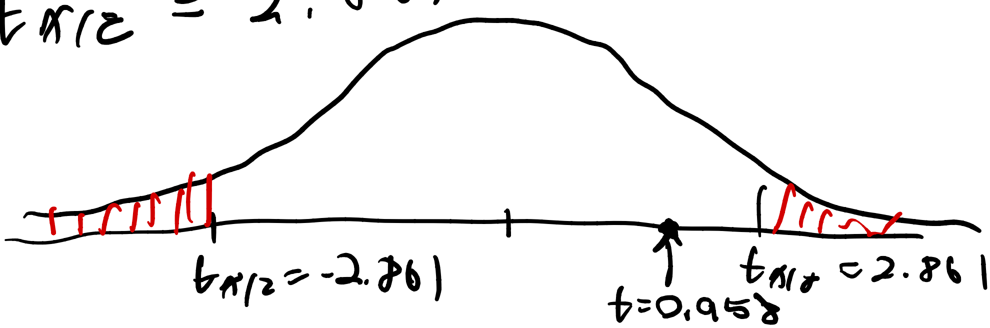
$H_a : \mu \neq 2.0$

$$t = \frac{\bar{x} - \mu_0}{(s/\sqrt{n})} = \frac{2.1 - 2.0}{0.33/\sqrt{20}} \simeq 0.958$$

$n = 20 \Rightarrow df = 20 - 1 = 19$

$\alpha = 0.01 \Rightarrow \alpha/2 = 0.005$

$t_{\alpha/2} = 2.861$ $\quad -t_{\alpha/2} = -2.861$



Hence, t lies OUTSIDE reject region
$\Rightarrow$ fail to reject H_0

$\boxed{\text{10.4 Hypothesis test for } p}$

Same as z-test, but now

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

Warning: conditions of binomial distribution need to be met and $np, n(1-p) \geq 10$

e.g. WSJ poll has $\hat{p}_{Trump} = 0.49$.
To win election, need ($\hat{p}_{Harris} = 0.46$)
about $p = 0.47$. With significance $\alpha = 0.1$ can we conclude either candidate will win?

$H_0: p_0 > 0.47$ $n = 1500$
(right tailed test)

Trump, $Z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)} / \sqrt{n}} = 1.55$

Harris, $Z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)} / \sqrt{n}} = -0.78$

$p_{trump} = 0.0606 \leq 0.1 = \alpha \Rightarrow$ reject H_0

$p_{harris} = 0.7823 \geq 0.1 = \alpha \Rightarrow$ fail to reject H_0

10.5 Hypothesis Test for σ

Can use χ^2 test if population $\approx$ normal

$$\chi^2 = \frac{(n-1) s^2}{\sigma_0^2}$$ is test statistic for population variance

where n = sample size

σ_0 = presumed value in H_0

s^2 = sample variance

Rejection regions (Method 1)

reject H_0 if

- $\chi^2 \leq \chi^2_{(1-\alpha)}$ (for left-tail)
- $\chi^2 \geq \chi^2_{\alpha}$ (for right-tail)
- $\chi^2 \leq \chi^2_{(1-\alpha/2)}$ or $\chi^2 \geq \chi^2_{\alpha/2}$ (2 tailed)

Or use p-value w/ calculator

if p-value $\leq \alpha$, reject H_0

if p-value $\geq \alpha$, fail to reject H_0

eg. Drug company claims its variance in active ingredient less than leading brand

$\sigma_0 = 0.000900$ $n = 100$

$\bar{x} = 2.470$ mg. $s = 0.026 = \partial$

$s^2 = 0.000676$

$\alpha = 0.01$ $H_0: \sigma_0^2 = 0.0009$

$H_a: \sigma_0^2 < 0.0009$

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(100-1)\,0.000676}{0.0009}$$

$$= 74.36$$

rejection region $\Rightarrow$ $df = 100 - 1 = 99$

closest is χ^2 for $df = 100$.

$\chi^2_{0.99} = 70.065$ is critical value

Not in rej. region $\quad \chi^2 \leq \chi^2_{(1-\alpha)^2}$

hence fail to reject H_0

e.g. tree farmer wants to know variance of trees (2 years old). $\sigma_0 = 16$ ft according to botanists. Simple random sample of $n = 40$; found $s = 3.2$ ft. Does this evidence, at 0.10 level of significance, support the farmer's claim that the variance after 2 yrs is not 16 ft.

Step 1

$H_0: \sigma_0^2 = 16$

$H_a: \sigma_0^2 \neq 16$

Step 2: χ^2-test w/ $\alpha = 0.10$

Step 3:
$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(40-1)(3.2)^2}{16}$$

$$= 24.96$$

Step 4: rej. region $\chi^2 \leq \chi^2_{(1-\alpha/2)}$ or $\chi^2 \geq \chi^2_{\alpha/2}$

df = 40 - 1 = 39 (approximately 40)

$\chi^2_{0.95} = 26.509$

$\chi^2_{0.05} = 55.758$

Hence, decision to reject H_0

11.1 Hypothesis Testing: two population means (σ known, independent samples)

Two types of null hypothesis

$H_0: \mu_1 = \mu_2$

$H_a: \mu_1 > \mu_2$

or

$H_0: \mu_1 - \mu_2 = 20$

$H_a: \mu_1 - \mu_2 > 20$ (right-tailed)

Test statistic

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

where $\bar{x}_1$ and $\bar{x}_2$ are sample means

$\mu_1 - \mu_2$ is presumed value of difference between population means

σ_1 and σ_2 are 2 pop. std.
n_1 and n_2 are sample sizes
(As a reminder)

Reject H_0 if

$z \leq -z_\alpha$ for left-tailed

$z \geq z_\alpha$ for right-tailed

$|z| \geq z_{\alpha/2}$ for two-tailed

e.g. drug manufacturer claims new drug w/ diet and excercise lowers cholesterol by over 20 points more than diet + excercise alone.

$n_1 = n_2 = 55$

$\overline{x}_1 = 44.7$ (diet/excercise/drug)

$\sigma_1 = 6.8$

(diet/excercise)

$\overline{x}_2 = 23.1$

$\sigma_2 = 5.3$

Use $\alpha = 0.01$

State null/alternative hypothesis:

$$H_0 : \mu_1 - \mu_2 = 20$$

$$H_a : \mu_1 - \mu_2 > 20$$

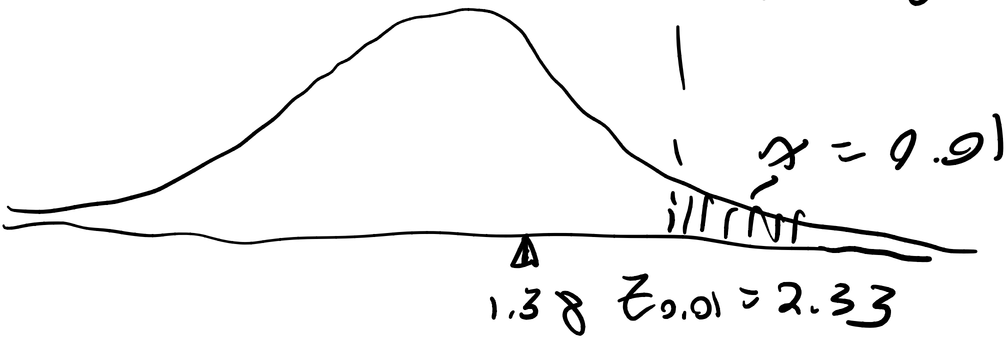
$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1^2} + \frac{\sigma_2^2}{n_2^2}}}$$

$$= \frac{(44.7 - 23.1) - 20}{\sqrt{\frac{6.8^2}{55} + \frac{5.3^2}{55}}}$$

$$\approx 1.38$$

Look up critical z-value for $\alpha = 0.01$

(one right tail)



or, find are to right of $z = 1.38$

p-value $\geq \alpha$ $\Rightarrow$ fail to reject H_0