Chapter 7: Hypothesis Testing Procedures

Dr. Daniel B. Rowe
Professor of Computational Statistics
Department of Mathematical and Statistical Sciences
Marquette University

Hypothesis Testing

We make decisions every day in our lives.

Should I believe A or should I believe B (not A)?

Two Competing Hypotheses. A and B.

Null Hypothesis ($\mathbf{H_0}$): No difference, no association, or no effect.

Alternative Hypothesis ($\mathbf{H_1}$): Investigators belief.

The Alternative Hypothesis is always set up to be what you want to build up evidence to prove.

7.1 Introduction to Hypothesis Testing

$$\hat{p} = \frac{x}{n}$$

Example: Friend's Party.

H_0: The party will be boring.

vs.

H_1: The party will be fun.

I wish that every time I had to make a decision, I could calculate a measure and use this measure (test statistic) to decide what to do.

Maybe use $\hat{p}$ the sample proportion of fun parties friend has had?
I might believe the party will be fun if $\hat{p}$ is "large."

7.1 Introduction to Hypothesis Testing

$$\bar{X} = \frac{1}{n}\sum X$$

Example: Men's Weight.

H_0: The mean weight of men is equal to 191 lbs. $\mu = 191$ lbs ($\leq$)

vs.

H_1: The mean weight of men is greater than 191 lbs. $\mu > 191$ lbs

I wish that every time I had to make a decision, I could calculate a measure and use this measure (test statistic) to decide what to do.

Maybe use $\bar{X}$ the sample mean weight of men?

I might believe men's mean weight ≥ 191 if $\bar{X}$ is "large."

7.1 Introduction to Hypothesis Testing

Example: H_0: μ = 191 lbs *vs*. H_1: $\mu \geq$ 191 lbs

To test the hypothesis, take a sample of n=100 men's weights.

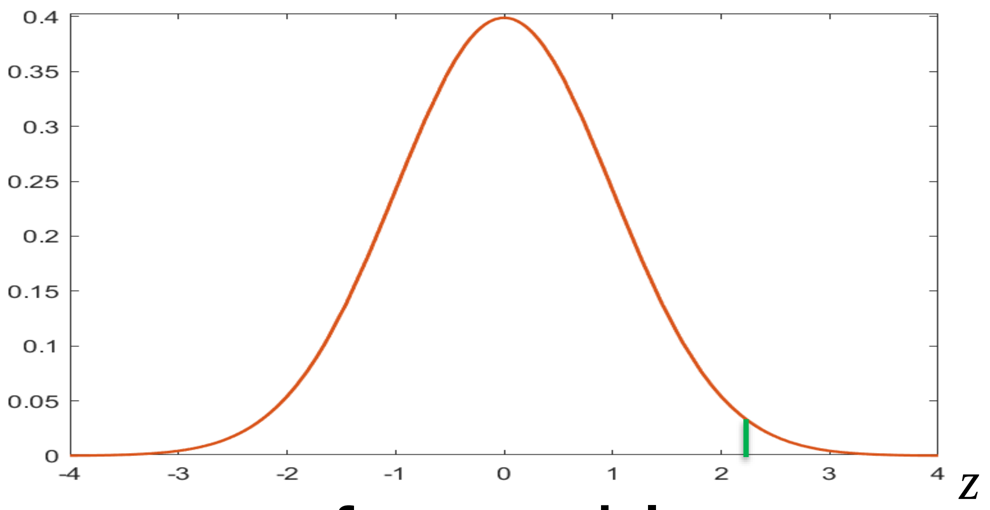
Suppose n=100 , $\bar{X}$ =197.1 lbs, and s=25.6 lbs.

Is 197.1 statistically larger than 191?

In hypothesis testing we assume H_0 is true, then see how likely $\bar{X}$ is.

$$P(\bar{X} > 197.1) = P\left(\frac{\bar{X} - 191}{25.6/\sqrt{100}} > \frac{197.1 - 191}{25.6/\sqrt{100}} \right)$$

$$P(z > 2.38) = 1 - 0.9913 = 0.0087$$ ← very unlikely



Assumed that $\bar{X}$ was normal and used z because n>30. 2.38 from table

7.1 Introduction to Hypothesis Testing

Example: H_0: $\mu = 191$ lbs *vs*. H_1: $\mu \geq 191$ lbs

To test the hypothesis, take a sample of n=100 men's weights.

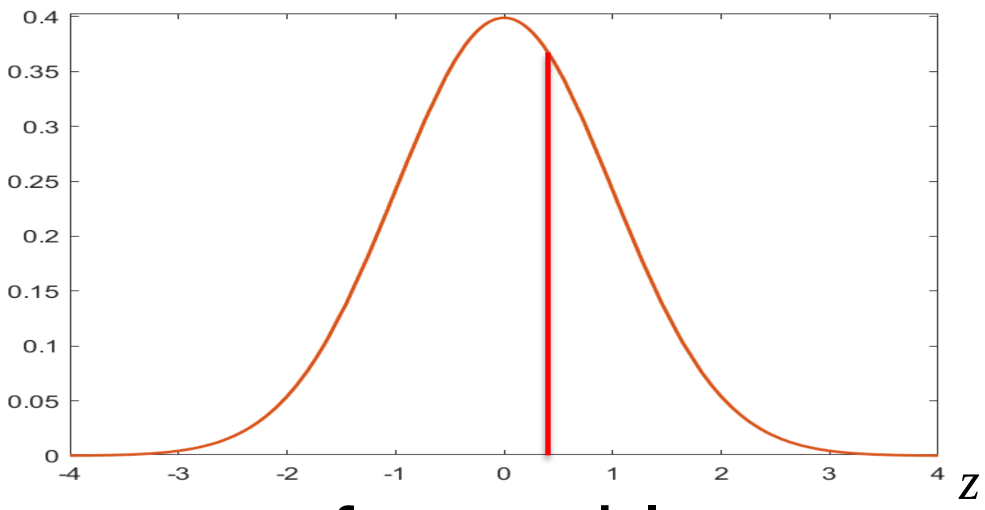
Suppose n=100 , $\bar{X} = 192.1$ lbs, and s=25.6 lbs.

Is 192.1 statistically larger than 191?

In hypothesis testing we assume H_0 is true, then see how likely $\bar{X}$ is.

$$P(\bar{X} > 192.1) = P\left(\frac{\bar{X} - 191}{25.6/\sqrt{100}} > \frac{192.1 - 191}{25.6/\sqrt{100}} \right)$$

$$P(z > 0.43) = 1 - 0.6664 = 0.3336$$ ← somewhat unlikely



Assumed that $\bar{X}$ was normal and used z because n>30. 0.43 from table

7.1 Introduction to Hypothesis Testing

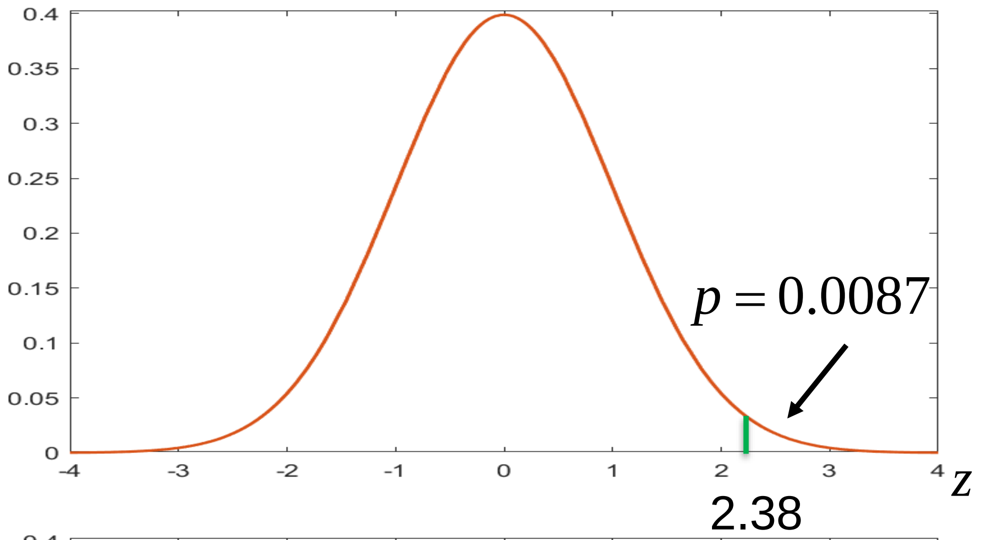
Where do we draw the line?

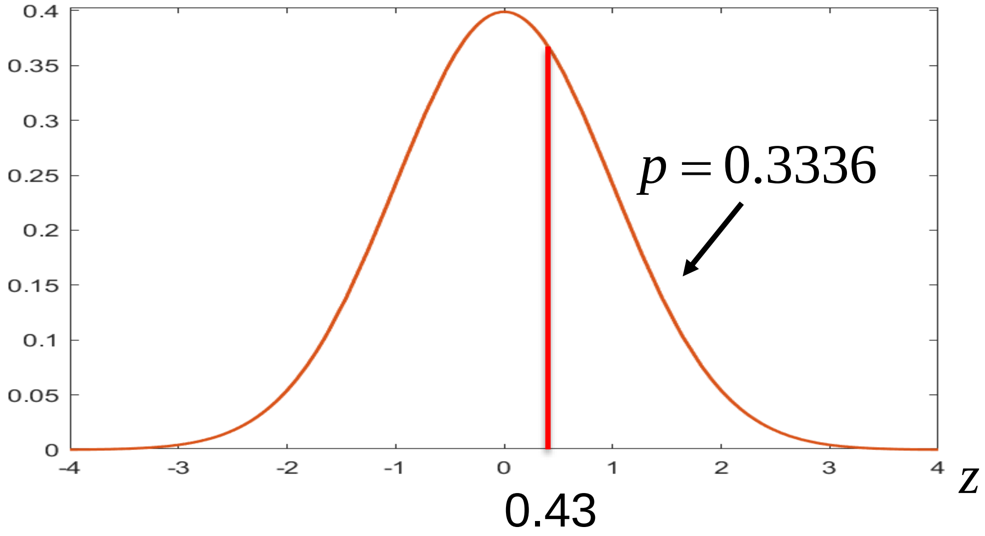
Suppose n=100 , $\bar{X}$ =197.1 lbs, and s=25.6 lbs.

$$P(z > 2.38) = 1 - 0.9913 = 0.0087$$ ← very unlikely

Suppose n=100 , $\bar{X}$ =192.1 lbs, and s=25.6 lbs.

$$P(z > 0.43) = 1 - 0.6664 = 0.3336$$ ← somewhat unlikely





We need a scientific way to select a cut-off α (probability) or z-*value* (critical value).

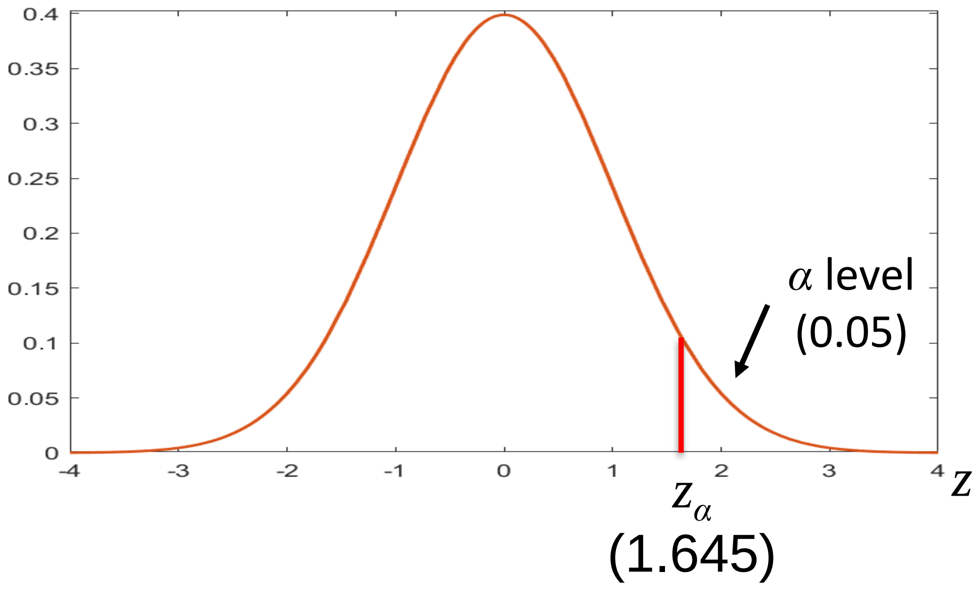
cut-off(s) called **critical value(s)** and depend on significance level α

7.1 Introduction to Hypothesis Testing

We will select a **Level of Significance** α.

$$\alpha = P(Reject\ \mathrm{H}_0 \mid \mathrm{H}_0\ is\ true)$$

Find z-$value$, z_α that corresponds to this α level.



Reject the Null Hypothesis H_0 in favor of the Alternative Hypothesis H_1

When z-$value > z_\alpha$ or p-$value < \alpha$.

This will be our scientific way to determine whether to believe the null hypothesis H_0 or alternative hypothesis H_1.

7.1 Introduction to Hypothesis Testing

The hypothesis testing process consists of 5 Steps.

Step 1: Set up the hypotheses and determine the level of significance.

Step 2: Select the appropriate test statistic.

Step 3: Set-up the decision rule.

Step 4: Compute the test statistic.

Step 5: Conclusion.

7.1 Introduction to Hypothesis Testing

The hypothesis testing process consists of 5 Steps.

Step 1: Set up the hypotheses and determine the level of significance.
State the null and the alternative hypotheses.
H_0: Null Hypothesis (no change, no difference)
vs.
H_1: Research Hypothesis (investigators belief, what we want to prove)

Select a level of significance α. $\alpha=0.05$

7.1 Introduction to Hypothesis Testing

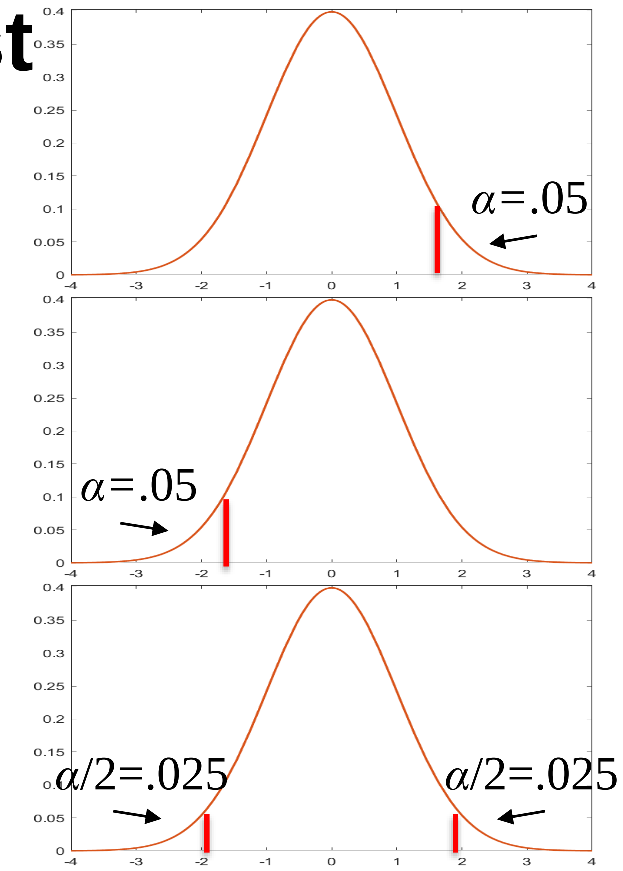
The hypothesis testing process consists of 5 Steps.

Step 1: Set up the hypotheses and determine the level of significance. There are three possible pairs.

$H_0: \mu = \mu_0$ *vs*. $H_1: \mu > \mu_0$ (prove greater than, **upper tailed test**)
$\quad\ \ \le$ reject for "large" $\bar{X}$ or z's

$H_0: \mu = \mu_0$ *vs*. $H_1: \mu < \mu_0$ (prove less than, **lower tailed test**)
$\quad\ \ \ge$ reject for "small" $\bar{X}$ or z's

$H_0: \mu = \mu_0$ *vs*. $H_1: \mu \neq \mu_0$ (prove not equal to, **two-tailed test**)
reject for "large" or "small" $\bar{X}$ or z's



7.1 Introduction to Hypothesis Testing

The hypothesis testing process consists of 5 Steps.

Step 2: Select the appropriate test statistic.
The test statistic is a single (decision) number.

n large

$$z = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$$

n small

$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} \qquad df=n-1$$

Use the test statistic that depends on data and null hypothesis with a critical value z_a (or $t_{a,df}$) that depends on significance level α to make decision.

$a = \alpha$ or $\alpha/2$

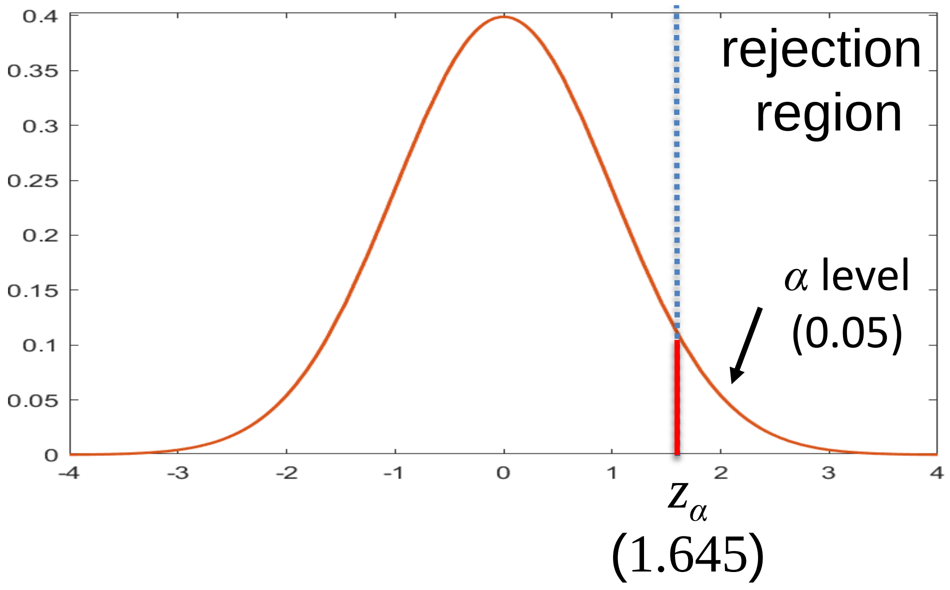
We will test hypotheses on various parameters with various test statistics.

7.1 Introduction to Hypothesis Testing

The hypothesis testing process consists of 5 Steps.

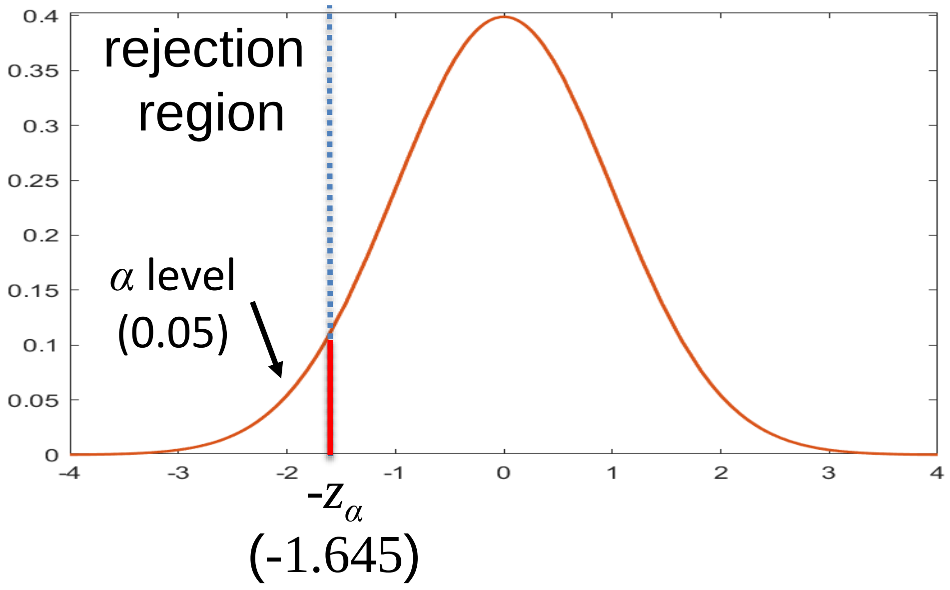
Step 3: Set-up the decision rule.

$H_0: \mu = \mu_0$ vs. $H_1: \mu > \mu_0$



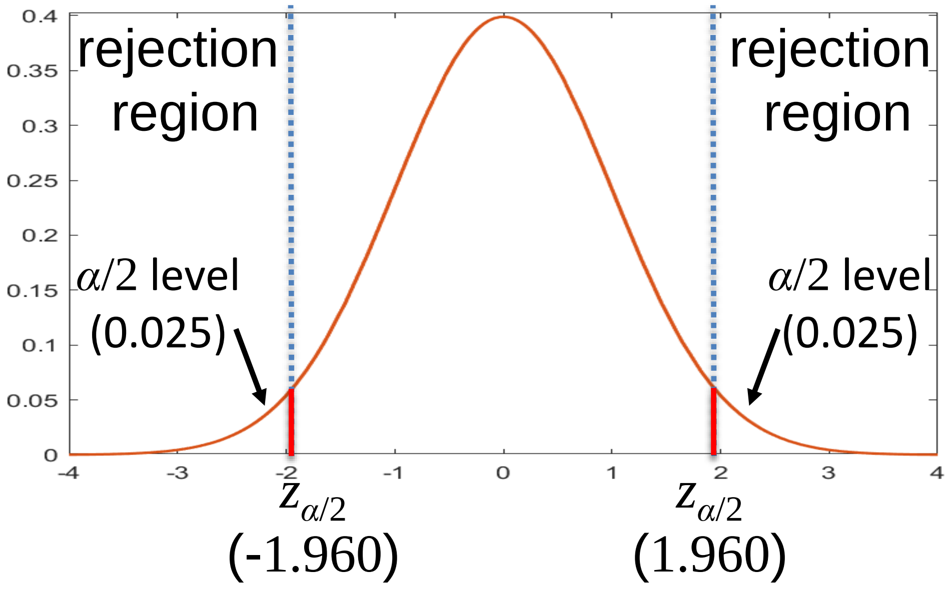
Reject H_0 if $z \geq z_\alpha$

$H_0: \mu = \mu_0$ vs. $H_1: \mu < \mu_0$



Reject H_0 if $z \leq z_\alpha$

$H_0: \mu = \mu_0$ vs. $H_1: \mu \neq \mu_0$



Reject H_0 $z \leq z_{\alpha/2}$ or $z \geq z_{\alpha/2}$

7.1 Introduction to Hypothesis Testing

The hypothesis testing process consists of 5 Steps.

Step 4: Compute the test statistic.

Use sample data $x_1,\ldots,x_n$ and hypothesized value μ_0 to compute z (or t).

Compare test statistic z (or t) to critical value(s) $z_{\alpha/2}$ (or $t_{\alpha/2,df}$) with rule.

Step 5: Conclusion.

Make a decision, reject H_0 or not to reject H_0.

Interpret the results.

7.1 Introduction to Hypothesis Testing

There are two types of error we can make.

Type I error rate.

$$\alpha = P(\text{Type I Error}) = P(\text{Reject } \mathrm{H}_0 | \mathrm{H}_0 \text{ is true})$$

Sometimes called the false positive rate.

Type II error rate.

$$\beta = P(\text{Type II Error}) = P(\text{Do Not Reject } \mathrm{H}_0 | \mathrm{H}_0 \text{ is false})$$

Sometimes called the false negative rate.

	H_0 True	H_0 False
Fail to Reject H_0	Correct Decision $(1\text{-}\alpha)$	Type II Error (β)
Reject H_0	Type I Error (α)	Correct Decision $(1\text{-}\beta)$

7.2 tests with One Sample, Continuous Outcome

Covers same material as Section 7.1 but additional small sample test with t statistic.

TABLE 7-4	Test Statistic for Testing $H_0: \mu = \mu_0$	
$n \geq 30$	$z = \dfrac{\overline{X} - \mu_0}{s/\sqrt{n}}$	(Find critical value in Table 1C)
$n < 30$	$t = \dfrac{\overline{X} - \mu_0}{s/\sqrt{n}}$	(Find critical value in Table 2, $df = n - 1$)

7.3 Tests with One Sample, Dichotomous Outcome

To test hypothesis on a proportion, we follow the same 5 Steps.

Step 1: Set up the hypotheses and determine the level of significance.

$H_0: p = p_0$ vs. $H_1: p > p_0$, $H_0: p = p_0$ vs. $H_1: p < p_0$, $H_0: p = p_0$ vs. $H_1: p \neq p_0$

Step 2: Select the appropriate test statistic.

Assume n is large. $z = (\hat{p} - p_0) \Big/ \sqrt{p_0(1 - p_0)/n}$ $\hat{p} = \dfrac{x}{n}$

Step 3: Set-up the decision rule.

Reject H_0 if $z \geq z_\alpha$, Reject H_0 if $z \leq z_\alpha$, Reject H_0 $z \geq z_{\alpha/2}$ or $z \leq z_{\alpha/2}$

Step 4: Compute the test statistic.

$$z = a \ number$$

Step 5: Conclusion.

Compare test statistic to critical value(s). Make a decision.

7.3 Tests with One Sample, Dichotomous Outcome

Example: Is proportion of children using dental service different from 0.86?

Step 1: Null and Alternative Hypotheses.

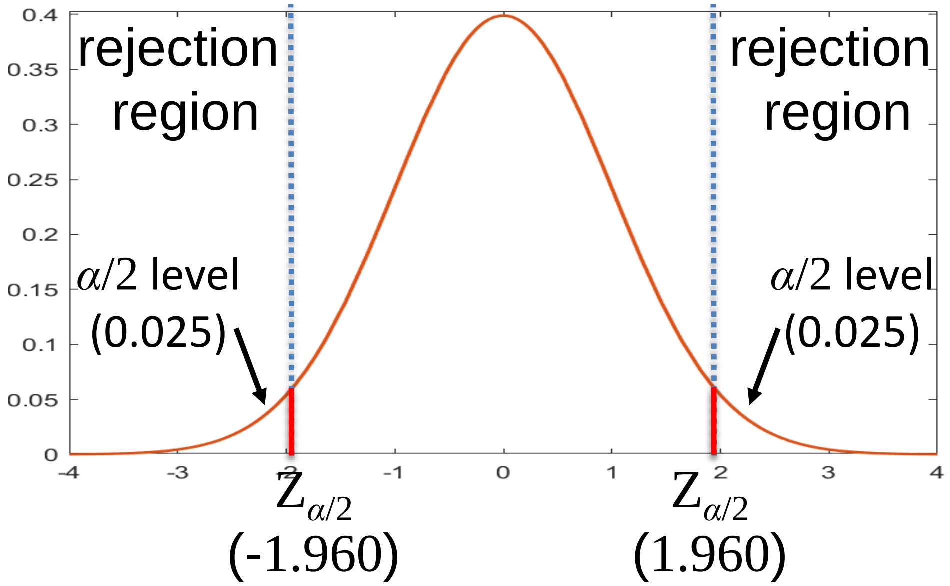
H_0: $p = 0.86$ *vs*. H_1: $p \neq 0.86$

Step 2: Test Statistic.

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

Step 3: Decision Rule. α=0.05

Reject H_0 if $z \leq$ -1.960 or $z \geq$ 1.960.



Step 4: Compute test statistic. n=125, x=64, $\hat{p} = x/n = 0.512$.

$$z = (0.512 - 0.86) \Big/ \sqrt{0.86(1-0.86)/125} = -11.21$$

Step 5: Conclusion

Because $z \leq$ -1.96, reject and conclude proportion different from 0.86.

7.4 Tests with One Sample, Categorical and Ordinal Outcomes

There are cases with more than two Yes/No categories. Binomial

Assume that we have n items classified into one of k categories.

We have a hypothesis about the true proportions for each category.

We want to test to see if our hypothesis is correct, or something different.

We can do this with a scientific statistical hypothesis test.

7.4 Tests with One Sample, Categorical and Ordinal Outcomes

To test hypothesis on a proportion, we follow the same 5 Steps.

Step 1: Set up the hypotheses and determine the level of significance.

$H_0: p_1 = p_{01}, \ldots, p_k = p_{0k}$ vs. H_1: H_0 false (only one pair)

Step 2: Select the appropriate test statistic.

$$\chi^2 = \sum (O-E)^2 / E \qquad df=k\text{-}1 \qquad E_i = np_{0i}$$

Step 3: Set-up the decision rule.

Reject H_0 if $\chi^2 \geq \chi^2_{\alpha,df}$.

Step 4: Compute the test statistic.

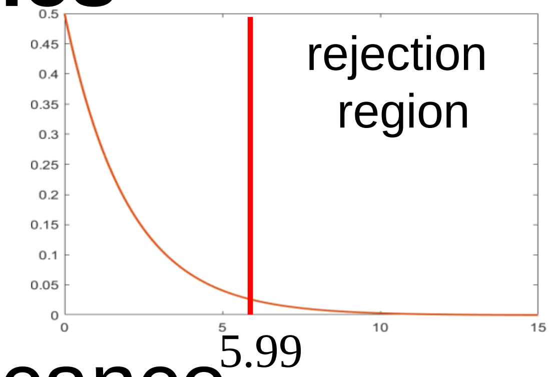
$$\chi^2 = a\ number$$

Step 5: Conclusion.

Compare test statistic to critical value. Make a decision.

7.4 Tests with One Sample, Categorical and Ordinal Outcomes

Book has df=1 figure.



Example: Health Survey. $n = 470$

Step 1: Set up the hypotheses and determine the level of significance.

H_0: $p_1 = 0.60$, $p_2 = 0.25$, $p_3 = 0.15$ *vs.* H_1: H_0: false (only one pair) α=0.05

Step 2: Select the appropriate test statistic.

$$\chi^2 = \sum (O-E)^2 / E \qquad df=k\text{-}1 \qquad E_i = np_{0i}$$

Step 3: Set-up the decision rule.

Reject H_0 if $\chi^2 \geq \chi^2_{0.05,2}$=5.99. ← Table 3 next slide

	No Regular Exercise	Sporadic Exercise	Regular Exercise	Total
(O)	255	125	90	470
(E)	470(0.60) = 282	470(0.25) = 117.5	470(0.15) = 70.5	470

Step 4: Compute the test statistic.

$$\chi^2 = \frac{(255-282)^2}{282} + \frac{(125-177.5)^2}{177.5} + \frac{(90-70.5)^2}{70.5} = 8.46$$

Step 5: Conclusion.

Since χ^2=8.46 $\geq \chi^2_{0.05,2}$ =5.99, reject H_0 conclude p's not what we hypothesize.

7.4 Tests with One Sample, Categorical and Ordinal Outcomes

TABLE 3. Critical Values of the χ2 Distribution

Table entries represent values from χ^2 distribution with upper tail area equal to α.
$P(\chi^2_{df} > \chi^2) = \alpha$, e.g., $P(\chi^2_3 > 7.81) = 0.05$

df	.10	.05	.025	.01	.005
1	2.71	3.84	5.02	6.63	7.88
2	4.61	5.99	7.38	9.21	10.60
3	6.25	7.81	9.35	11.34	12.84
4	7.78	9.49	11.14	13.28	14.86
5	9.24	11.07	12.83	15.09	16.75
6	10.64	12.59	14.45	16.81	18.55
7	12.02	14.07	16.01	18.48	20.28
8	13.36	15.51	17.53	20.09	21.95
9	14.68	16.92	19.02	21.67	23.59
10	15.99	18.31	20.48	23.21	25.19
11	17.28	19.68	21.92	24.72	26.76
12	18.55	21.03	23.34	26.22	28.30
13	19.81	22.36	24.74	27.69	29.82
14	21.06	23.68	26.12	29.14	31.32
15	22.31	25.00	27.49	30.58	32.80
16	23.54	26.30	28.85	32.00	34.27
17	24.77	27.59	30.19	33.41	35.72
18	25.99	28.87	31.53	34.81	37.16
19	27.20	30.14	32.85	36.19	38.58
20	28.41	31.41	34.17	37.57	40.00
21	29.62	32.67	35.48	38.93	41.40
22	30.81	33.92	36.78	40.29	42.80
23	32.01	35.17	38.08	41.64	44.18
24	33.20	36.42	39.36	42.98	45.56

Questions?

Homework 7

Read Chapter 7.

Problems # 4, *, 9

* A doctor believes that less than 20% of patients have a certain disease. In a random sample of n=100 patients, x=17 had the disease. Test the hypotheses H_0: $p \geq 0.20$ *vs*. H_1: $p < 0.20$ at α=0.025.