Hypothesis Testing

Nonstatistical Hypothesis Testing…

In a trial a jury must decide between two hypotheses.

- H_0: The null hypothesis - The defendant is innocent
- H_1: The alternative hypothesis -The defendant is guilty

The jury does not know which hypothesis is true. They must make a decision on the basis of evidence presented.

Research Hypothesis Testing…

The FDA or "science" needs to decide on a new theory, drug, treatment…

- H_0: The null hypothesis - the current theory, drug, treatment, is as good or better
- H_1: The alternative hypothesis - the new theory, drug, treatment, should replace the old one

Researchers do not know which hypothesis is true. They must make a decision on the basis of evidence presented.

Hypothesis Testing Terms

- Convicting the defendant is *rejecting the null hypothesis in favor of the alternative hypothesis*.
- There is enough evidence to conclude that the defendant is guilty (i.e., there is enough evidence to support the alternative hypothesis).
- If the jury acquits it is stating that *there is not enough evidence to support the alternative hypothesis*.
- The jury is not saying that the defendant is innocent, only that there is not enough to convict.
- We don't accept the null hypothesis, we just don't have sufficient evidence to reject it.

Hypothesis Testing…

- Two hypotheses, the null and the alternative
- Begins with the assumption that the null hypothesis is true.
- Is there enough evidence to infer that the alternative hypothesis is true, or the null is not likely to be true.
- There are two possible decisions:
  - Enough evidence to support the alternative hypothesis. Reject the null.
  - Not enough evidence to support the alternative hypothesis. Fail to reject the null.

Example: Is the Coin Biased?

- Flip a coin 10 times and get 2 HEAds
- Flip a coin 100 times and get 20 HEAds
- Flip a coin 100 times and get 37 HEAds
- Flip a coin 100 times and get 47 HEAds

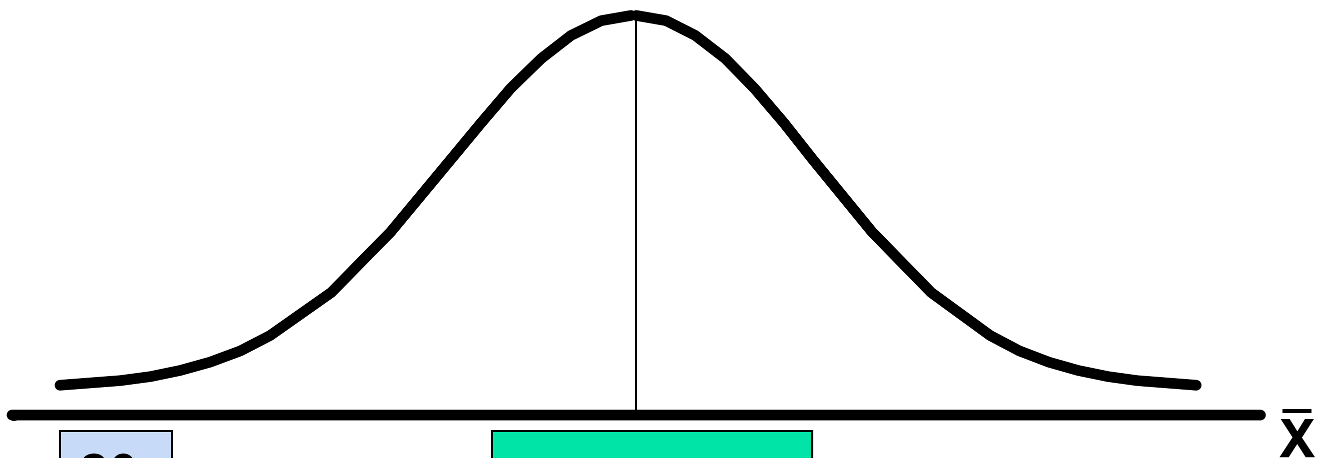
We reject the prior assumption that the coin is unbiased if the number of HEADs is very unlikely under that assumption

Example: Is the Coin Biased?

- H_0 : Probability of HEAD = ½
- Possible alternatives:
  - H_1 : Probability of HEAD $\neq$ ½
  - H_1 : Probability of HEAD is > ½
  - H_1 : Probability of HEAD is < ½

Example: Is the Coin Biased?

Number of HEADs in 100 flipd



20

μ = 50
If H_0 is true

If it is unlikely that we get this value ...

... if in fact the coin was unbiased..

... Therefore we reject the null hypothesis that the coin is unbiased

Level of Significance and the Rejection Region

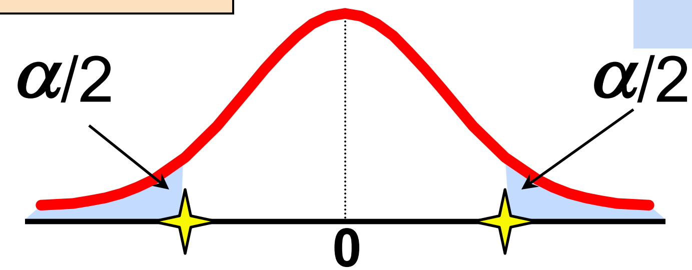
Level of significance = α

Represents critical value

Rejection region is shaded

H_0: $\mu = 1/2$

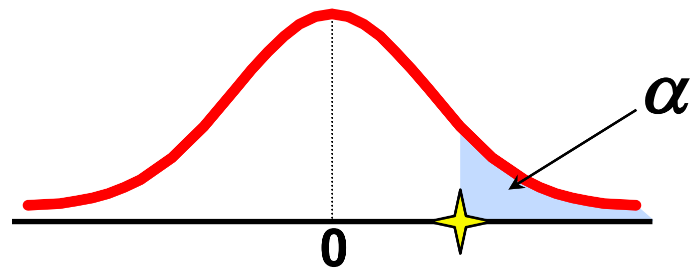
H_1: $\mu \neq 1/2$



Two-tail test

H_0: $\mu \leq 1/2$

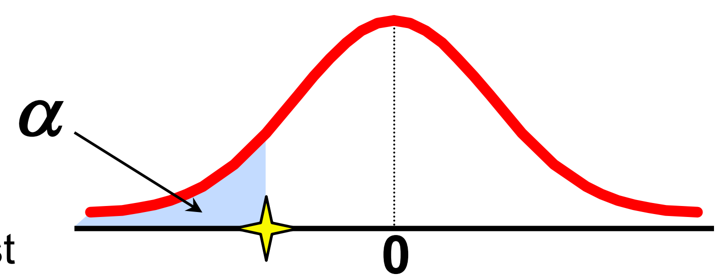
H_1: $\mu > 1/2$



Upper-tail test

H_0: $\mu \geq 1/2$

H_1: $\mu < 1/2$



Lower-tail test

What is a Hypothesis?

- A hypothesis is a claim (fact, model, parameter) that can be tested.

  - population mean

Example: The mean monthly cell phone bill of this city is $\mu = \$42$

  - population proportion

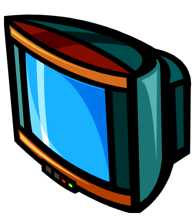
Example: The proportion of adults in this city with cell phones is $p = .68$

The Null Hypothesis, H_0

- States the assumption (numerical) to be tested

> Example: The average number of TV sets in U.S. Homes is equal to three ($H_0: \mu = 3$)

- Is always about a model, not about a sample statistic

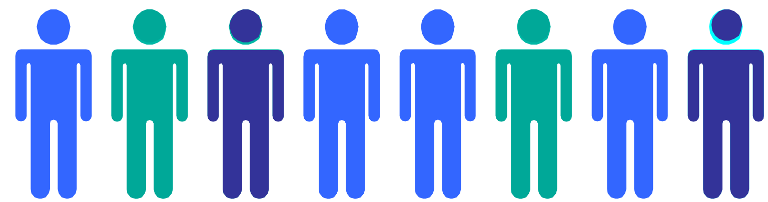


$H_0: \mu = 3$

$\cancel{H_0: \bar{X} = 3}$

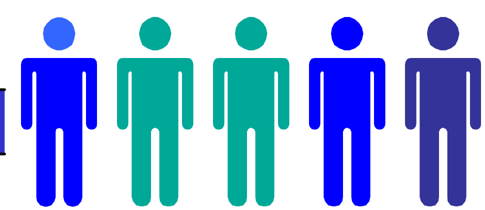
Hypothesis Testing Process

Claim: the population mean age is 50.
(Null Hypothesis:
H_0: $\mu = 50$)



Population

Now select a random sample



Sample

Suppose the sample mean age is 20: $\bar{X} = 20$

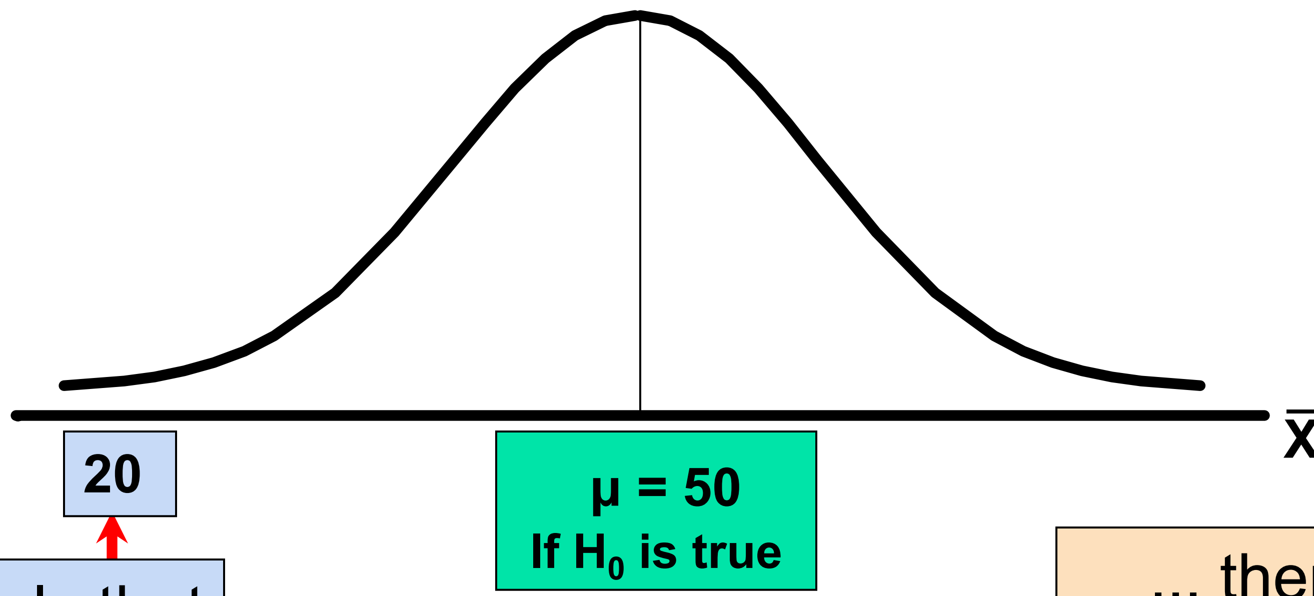
Is $\bar{X} = 20$ likely if $\mu = 50$?

If not likely,

REJECT
Null Hypothesis

Reason for Rejecting H_0

Sampling Distribution of $\bar{X}$



If it is unlikely that we would get a sample mean of this value ...

... if in fact this were the population mean…

... then we reject the null hypothesis that $\mu = 50$.

Level of Significance, α

- **Defines the unlikely values of the sample statistic if the null hypothesis is true**
  - Defines rejection region of the sampling distribution
- Is designated by α , (level of significance)
  - Typical values are .01, .05, or .10
- Is selected by the researcher at the beginning
- Provides the critical value(s) of the test

Level of Significance and the Rejection Region

Level of significance = α

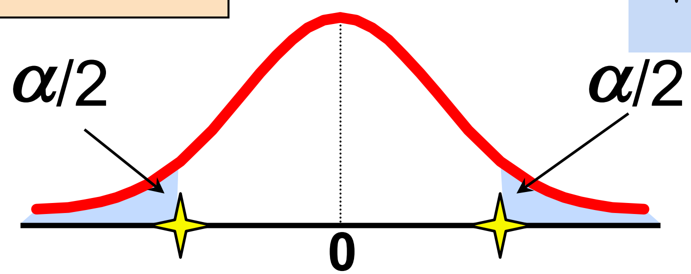
Represents critical value

Rejection region is shaded

$H_0: \mu = 3$

$H_1: \mu \neq 3$

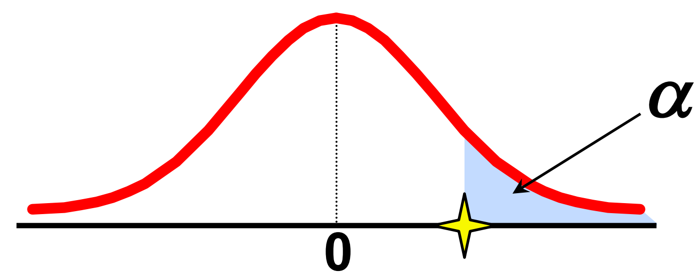
Two-tail test



$H_0: \mu \leq 3$

$H_1: \mu > 3$

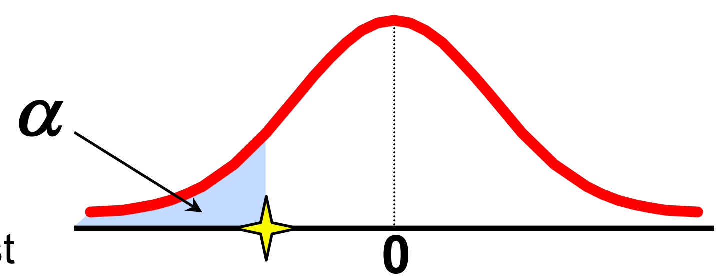
Upper-tail test



$H_0: \mu \geq 3$

$H_1: \mu < 3$

Lower-tail test



Errors in Making Decisions

- **Type I Error**
  - Reject a true null hypothesis
  - Considered a serious type of error

The probability of Type I Error is α

- Called level of significance of the test
- Set by researcher in advance

Errors in Making Decisions

(continued)

- **Type II Error**
  - Fail to reject a false null hypothesis

The probability of Type II Error is β

Outcomes and Probabilities

Possible Hypothesis Test Outcomes

	Actual Situation	
Decision	H_0 True	H_0 False
Do Not Reject **H_0**	**No error** **(1 - α)**	**Type II Error** **(β)**
Reject **H_0**	**Type I Error** **(α)**	**No Error** **(1 - β)**

Key:
Outcome
(Probability)

Type I & II Error Relationship

- Type I and Type II errors can not happen at the same time
  - Type I error can only occur if H_0 is true
  - Type II error can only occur if H_0 is false

If Type I error probability (α) ⇧, then

Type II error probability (β) ⇩

Factors Affecting Type II Error

- All else equal,
  - $\beta \uparrow$ when the difference between hypothesized parameter and its true value $\downarrow$
  - $\beta \uparrow$ when $\alpha \downarrow$
  - $\beta \uparrow$ when $\sigma \uparrow$
  - $\beta \uparrow$ when $n \downarrow$

Power of the Test

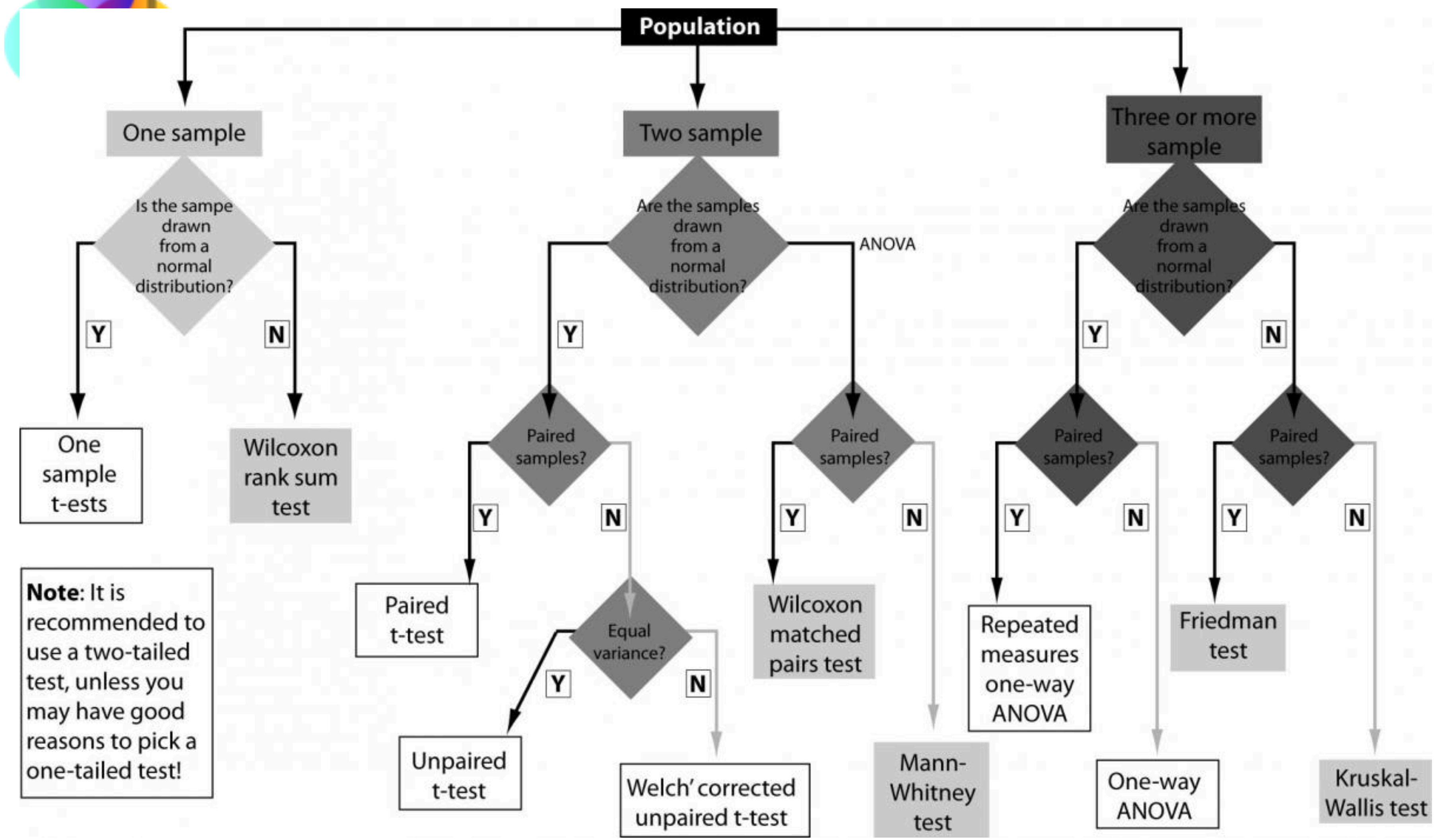
- The power of a test is the probability of rejecting a null hypothesis that is false

- i.e., Power = P(Reject H_0 | H_1 is true)

  - Power of the test increases as the sample size increases

Select your test

- Testing is a bit like finding the right recipe based on these ingredients:
  - Question
  - Data type
  - Sample size
  - Variance known? Variance of several groups equal?
- Good news: Plenty of tables available, e.g.,
  - http://www.ats.ucla.edu/stat/mult_pkg/whatstat/default.htm (with examples in R, SAS, Stata, SPSS)
  - http://sites.stat.psu.edu/~ajw13/stat500_su_res/notes/lesson14/images/summary_table.pdf



Population
One sample
Is the sampe drawn from a normal distribution?
Y
N
One sample t-ests
Wilcoxon rank sum test
Note: It is recommended to use a two-tailed test, unless you may have good reasons to pick a one-tailed test!
Two sample
Are the samples drawn from a normal distribution?
ANOVA
Y
Paired samples?
Y
N
Paired t-test
Equal variance?
Y
N
Unpaired t-test
Welch' corrected unpaired t-test
Paired samples?
Y
N
Wilcoxon matched pairs test
Mann-Whitney test
Three or more sample
Are the samples drawn from a normal distribution?
Y
N
Paired samples?
Y
N
Repeated measures one-way ANOVA
One-way ANOVA
Paired samples?
Y
N
Friedman test
Kruskal-Wallis test

Example of a table of tests

Summary Table for Statistical Techniques

	Inference	Parameter	Statistic	Type of Data	Examples	Analysis	Minitab Command	Conditions
1	Estimating a Mean	One Population Mean μ	Sample mean $\bar{y}$	Numerical	• What is the average weight of adults? • What is the average cholesterol level of adult females?	1-sample t-interval $\bar{y} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$	Stat >Basic statistics >1-sample t	• data approximately normal or • have a large sample size ($n \geq 30$)
2	Test about a Mean	One Population Mean μ	Sample mean $\bar{y}$	Numerical	• Is the average GPA of juniors at Penn State higher than 3.0? • Is the average Winter temperature in State College less than 42° F?	$H_o: \mu = \mu_o$ $H_a: \mu \neq \mu_o$ or $H_a: \mu > \mu_o$ or $H_a: \mu < \mu_o$ The one sample t test: $t = \frac{\bar{y} - \mu_0}{s/\sqrt{n}}$	Stat >Basic statistics >1-sample t	• data approximately normal or • have a large sample size ($n \geq 30$)
3	Estimating a Proportion	One Population Proportion π	Sample Proportion $\hat{\pi}$	Categorical (Binary)	• What is the proportion of males in the world? • What is the proportion of students that smoke?	1-proportion Z-interval $\hat{\pi} \pm z_{\alpha/2} \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}$	Stat >Basic statistics >1-sample proportion	• have at least 5 in each category
4	Test about a Proportion	One Population Proportion π	Sample Proportion $\hat{\pi}$	Categorical (Binary)	• Is the proportion of females different from 0.5? • Is the proportion of students who fail Stat 500 less than 0.1?	$H_o: \pi = \pi_o$ $H_a: \pi \neq \pi_o$ or $H_a: \pi > \pi_o$ or $H_a: \pi < \pi_o$ The one proportion Z-test: $z = \frac{\hat{\pi} - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$	Stat >Basic statistics >1-sample proportion	• $n\,\pi_o \geq 5$ and $n\,(1-\pi_o) \geq 5$

http://sites.stat.psu.edu/~ajw13/stat500_su_res/notes/lesson14/images/summary_table.pdf

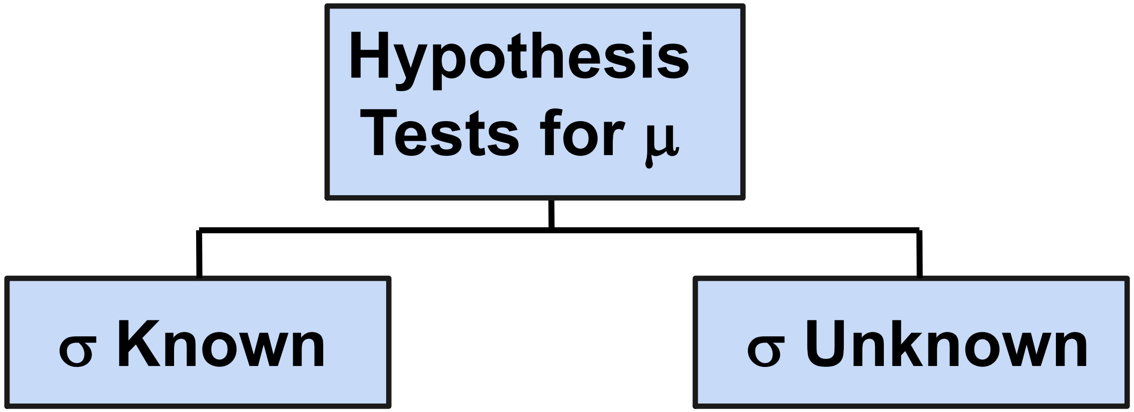
A common test: One-sample t-test

- When: Estimating a mean, comparing mean to a hypothetical value
- Requirements: Data approx. normal or sample size > 30
- Setup:

$H_0 : \mu \leq \mu_0$ $H_1 : \mu > \mu_0$	$H_0 : \mu = \mu_0$ $H_1 : \mu \neq \mu_0$	$H_0 : \mu \geq \mu_0$ $H_1 : \mu < \mu_0$
	$T = \sqrt{n}\frac{\bar{X} - \mu_0}{S} \sim t_{n-1}$	
$t = \sqrt{n}\frac{\bar{x} - \mu_0}{s}$	$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$	$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$

- Evaluation: Compare t-statistic (table, excel) to your value and accept / reject null hypothesis

Hypothesis Tests for the Mean



Test of Hypothesis for the Mean (σ Known)

- Convert sample result ($\bar{x}$) to a z value

Hypothesis Tests for μ

σ Known | **σ Unknown**

Consider the test

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

(Assume the population is normal)

The decision rule is:

$$\text{Reject } H_0 \text{ if } z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} > z_\alpha$$

Decision Rule

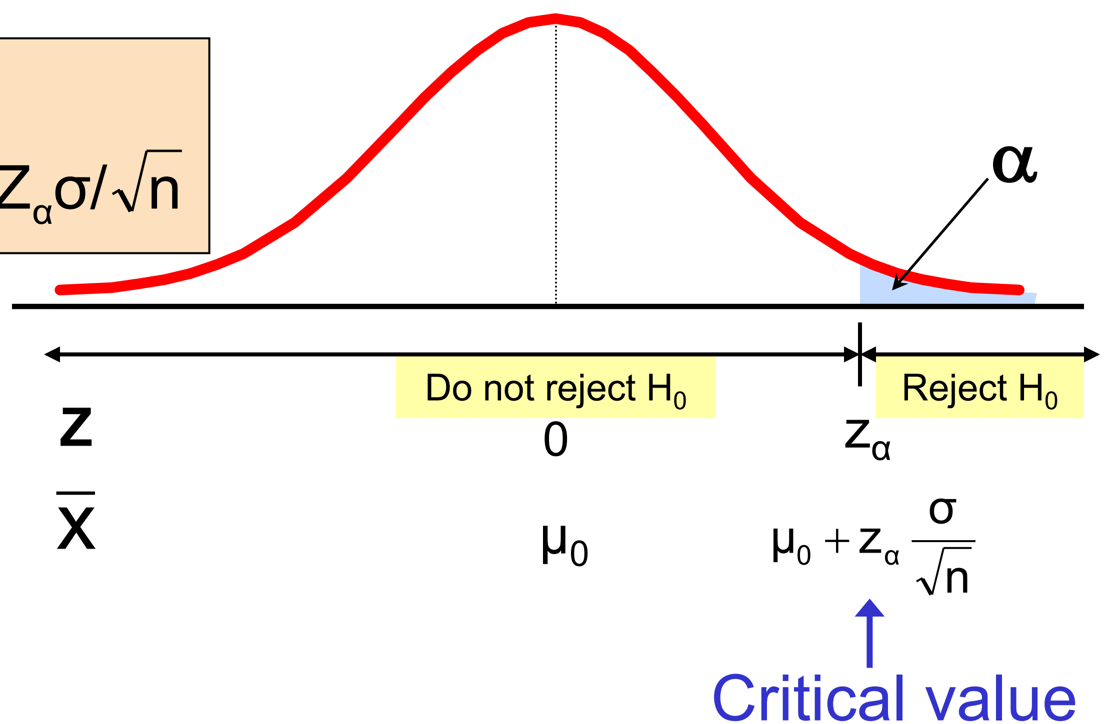
$H_0: \mu = \mu_0$

$H_1: \mu > \mu_0$

$$\text{Reject } H_0 \text{ if } z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} > z_\alpha$$

Alternate rule:

$$\text{Reject } H_0 \text{ if } \bar{X} > \mu_0 + Z_\alpha \sigma / \sqrt{n}$$



p-Value Approach to Testing

- p-value: Probability of obtaining a test statistic more extreme (≤ or ≥) than the observed sample value given H_0 is true
  - Also called observed level of significance
  - Smallest value of α for which H_0 can be rejected

p-Value Approach to Testing

(continued)

- Convert sample result (e.g., $\bar{x}$) to test statistic (e.g., z statistic)

- Obtain the p-value
  - For an upper tail test:

$$\text{p-value} = P(Z > \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}\text{, given that } H_0 \text{ is true})$$

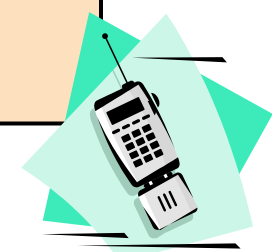
$$= P(Z > \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \mid \mu = \mu_0)$$

- Decision rule: compare the p-value to α

  - If p-value $< \alpha$, reject H_0
  - If p-value $\geq \alpha$, do not reject H_0

Example: Upper-Tail Z Test for Mean (σ Known)

A phone industry manager thinks that customer monthly cell phone bill have increased, and now average over \$52 per month. The company wishes to test this claim. (Assume σ = 10 is known)



Form hypothesis test:

H_0: $\mu \le 52$ the average is not over \$52 per month

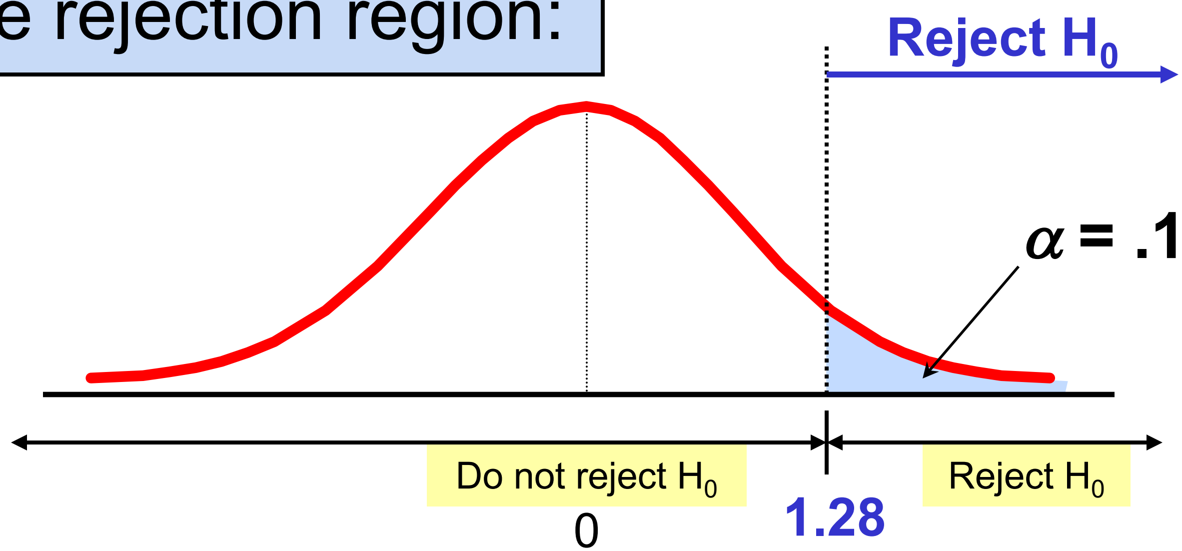
H_1: $\mu > 52$ the average **is** greater than \$52 per month (i.e., sufficient evidence exists to support the manager's claim)

Example: Find Rejection Region

(continued)

- Suppose that α = .10 is chosen for this test

Find the rejection region:



$$\text{Reject } H_0 \text{ if } z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} > 1.28$$

Example: Sample Results

(continued)

Obtain sample and compute the test statistic

Suppose a sample is taken with the following results: $n = 64$, $\bar{x} = 53.1$ (σ=10 was assumed known)

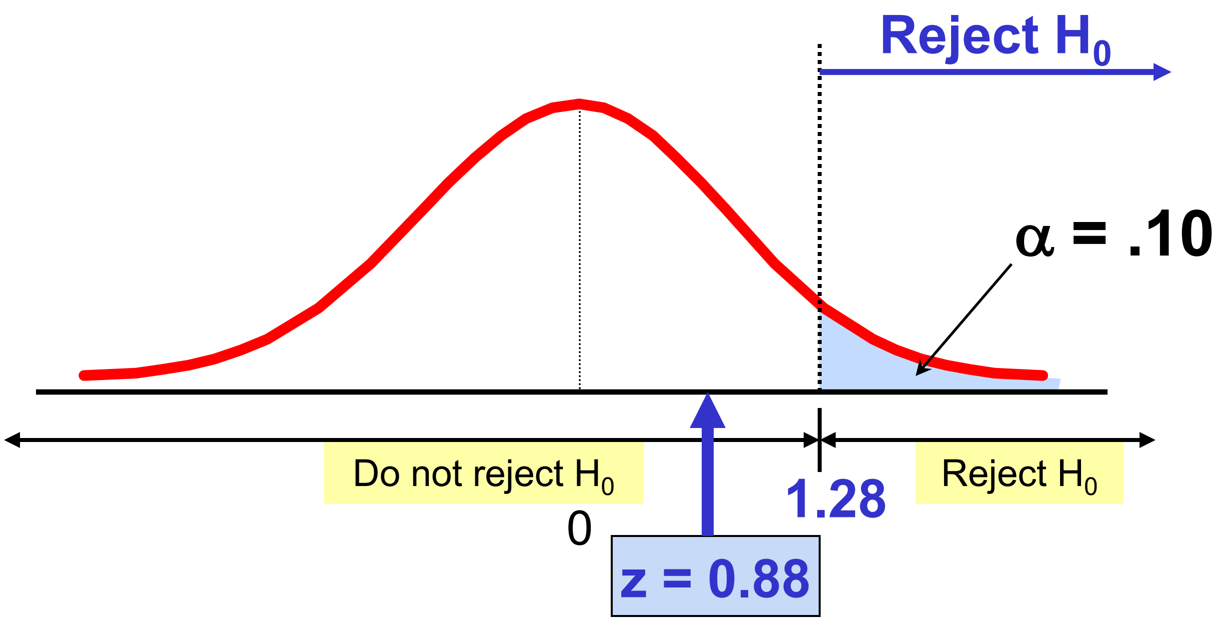
- Using the sample results,

$$z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{53.1 - 52}{\frac{10}{\sqrt{64}}} = 0.88$$

Example: Decision

(continued)

Reach a decision and interpret the result:



Do not reject H_0 since z = 0.88 < 1.28

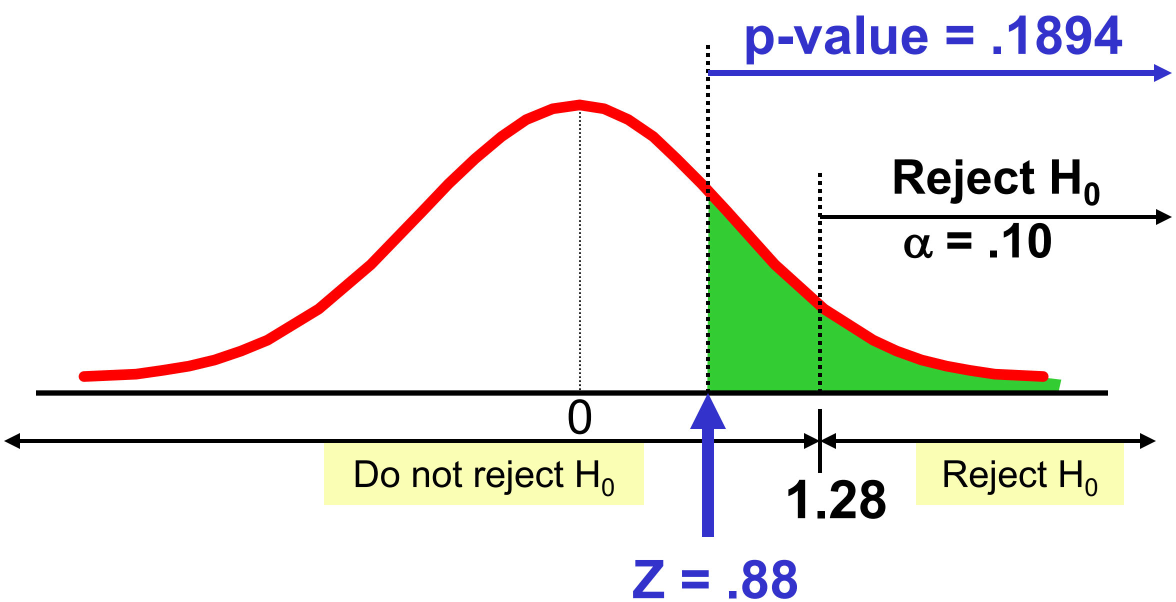
i.e.: there is not sufficient evidence that the mean bill is over $52

Example: p-Value Solution

(continued)

Calculate the p-value and compare to α

(assuming that μ = 52.0)



$$P(\bar{x} \geq 53.1 \mid \mu = 52.0)$$

$$= P\left(z \geq \frac{53.1 - 52.0}{10/\sqrt{64}}\right)$$

$$= P(z \geq 0.88) = 1 - .8106$$

$$= .1894$$

Do not reject H_0 since p-value = .1894 > α = .10

One-Tail Tests

- In many cases, the alternative hypothesis focuses on one particular direction

H_0: $\mu \leq 3$

H_1: $\mu > 3$

This is an upper-tail test since the alternative hypothesis is focused on the upper tail above the mean of 3

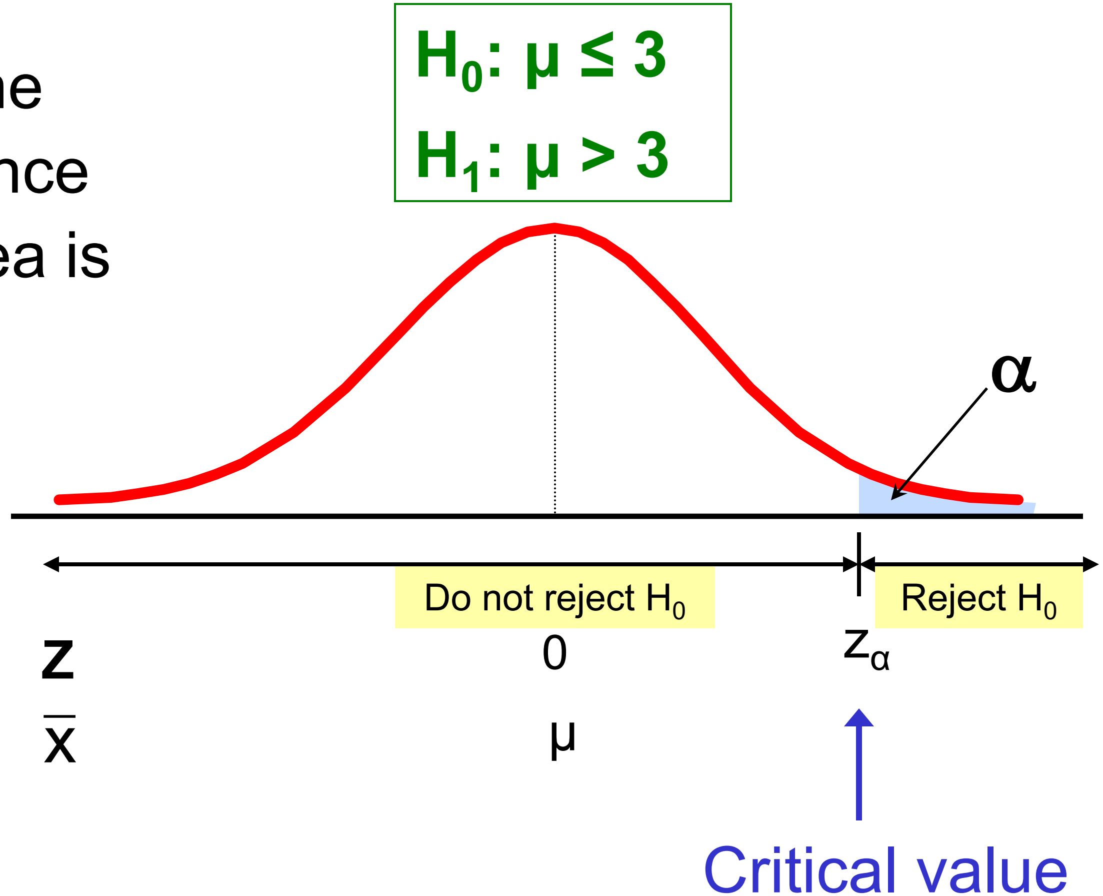
H_0: $\mu \geq 3$

H_1: $\mu < 3$

This is a lower-tail test since the alternative hypothesis is focused on the lower tail below the mean of 3

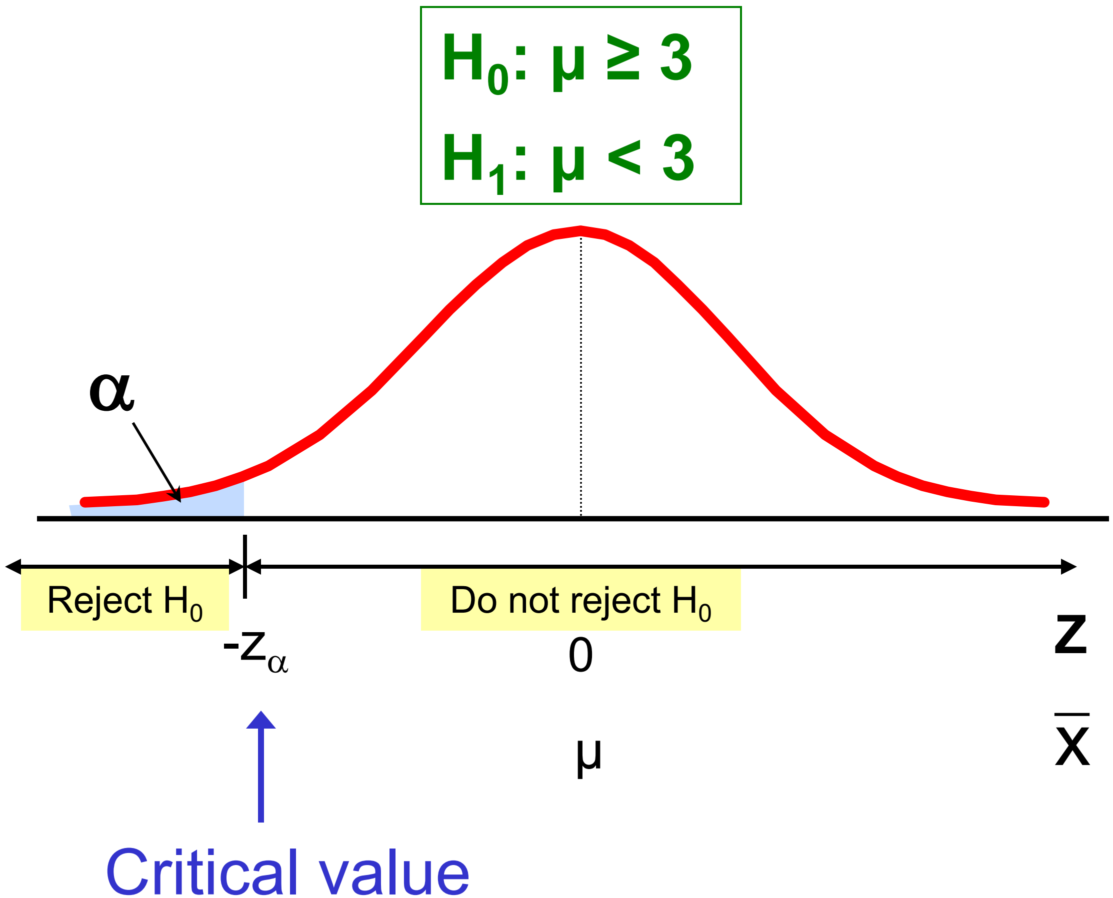
Upper-Tail Tests

- There is only one critical value, since the rejection area is in only one tail



Lower-Tail Tests

- There is only one critical value, since the rejection area is in only one tail

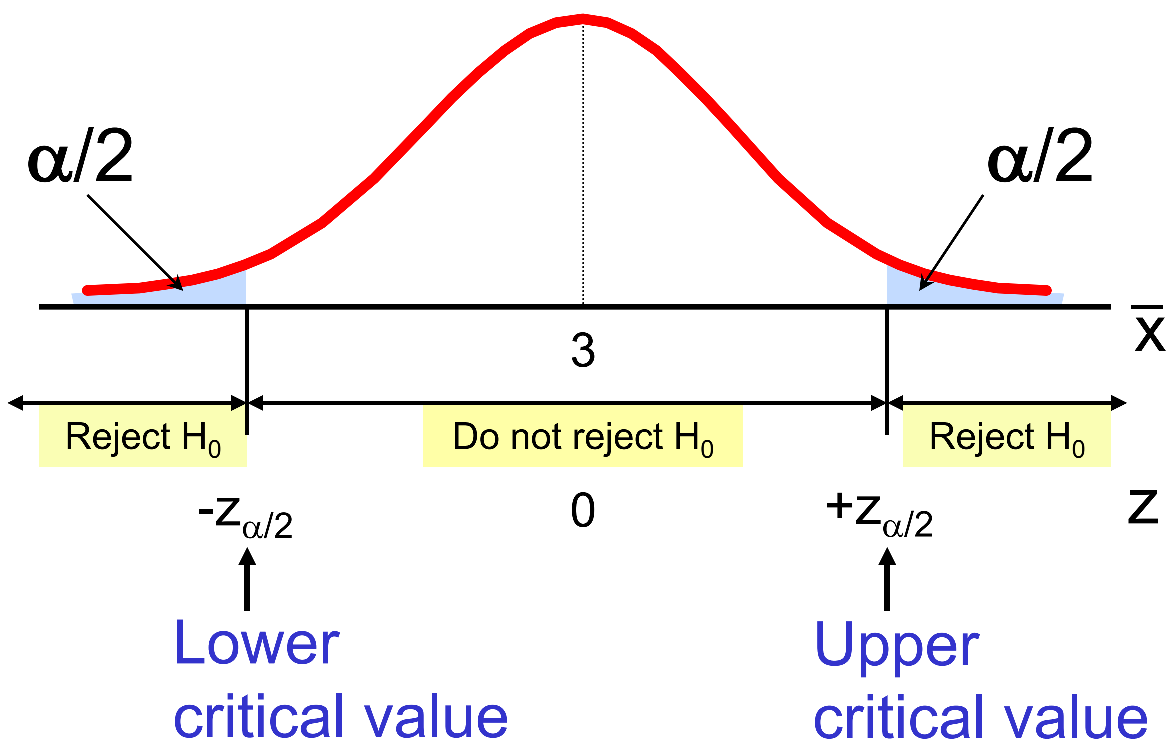


Two-Tail Tests

- In some settings, the alternative hypothesis does not specify a unique direction

H_0: $\mu = 3$
H_1: $\mu \neq 3$

- There are two critical values, defining the two regions of rejection



Hypothesis Testing Example

Test the claim that the true mean # of TV sets in US homes is equal to 3. (Assume $\sigma = 0.8$)

- State the appropriate null and alternative hypotheses
  - H_0: $\mu = 3$, H_1: $\mu \neq 3$ (This is a two tailed test)
- Specify the desired level of significance
  - Suppose that $\alpha = .05$ is chosen for this test
- Choose a sample size
  - Suppose a sample of size n = 100 is selected

Hypothesis Testing Example

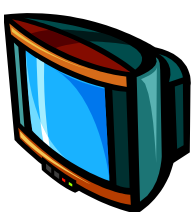
(continued)

- Determine the appropriate technique
  - σ is known so this is a z test
- Set up the critical values
  - For α = .05 the critical z values are ± 1.96
- Collect the data and compute the test statistic
  - Suppose the sample results are

    $n = 100$, $\bar{x} = 2.84$ ($\sigma = 0.8$ is assumed known)

So the test statistic is:

$$z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{2.84 - 3}{\frac{0.8}{\sqrt{100}}} = \frac{-.16}{.08} = -2.0$$

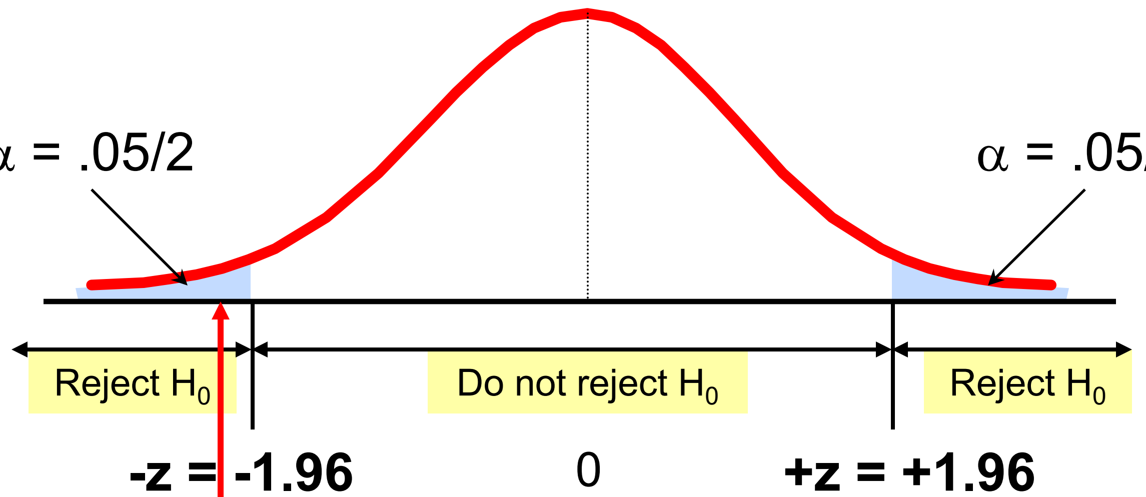


Hypothesis Testing Example

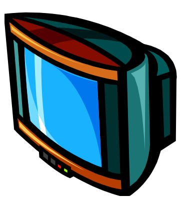
(continued)

- Is the test statistic in the rejection region?

Reject H_0 if $z < -1.96$ or $z > 1.96$; otherwise do not reject H_0



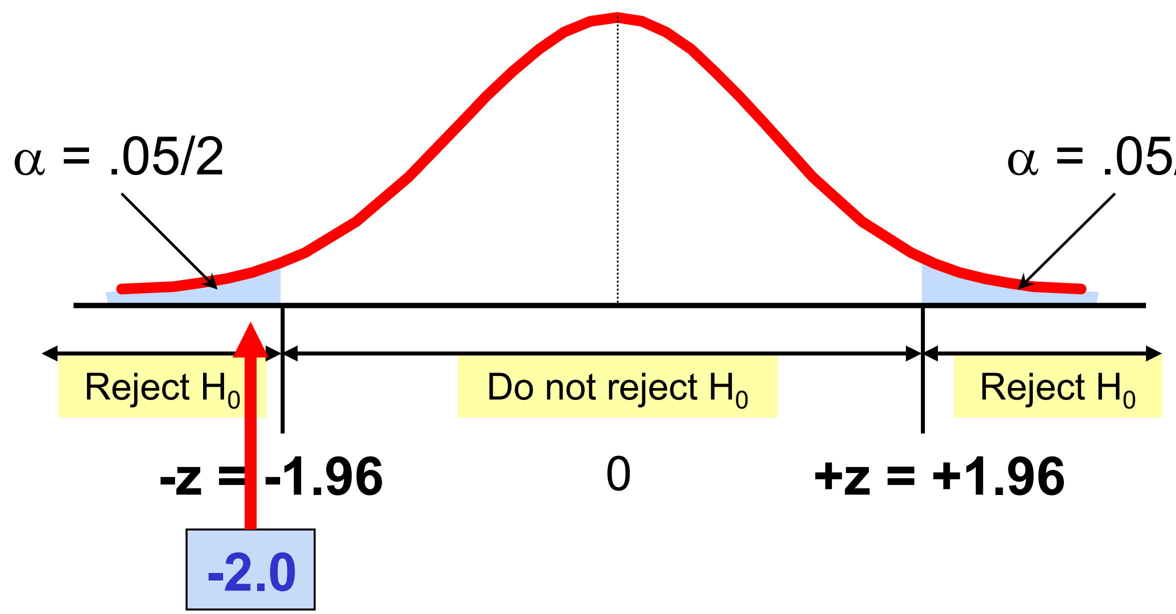
Here, $z = -2.0 < -1.96$, so the test statistic is in the rejection region



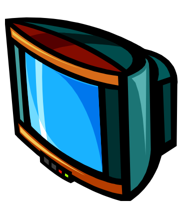
Hypothesis Testing Example

(continued)

- Reach a decision and interpret the result



Since z = -2.0 < -1.96, we <u>reject the null hypothesis</u> and conclude that there is sufficient evidence that the mean number of TVs in US homes is not equal to 3



Example: p-Value

- Example: How likely is it to see a sample mean of 2.84 (or something further from the mean, in either direction) if the true mean is $\mu = 3.0$?

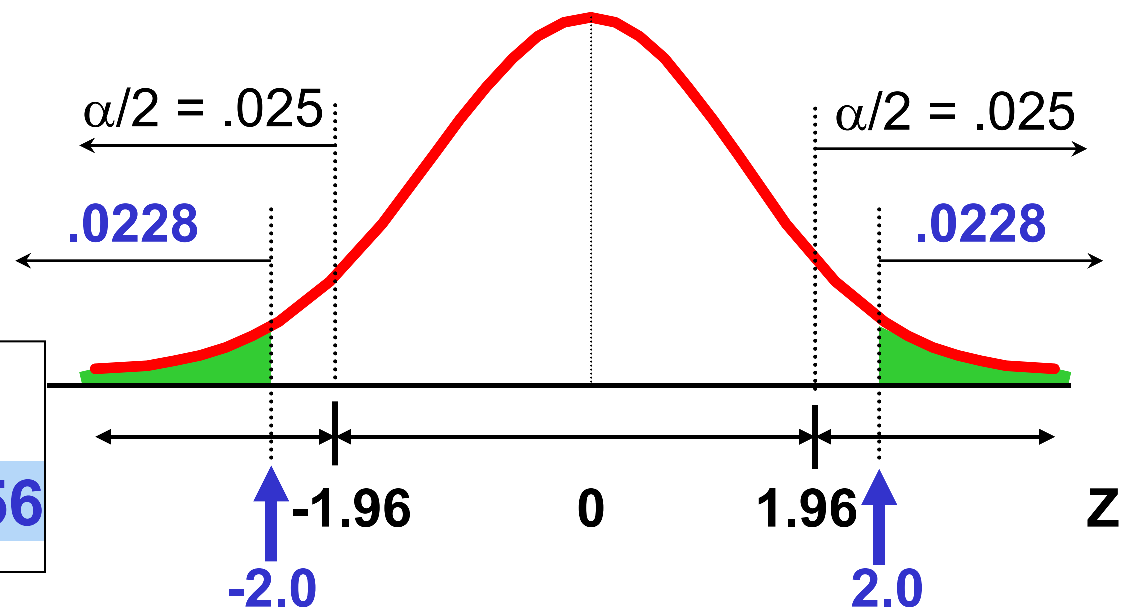
$\bar{x} = 2.84$ is translated to a z score of $z = -2.0$

$P(z < -2.0) = .0228$

$P(z > 2.0) = .0228$

p-value

= .0228 + .0228 = .0456



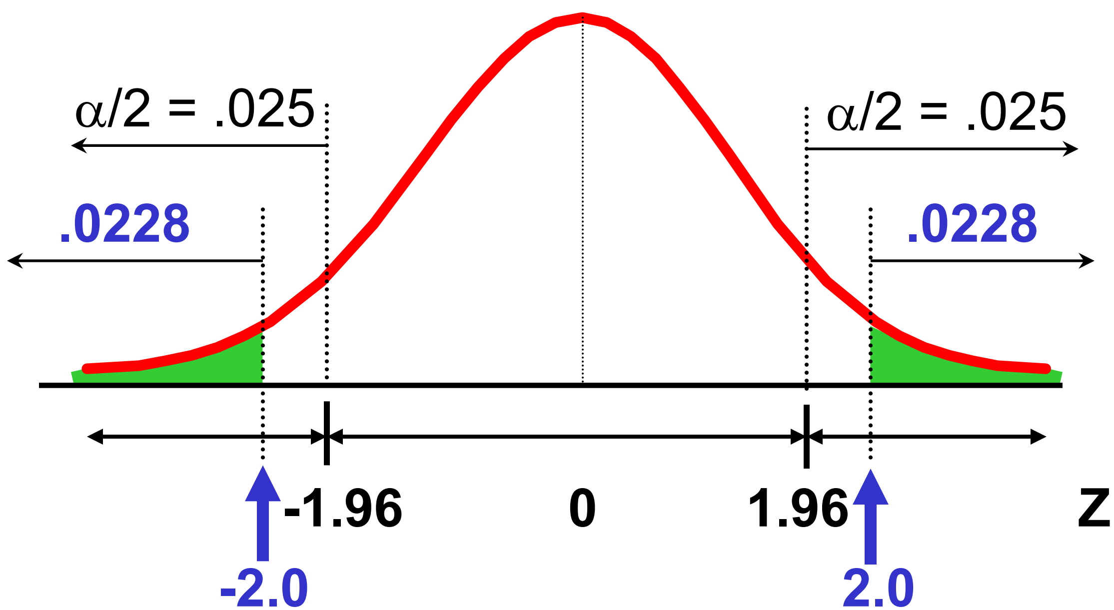
Example: p-Value

(continued)

- Compare the p-value with α
  - If p-value $<$ α , reject H_0
  - If p-value $\geq$ α , do not reject H_0

Here: p-value = .0456
α = .05

Since .0456 < .05, we reject the null hypothesis



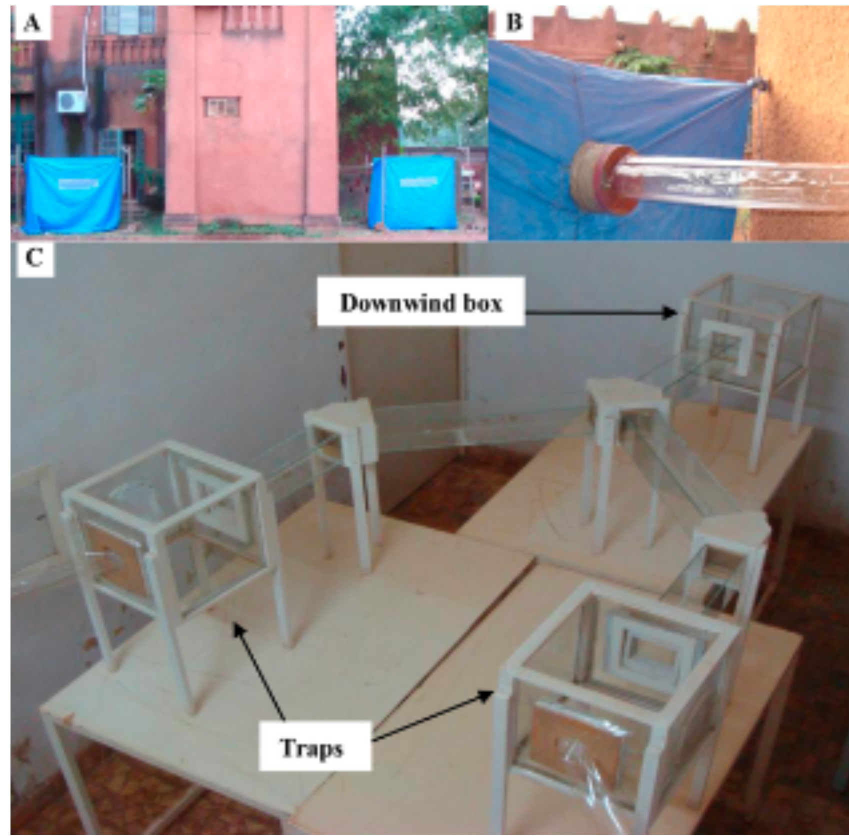
An Alternative Way

Beer Consumption [redacted] Human Attractiveness to Malaria Mosquitoes

Beer (25):
27 20 21 26 27 31 24 21 20 19
23 24 28 19 24 29 18 20 17 31
20 25 28 21 27
Mean: 23.6

Water (18):
21 22 15 12 21 16 19 15 22 24
19 23 13 22 20 24 18 20
Mean: 19.2

Is a difference of 4.4 significant?



Beer Consumption #Human Attractiveness to Malaria Mosquitoes
Thierry Lefèvre , Louis-Clément Gouagna, Kounbobr Roch Dabiré, Eric Elguero, Didier Fontenille, François Renaud, Carlo Costantini, Frédéric Thomas

Permutation Test

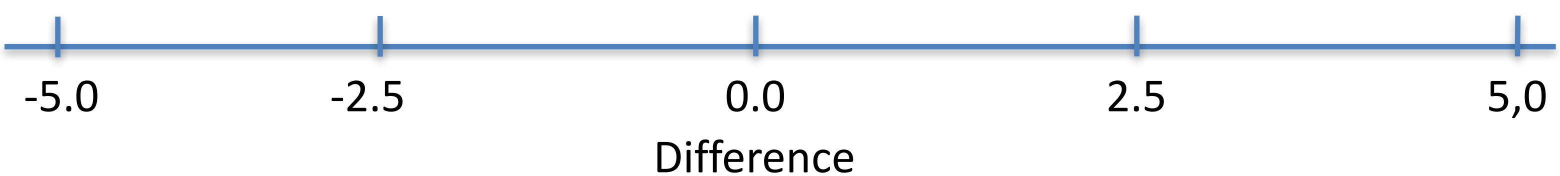
Beer (25)

27	23	20	31	29
20	24	25	24	18
21	28	28	21	20
26	19	21	20	17
27	24	27	19	31

Water (18)

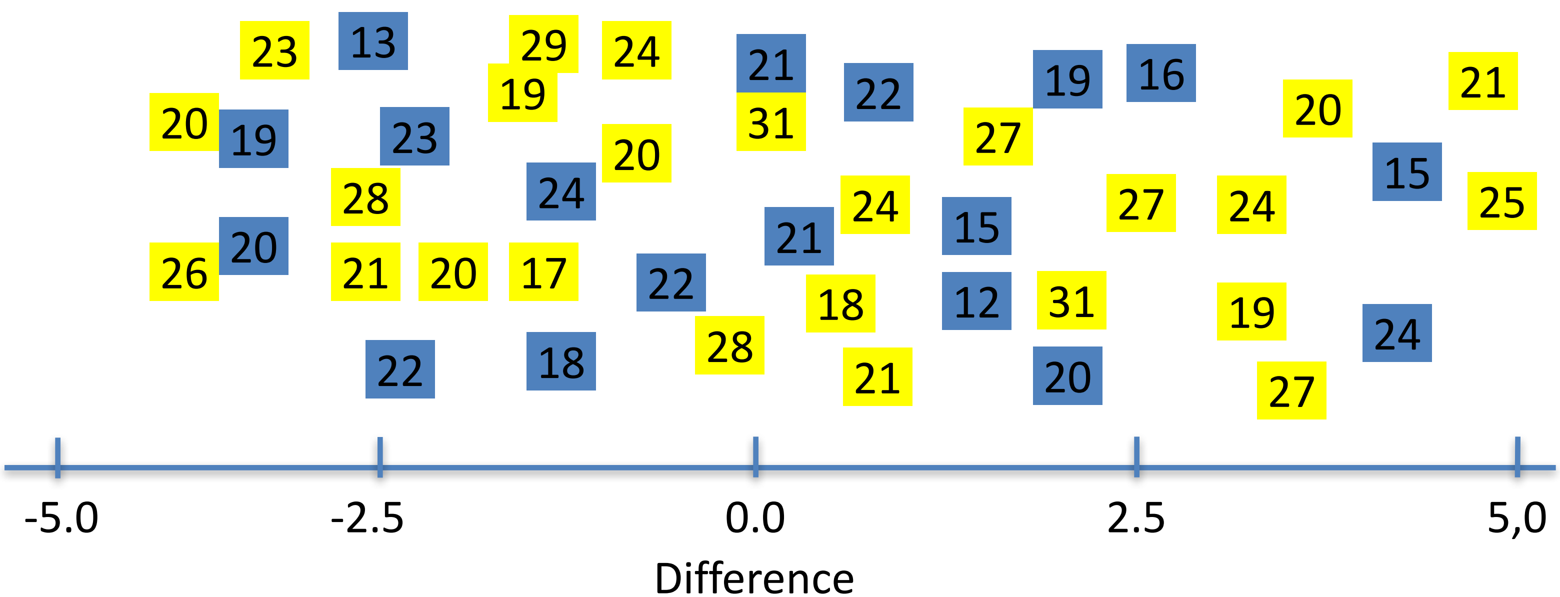
21	19	16	24
22	23	19	18
15	13	15	20
12	22	22	
21	20	24	

Difference: 4.4



Permutation Test

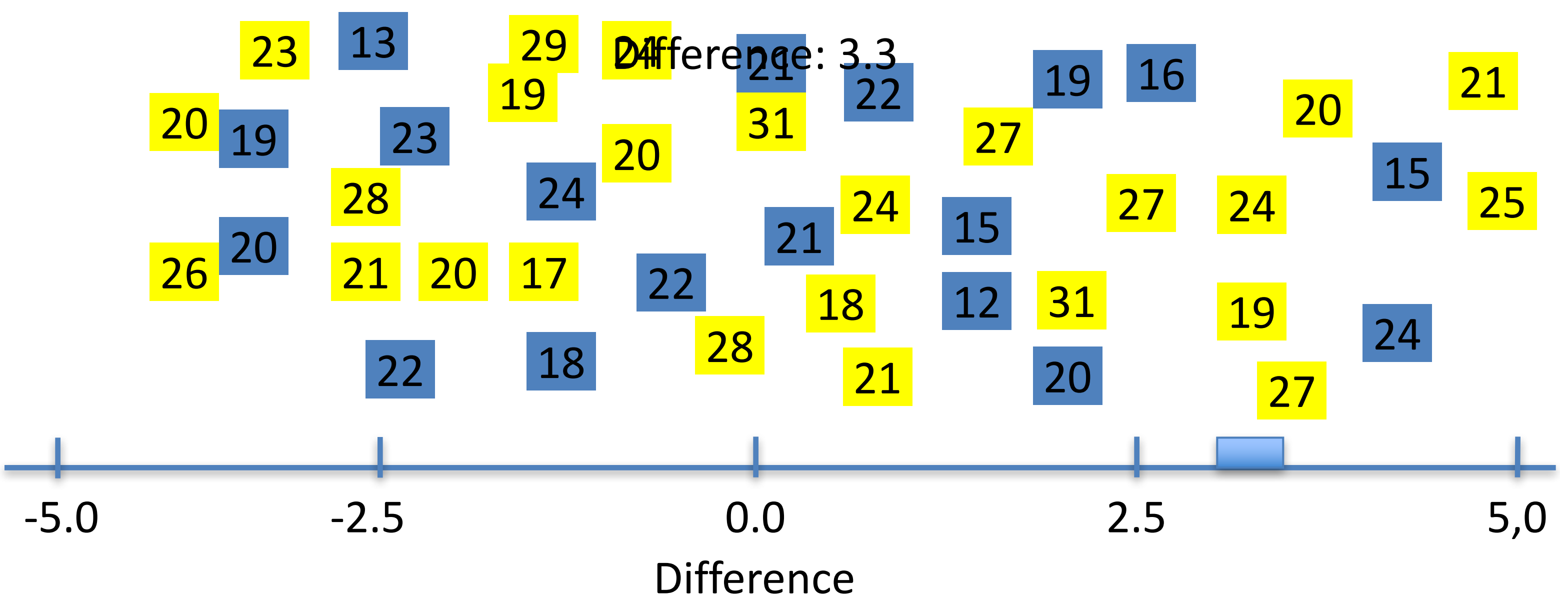
Beer (25) Water (18)



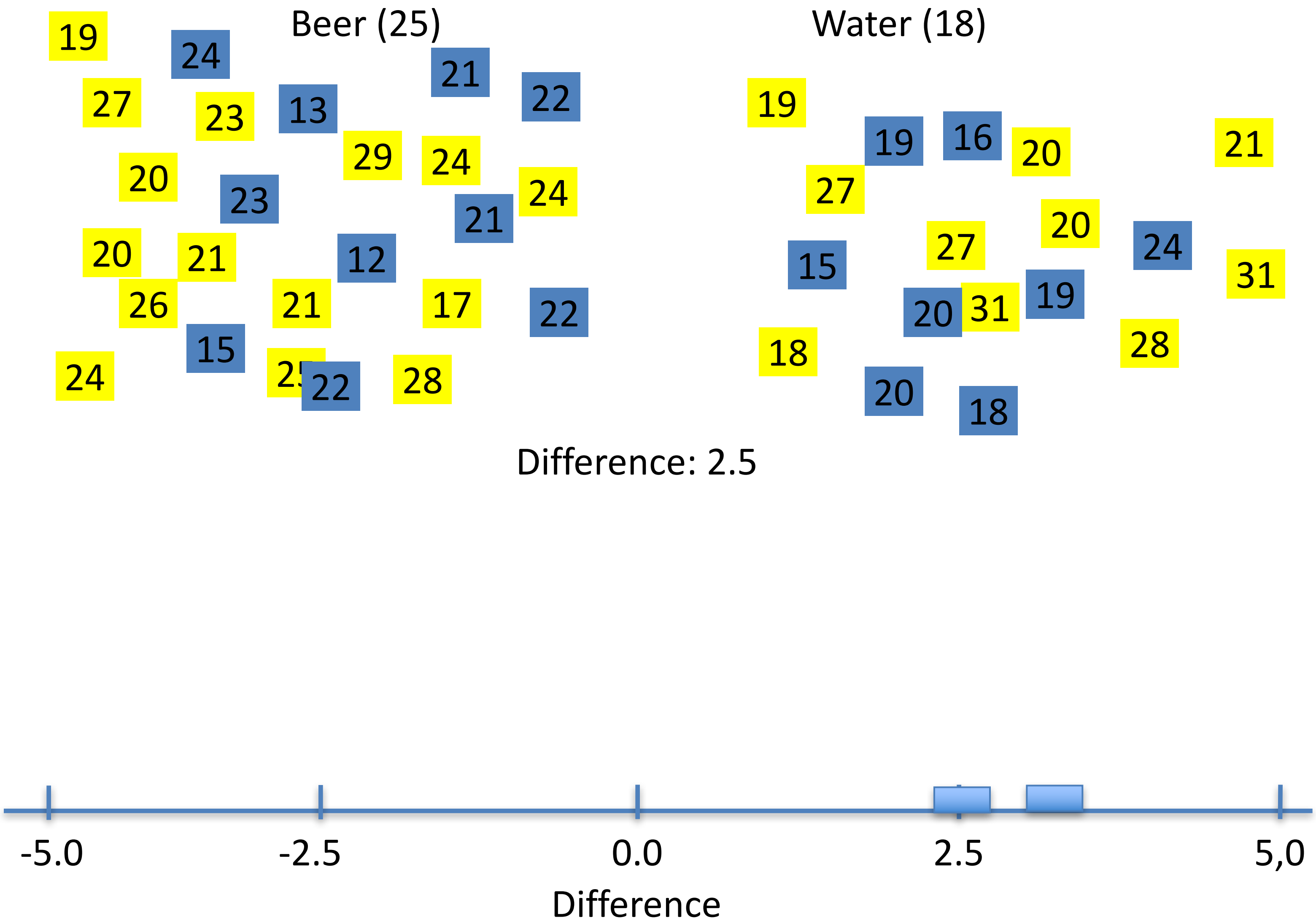
Permutation Test

Beer (25)

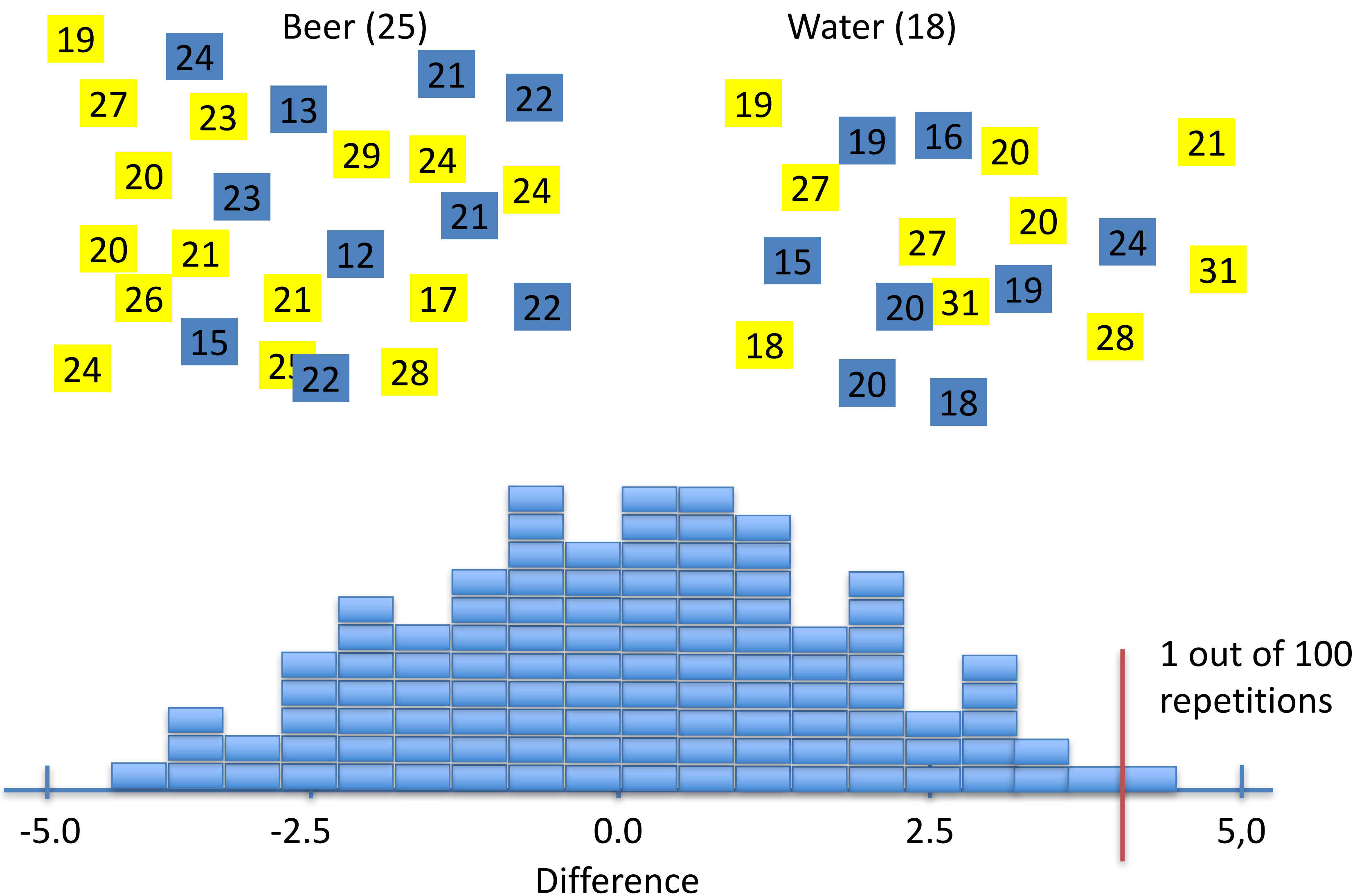
Water (18)



Permutation Test



Permutation Test



t Test of Hypothesis for the Mean (σ Unknown)

- Convert sample result ($\bar{x}$) to a t test statistic

Hypothesis Tests for μ

σ Known | σ Unknown

Consider the test

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

(Assume the population is normal)

The decision rule is:

$$\text{Reject } H_0 \text{ if } t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} > t_{n-1,\alpha}$$

t Test of Hypothesis for the Mean (σ Unknown)

(continued)

- For a two-tailed test:

  Consider the test

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

(Assume the population is normal, and the population variance is unknown)

The decision rule is:

Reject H_0 if $t = \dfrac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} < -t_{n-1,\,\alpha/2}$ or if $t = \dfrac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} > t_{n-1,\,\alpha/2}$

What is a T-distribution?

- A t-distribution is like a Z distribution, except has slightly fatter tails to reflect the uncertainty added by estimating σ.
- The bigger the sample size (i.e., the bigger the sample size used to estimate σ), then the closer t becomes to Z.
- If $n>100$, t approaches Z.

The T probability density function

$$f(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{\sqrt{\nu\pi}} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$$

Where:

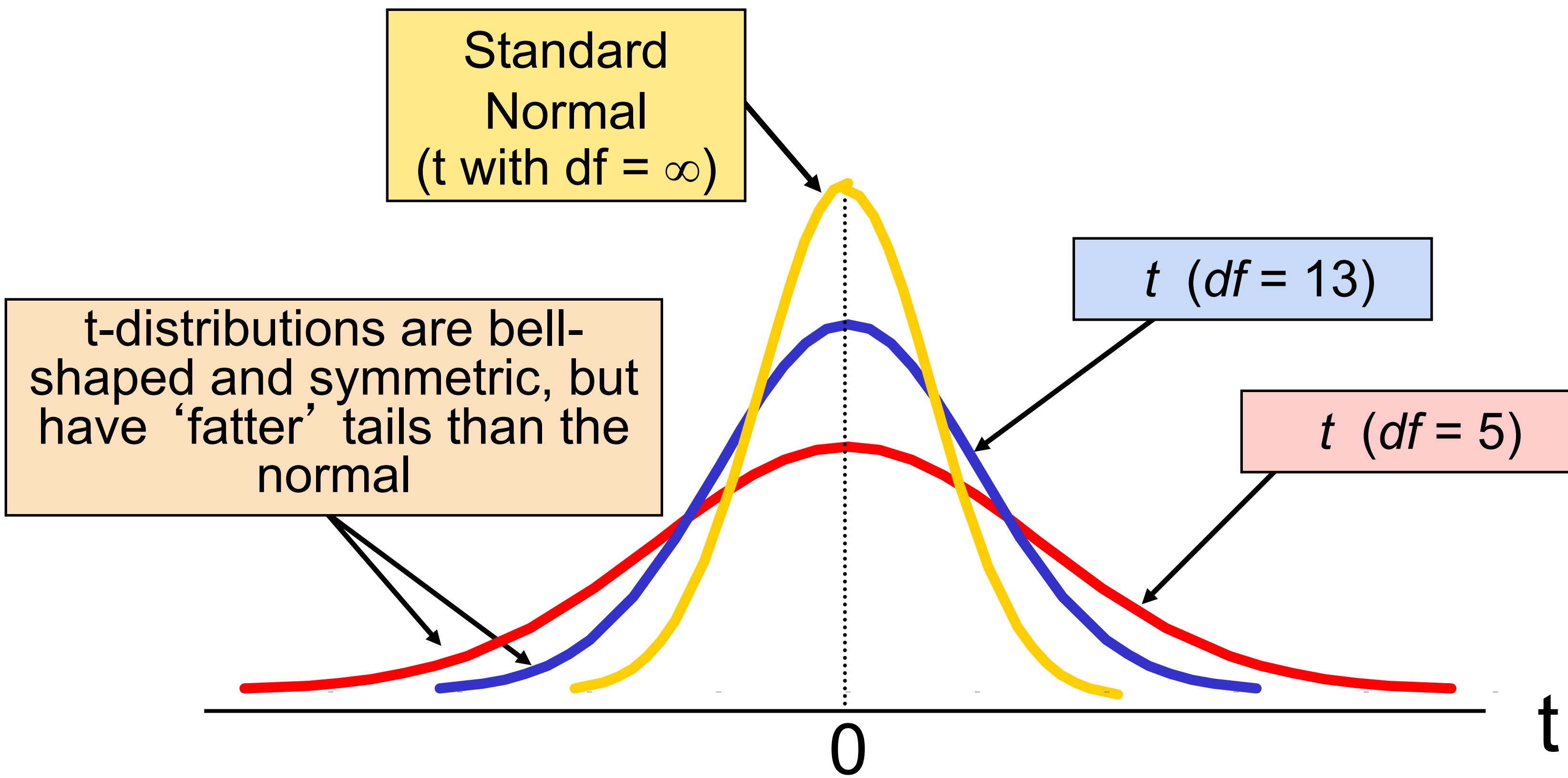
v is the degrees of freedom

(gamma) is the Gamma function

is the constant Pi (3.14...)

Student's t Distribution

Note: $t \longrightarrow Z$ as n increases



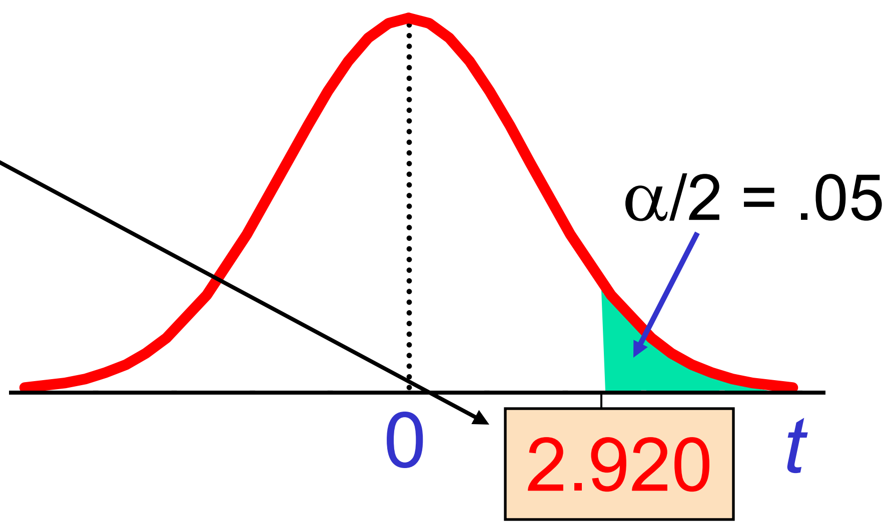
from "Statistics for Managers" Using Microsoft® Excel 4th Edition, Prentice-Hall 2004

Student's t Table

	Upper Tail Area		
df	.25	.10	.05
1	1.000	3.078	6.314
2	0.817	1.886	2.920
3	0.765	1.638	2.353

The body of the table contains t values, not probabilities

Let: n = 3
df = n - 1 = 2
α = .10
$\alpha/2$ =.05



from "Statistics for Managers" Using Microsoft® Excel 4th Edition, Prentice-Hall 2004

t distribution values

With comparison to the Z value

Confidence Level	t (10 d.f.)	t (20 d.f.)	t (30 d.f.)	Z
.80	1.372	1.325	1.310	1.28
.90	1.812	1.725	1.697	1.64
.95	2.228	2.086	2.042	1.96
.99	3.169	2.845	2.750	2.58

Note: $t \longrightarrow Z$ as n increases

from "Statistics for Managers" Using Microsoft® Excel 4th Edition, Prentice-Hall 2004

Summary: Single population mean (small n, normality)

- Hypothesis test:

$$t_{n-1} = \frac{\text{observed mean} - \text{null mean}}{\frac{s_x}{\sqrt{n}}}$$

- Confidence Interval

$$\text{confidence interval} = \text{observed mean} \pm \mathrm{t}_{\mathrm{n\text{-}1},\alpha/2} * \left(\frac{s_x}{\sqrt{n}}\right)$$

Summary: Single population mean (large n)

- Hypothesis test:

$$Z \cong t_{n-1} = \frac{\text{observed mean} - \text{null mean}}{\frac{s_x}{\sqrt{n}}}$$

- Confidence Interval

$$\text{confidence interval} = \text{observed mean} \pm [\, \mathrm{t}_{\mathrm{n\text{-}1},\alpha/2} \cong \mathrm{Z}_{\alpha/2}\,] * (\frac{s_x}{\sqrt{n}})$$

Example: Two-Tail Test (σ Unknown)

The average cost of a hotel room in New York is said to be \$168 per night. A random sample of 25 hotels resulted in $\bar{x}$ = \$172.50 and

s = \$15.40. Test at the

α = 0.05 level.

(Assume the population distribution is normal)

H_0: $\mu = 168$

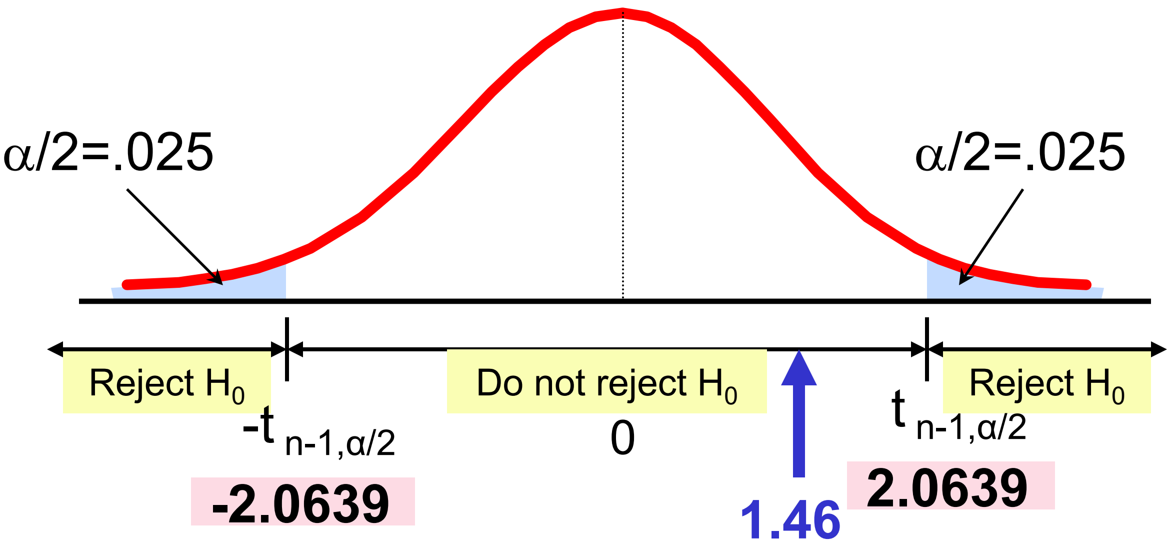
H_1: $\mu \neq 168$

Example Solution: Two-Tail Test

H_0: $\mu = 168$
H_1: $\mu \neq 168$

- $\alpha = 0.05$
- $n = 25$
- σ is unknown, so use a **t statistic**
- Critical Value:

$t_{24\,,\,.025} = \pm\, 2.0639$



$$t_{n-1} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{172.50 - 168}{\frac{15.40}{\sqrt{25}}} = 1.46$$

Do not reject H_0: not sufficient evidence that true mean cost is different than $168

Tests of the Population Proportion

- Involves categorical variables
- Two possible outcomes
  - "Success" (a certain characteristic is present)
  - "Failure" (the characteristic is not present)
- Fraction or proportion of the population in the "success" category is denoted by P
- Assume sample size is large

Proportions

(continued)

- Sample proportion in the success category is denoted by $\hat{p}$
  - $$\hat{p} = \frac{x}{n} = \frac{\text{number of successes in sample}}{\text{sample size}}$$
- When $nP(1 - P) > 9$, $\hat{p}$ can be approximated by a normal distribution with mean and standard deviation
  - $$\mu_{\hat{p}} = P \qquad \sigma_{\hat{p}} = \sqrt{\frac{P(1-P)}{n}}$$

Hypothesis Tests for Proportions

- The sampling distribution of $\hat{p}$ is approximately normal, so the test statistic is a z value:

$$z = \frac{\hat{p} - P_0}{\sqrt{\frac{P_0(1 - P_0)}{n}}}$$

Hypothesis Tests for P

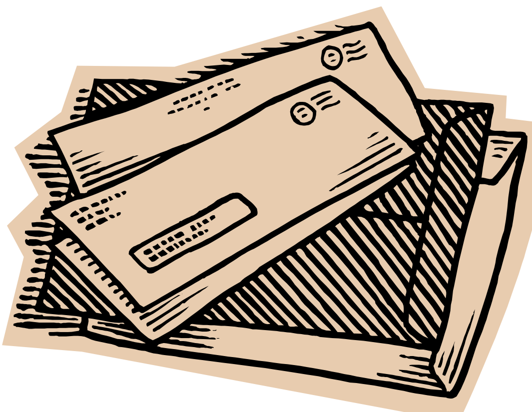
$nP(1 - P) > 9$

$nP(1 - P) < 9$

Exact test

Example: Z Test for Proportion

A marketing company claims that it receives 8% responses from its mailing. To test this claim, a random sample of 500 were surveyed with 25 responses. Test at the $\alpha = .05$ significance level.



Check:

Our approximation for P is

$\hat{p} = 25/500 = .05$

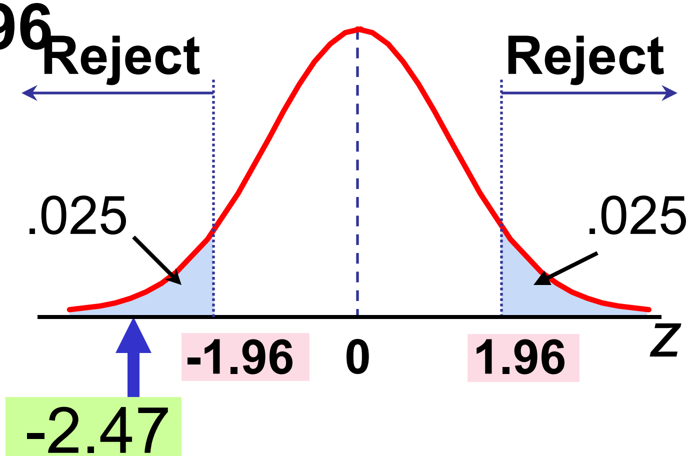
$nP(1 - P) = (500)(.05)(.95)$
$= 23.75 > 9$ ✓

Z Test for Proportion: Solution

H_0: $P = .08$
H_1: $P \neq .08$

α = .05
n = 500, $\hat{p}$ = .05

Critical Values: ± 1.96



Test Statistic:

$$z = \frac{\hat{p} - P_0}{\sqrt{\frac{P_0(1-P_0)}{n}}} = \frac{.05 - .08}{\sqrt{\frac{.08(1-.08)}{500}}} = -2.47$$

Decision:

Reject H_0 at α = .05

Conclusion:

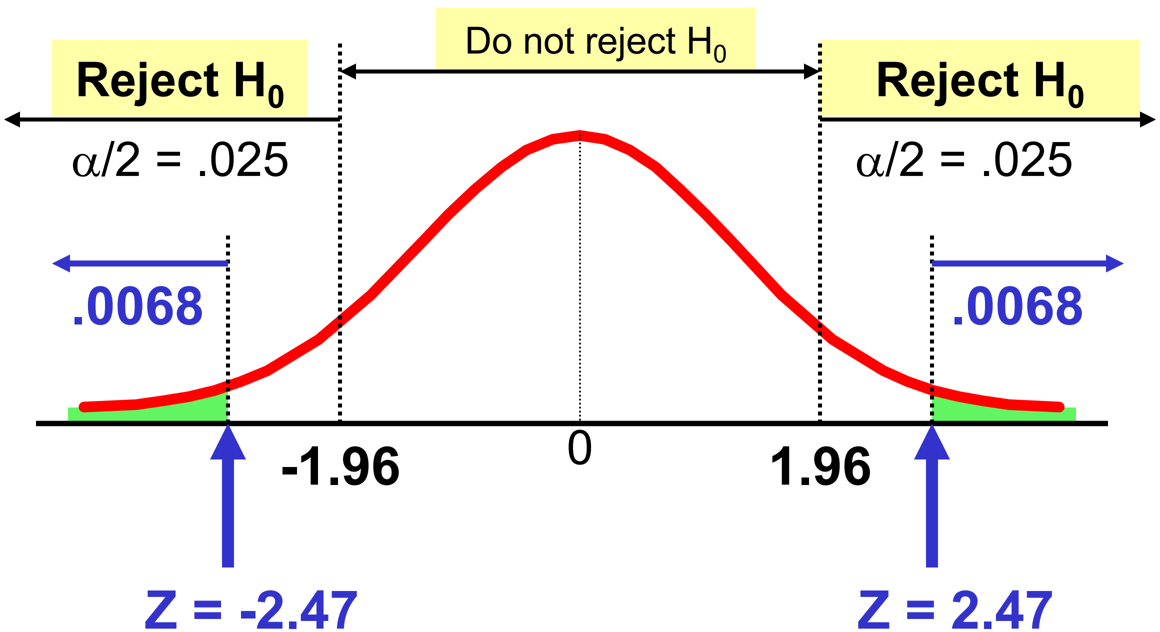
There is sufficient evidence to reject the company's claim of 8% response rate.

p-Value Solution

(continued)

Calculate the p-value and compare to α

(For a two sided test the p-value is always two sided)



p-value = .0136:

$$P(Z \le -2.47) + P(Z \ge 2.47)$$

$$= 2(.0068) = 0.0136$$

Reject H_0 since p-value = .0136 < α = .05

Power of the Test

- Recall the possible hypothesis test outcomes:

Key:
Outcome
(Probability)

	Actual Situation	
Decision	H_0 True	H_0 False
Do Not Reject H_0	**No error** **(1 - α)**	**Type II Error** **(β)**
Reject H_0	**Type I Error** **(α)**	**No Error** **(1 - β)**

- β denotes the probability of Type II Error
- 1 – β is defined as the power of the test

Power = 1 – β = the probability that a false null hypothesis is rejected

Type II Error

Assume the population is normal and the population variance is known. Consider the test

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

The decision rule is:

$$\text{Reject } H_0 \text{ if } z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} > z_\alpha$$

or

$$\text{Reject } H_0 \text{ if } \bar{x} = \bar{x}_c > \mu_0 + Z_\alpha \sigma / \sqrt{n}$$

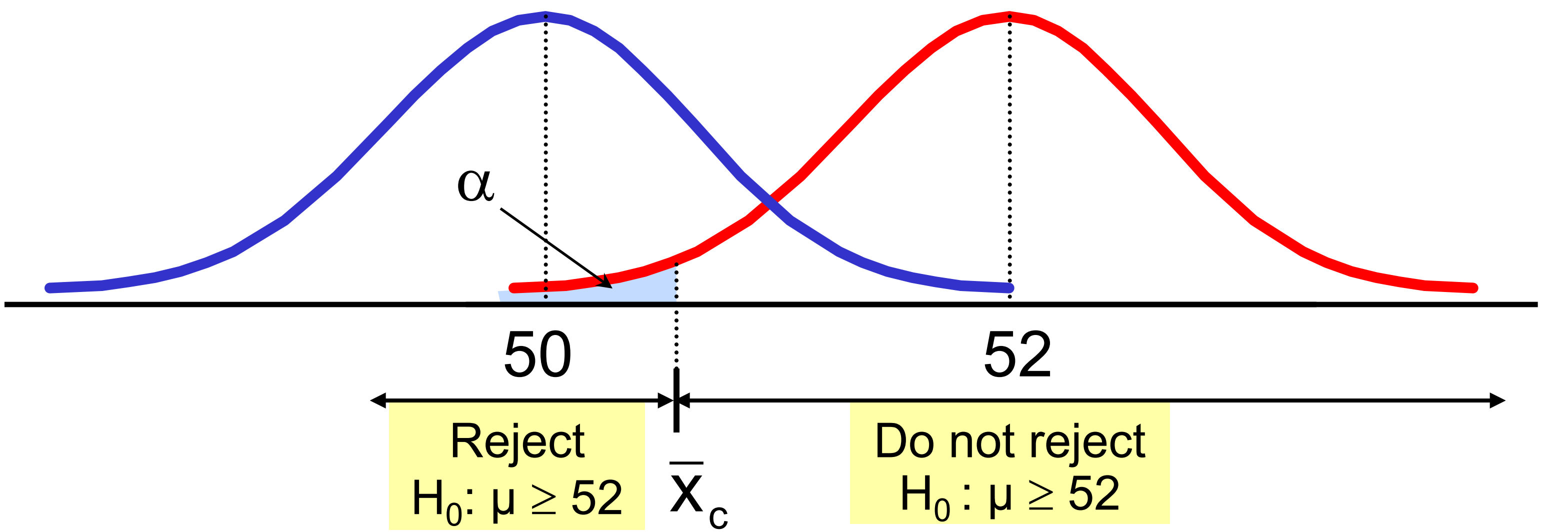
If the null hypothesis is false and the true mean is μ*, then the probability of type II error is

$$\beta = P(\bar{x} < \bar{x}_c \mid \mu = \mu^*) = P\left(z < \frac{\bar{x}_c - \mu^*}{\sigma / \sqrt{n}} \right)$$

Type II Error Example

- Type II error is the probability of failing to reject a false H_0

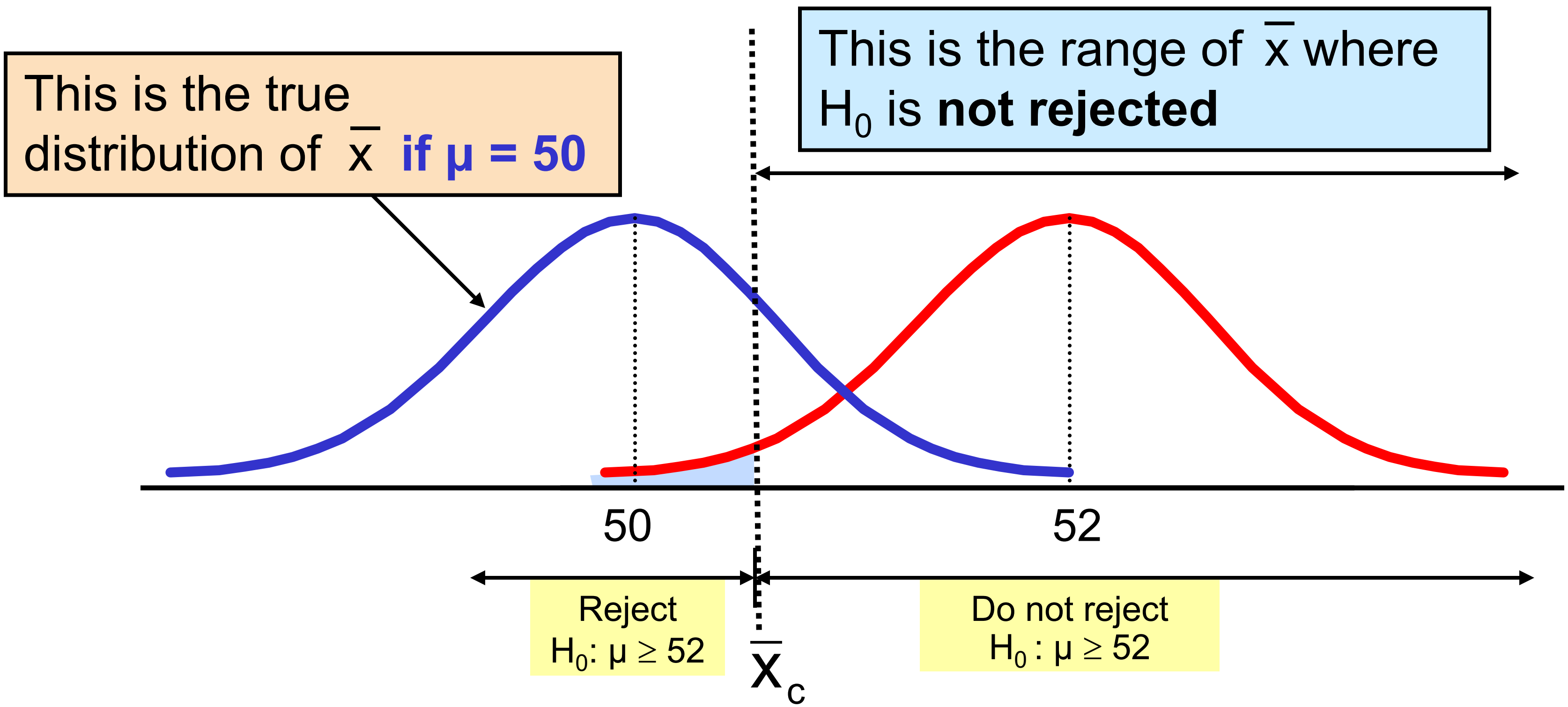
Suppose we fail to reject **H_0: $\mu \geq 52$**
when in fact the true mean is **$\mu^* = 50$**



Type II Error Example

(continued)

- Suppose we do not reject H_0: $\mu \geq 52$ when in fact the true mean is $\mu^* = 50$

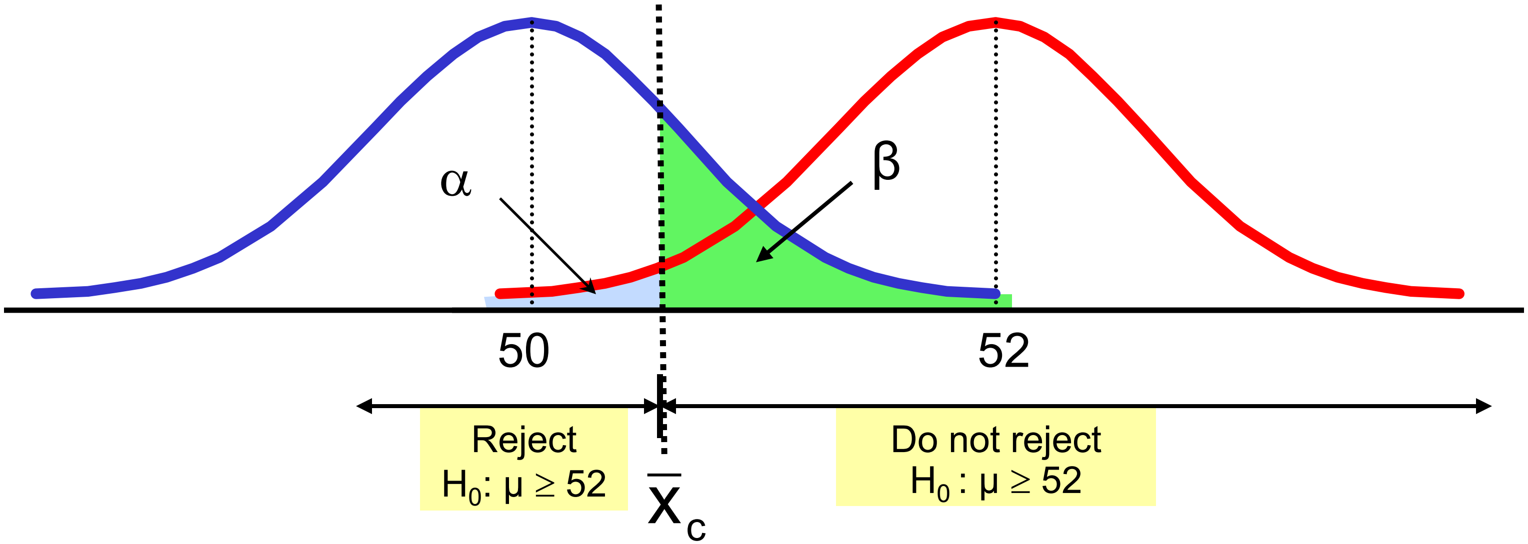


Type II Error Example

(continued)

- Suppose we do not reject H_0: $\mu \geq 52$ when in fact the true mean is $\mu^* = 50$

Here, $\beta = P(\bar{x} \geq \bar{x}_c)$ if $\mu^* = 50$



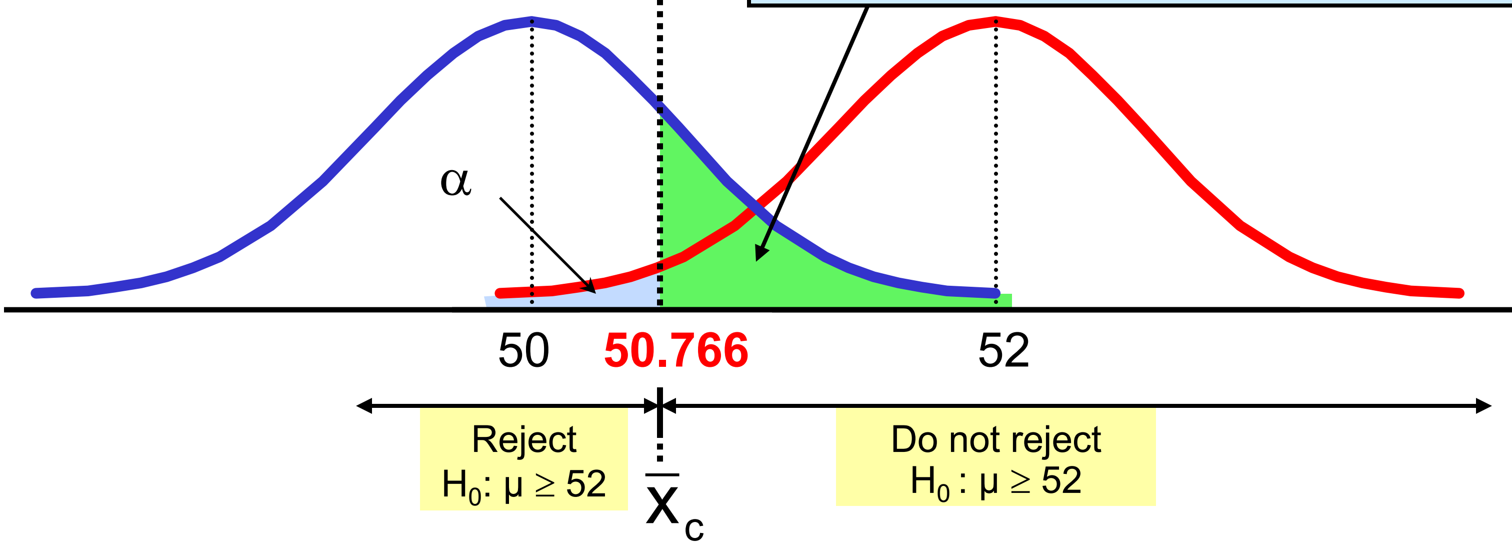
Calculating β

- Suppose n = 64 , σ = 6 , and α = .05

$$\bar{x}_c = \mu_0 - z_\alpha \frac{\sigma}{\sqrt{n}} = 52 - 1.645 \frac{6}{\sqrt{64}} = \boxed{50.766}$$

(for $H_0 : \mu \geq 52$)

So $\beta = P(\bar{x} \geq 50.766)$ if $\mu^* = 50$

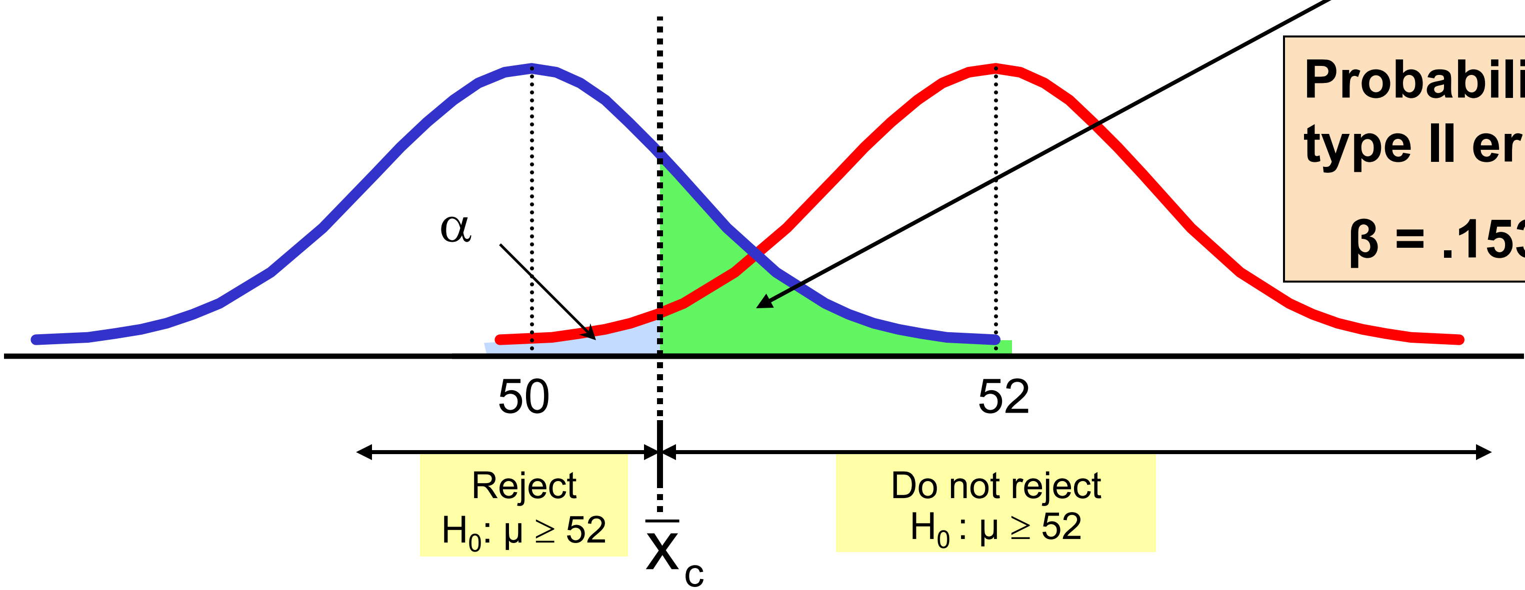


Calculating β

(continued)

- Suppose n = 64 , σ = 6 , and α = .05

$$P(\bar{x} \geq 50.766 \,|\, \mu^* = 50) = P\left(z \geq \frac{50.766 - 50}{6/\sqrt{64}} \right) = P(z \geq 1.02) = .5 - .3461 = \boxed{.1539}$$



Probability of type II error:

β = .1539

Power of the Test Example

If the true mean is $\mu^* = 50$,

- The probability of Type II Error = $\beta = 0.1539$
- The power of the test = $1 - \beta = 1 - 0.1539 = 0.8461$

Key:
Outcome
(Probability)

	Actual Situation	
Decision	H_0 True	H_0 False
Do Not Reject H_0	**No error** $1 - \alpha = 0.95$	**Type II Error** $\beta = 0.1539$
Reject H_0	**Type I Error** $\alpha = 0.05$	**No Error** $1 - \beta = 0.8461$

(The value of β and the power will be different for each μ^*)