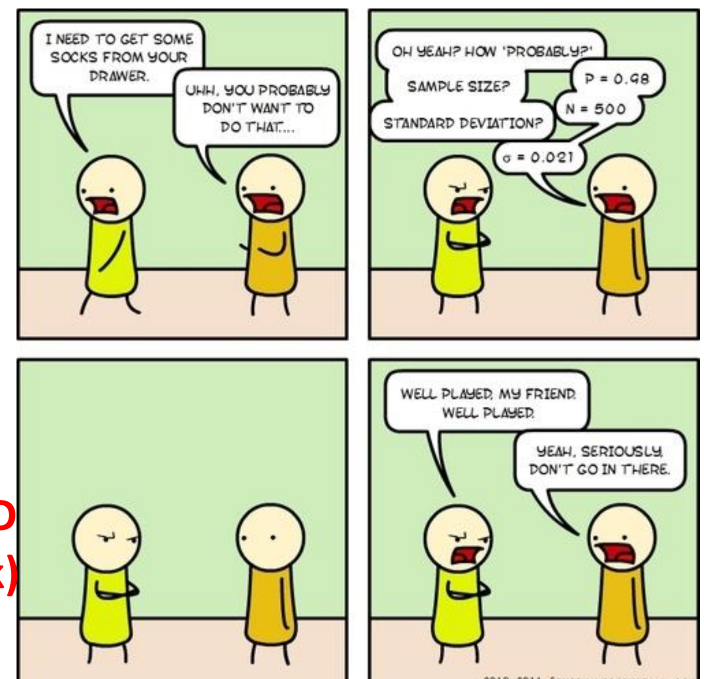
Past weeks:
Measures of central tendency (mean, mode, median)
Measures of dispersion (standard deviation, variance, range, etc).
Working with the normal curve

Last week:
Sample, population and sampling distributions
finished with estimation & “confidence intervals”…

Today: move onto Chapter 7:
Hypothesis Testing and the One Sample test..

AFTER TODAY< YOU CAN COMPLETE YOUR WRITTEN ASSIGNMENT # 1 (WHICH HAS BEEN POSTED ON MY WEBSITE): DUE NEXT CLASS (after study week)



Recall with Confidence Intervals, I gave you three formulas that are appropriate when calculating:

1. Working with means: when we know our "population standard deviation"

FORMULA 6.1

$$\text{c.i.} = \overline{X} \pm Z\left(\frac{\sigma}{\sqrt{n}}\right)$$

where c.i. = confidence interval

$\overline{X}$ = the sample mean

Z = the Z score as determined by the alpha level

$\frac{\sigma}{\sqrt{n}}$ = the standard deviation of the sampling distribution or the standard error of the mean

2. Working with means: when we <u>do not know</u> our "population standard deviation" but do know our sample standard deviation

FORMULA 6.2

$$\text{c.i.} = \overline{X} \pm Z\left(\frac{s}{\sqrt{n-1}}\right)$$

3. Working with proportions

FORMULA 6.3

$$\text{c.i.} = P_s \pm Z\sqrt{\frac{P_u(1 - P_u)}{n}}$$

Constructing Confidence Intervals

We want to construct an interval working with a sample whereby the true population parameter likely lies…

Procedures:

1. Set what is called our “alpha level”.
2. Find the associated Z score of the normal distribution that corresponds to this alpha (working with our sampling distribution).
3. Substitute values into the appropriate formula for constructing confidence intervals.. Several formulas are possible..

48 HOURS ONLY! Get FULL ACCESS to globeandmail.com

Canada's unemployment rate rises to 7.1 per cent as election approaches

RACHELLE YOUNGLAI
The Globe and Mail

Published Friday, Oct. 09, 2015 8:34AM EDT
Last updated Friday, Oct. 09, 2015 5:24PM EDT

157 Comments | 741 | 574 | 95 | 72 | G+1 14 | Print / License | AA

A random sample of 100,000 Canadians (this is roughly the size of Canada's Labour Force Survey) estimates that 7.1% of Canada's population is unemployed. Create a 95% CI on this statistic:

$n=100{,}000$ $P_s = .071$

Recall with Confidence Intervals, I gave you three formulas that are appropriate when calculating:

1. Working with means: when we know our "population standard deviation"

FORMULA 6.1

$$\text{c.i.} = \overline{X} \pm Z\left(\frac{\sigma}{\sqrt{n}}\right)$$

where c.i. = confidence interval

$\overline{X}$ = the sample mean

Z = the Z score as determined by the alpha level

$\frac{\sigma}{\sqrt{n}}$ = the standard deviation of the sampling distribution or the standard error of the mean

2. Working with means: when we <u>do not know</u> our "population standard deviation" but do know our sample standard deviation

FORMULA 6.2

$$\text{c.i.} = \overline{X} \pm Z\left(\frac{s}{\sqrt{n-1}}\right)$$

3. Working with proportions

FORMULA 6.3

$$\text{c.i.} = P_s \pm Z\sqrt{\frac{P_u(1 - P_u)}{n}}$$

A random sample of 100,000 Canadians (this is roughly the size of Canada's Labour Force Survey) estimates that 7.1% of Canada's population is unemployed. Create a 95% CI on this statistic:

n=100,000 Ps = .071

$$\text{c.i.} = P_s \pm Z\sqrt{\frac{P_u(1 - P_u)}{n}}$$

1. Set alpha: .05, we are working with a 95% confidence interval

2. Set your appropriate Z score : Z = 1.96

A random sample of 100,000 Canadians (this is roughly the size of Canada's Labour Force Survey) estimates that 7.1% of Canada's population is unemployed. Create a 95% CI on this statistic:

N=100,000 Ps = .071

$$\text{c.i.} = P_s \pm Z\sqrt{\frac{P_u(1-P_u)}{n}}$$

1. Set alpha: .05, we are working with a 95% confidence interval

2. Set your appropriate Z score : Z = 1.96

3. Use appropriate formula:

C.I. = .071 +/- 1.96 $\sqrt{.5(1-.5)/100.000}$
C.I. = .071 +/- 1.96 $\sqrt{(.25)/100{,}000)}$
C.I. = .071 +/- 1.96 $\sqrt{.0000025}$
C.I. = .071 +/- 1.96 (.0015811)

C.I. = .071 +/- .003

CI: 95% of the time, we anticipate that Canada's unemployment rate falls somewhere between 7.07 and 7.13

Again: Reminder, that in addition to working with proportions (and %s)

$$\text{c.i.} = P_s \pm Z\sqrt{\frac{P_u(1 - P_u)}{n}}$$

Might be given a mean, and its sample standard deviation (with sample size)
-> calculate a CI

$$\text{c.i.} = \overline{X} \pm Z\left(\frac{s}{\sqrt{n-1}}\right)$$

Might be given a mean, with its population standard deviation (with sample size)
-> calculate a CI

$$\text{c.i.} = \overline{X} \pm Z\left(\frac{\sigma}{\sqrt{n}}\right)$$

Could be asked for a:

95% CI or a 90 % CI or a 99% CI

The only difference relates to the Z value used:

1.96 or 1.645 or 2.575

Chapter 7

Hypothesis Testing I:
The One-Sample Case

Idea:

Obtain a statistic for a specific sample,..

Does it differ significantly from a given population parameter?

Eg. We take a "sample of King's students"…
calculate their "mean GPA"

Does this sample statistic differ significantly from
all students in Ontario (i.e. the population parameter)?

That's a Hypothesis Test I:
The One-Sample Case

Hypothesis testing is designed to detect *significant differences*: differences that did *not* occur by random chance.

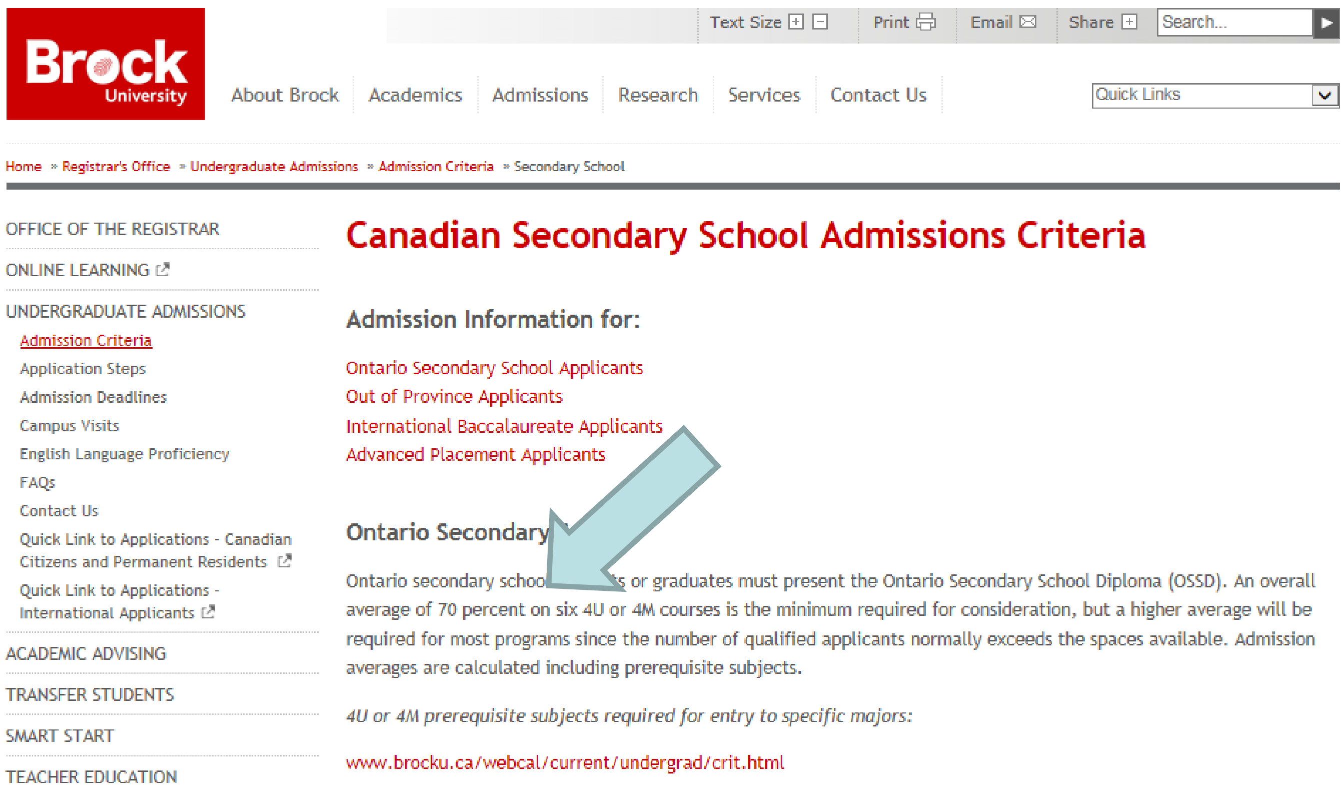
Is an observed difference "real", or is it merely
"sampling error" or "random noise" in our data?

In this presentation you will learn about:

- The basic logic of hypothesis testing
  - Hypothesis testing for single sample means
  - The Five-Step Model
- Other material covered:
  - One- vs. Two- tailed tests
  - Type I vs. Type II error
  - Student's *t* distribution
  - Hypothesis testing for single sample proportions

Hypothesis Testing – one sample test

- LET ME introduce it on a general conceptual level, with an example, and then afterwards approach it more systematically..



Text Size | Print | Email | Share | Search...

Brock University

About Brock | Academics | Admissions | Research | Services | Contact Us

Quick Links

Home » Registrar's Office » Undergraduate Admissions » Admission Criteria » Secondary School

- OFFICE OF THE REGISTRAR
- ONLINE LEARNING
- UNDERGRADUATE ADMISSIONS
  - Admission Criteria
  - Application Steps
  - Admission Deadlines
  - Campus Visits
  - English Language Proficiency
  - FAQs
  - Contact Us
  - Quick Link to Applications - Canadian Citizens and Permanent Residents
  - Quick Link to Applications - International Applicants
- ACADEMIC ADVISING
- TRANSFER STUDENTS
- SMART START
- TEACHER EDUCATION

Canadian Secondary School Admissions Criteria

Admission Information for:

Ontario Secondary School Applicants
Out of Province Applicants
International Baccalaureate Applicants
Advanced Placement Applicants

Ontario Secondary

Ontario secondary schoo ts or graduates must present the Ontario Secondary School Diploma (OSSD). An overall average of 70 percent on six 4U or 4M courses is the minimum required for consideration, but a higher average will be required for most programs since the number of qualified applicants normally exceeds the spaces available. Admission averages are calculated including prerequisite subjects.

4U or 4M prerequisite subjects required for entry to specific majors:

www.brocku.ca/webcal/current/undergrad/crit.html

Hypothesis Testing for Single Sample Means: An Example

- Brock University has been accused of inconsistent grading policies. Why? Grade inflation…
- The rumor is: their majors have much different GPAs than students in general – even though the work requirements & apparent motivation appear to be no different from other universities (if not weaker)
- We have data from Stats Can on all university students , i.e. the full population (but nothing specifically for Brock)
- We can only draw a “sample” of Brock students, and we want to do a “single sample test” , comparing the sample statistic with the population parameter.

Suppose we know from Stats Can:

- The value of the parameter, average GPA for *all* University students across Ontario, is 2.70 ($\mu = 2.70$), with a standard deviation of 0.70 ($\sigma = .70$).

Suppose we know from Stats Can:

- The value of the parameter, average GPA for *all* University students across Ontario, is 2.70 (μ = 2.70), with a standard deviation of 0.70 (σ = .70).

Then we take a random sample of 117 Brock majors, & we document a mean = 3.00

- There is a difference between the parameter (2.70) and the statistic (3.00)., but is it real???
  - Is it a “significant difference?”
  - The observed difference may have been caused by random chance.

- Formally, we can state the two **hypotheses** as:

Null Hypothesis (H_0)

"The difference is caused by random chance."

Note: The *Null Hypothesis* always states

there is "no significant difference."

OR

Alternative hypothesis (H_1)

"The difference is real".

Note: The *Alternate hypothesis*

always contradicts the $H_{0.}$

In other words:

H_0: The sample mean (3.00 with this specific sample) is not significantly different from the pop. mean (2.70).

- The difference is merely caused by random chance (sampling error)
- Note: more likely with small samples, right?

H_1: The difference is real (significant).

- Brock majors are different from all students.

We can test H_0 given our knowledge of the "sampling distribution" and Z scores

How do we do “significance tests”?

We always begin by assuming the H_0 is true (no real difference).

- & then ask, “What is the **probability** of getting a sample statistic if in fact H_0 is true?
- In other words, in this case:
- “What is the probability of this sample of Brock students having a mean of 3.00 if in fact all Brock majors in reality have a mean of 2.70” (i.e. no different from the mean of all Canadian students)?
- Think of our “sampling distribution”, i.e. if we were to repeatedly sample Brock students (again and again and again), what is the probability of a sample mean of 3 if in reality the mean for Brock students were 2.7.
- *We calculate a test statistic – called Z(obtained) in this context, which allows us to estimate this probability.*

- What is the test statistic?
- We work with the following *Z* score formula in our significance test in order to estimate the aforementioned probability.

$$Z(\text{obtained}) = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

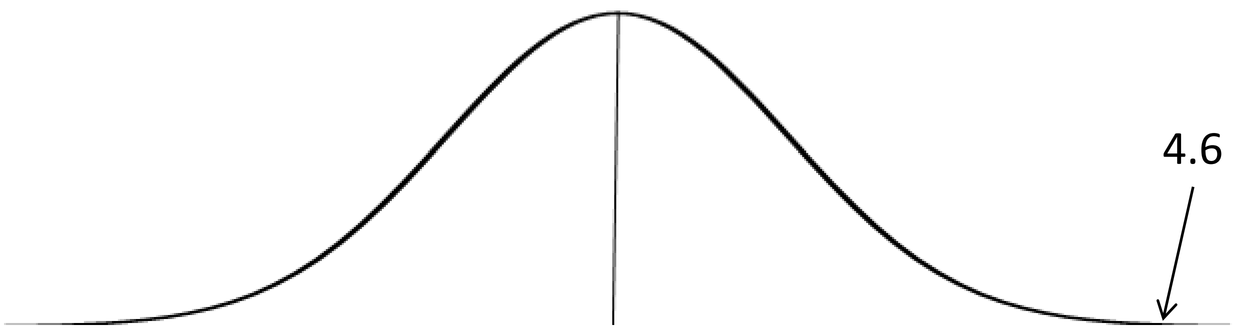
NOTE: this formula is equivalent to dividing the difference between the sample statistic and the population parameter by the standard error (the standard deviation of the sampling distribution)

We can then use Appendix A to determine the **probability** of getting this Z score in a normal distribution (in our sampling distribution).

With our example,
The sample mean = 3.00 (GPA for sample of Brock students)
With: μ = 2.70; σ=*0.70 (population mean & standard deviation)*

$$Z(\text{obtained}) = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}}$$

$$Z\ (\text{obtained}) = \frac{3.0 - 2.7}{.70 / \sqrt{117}} = 4.6$$

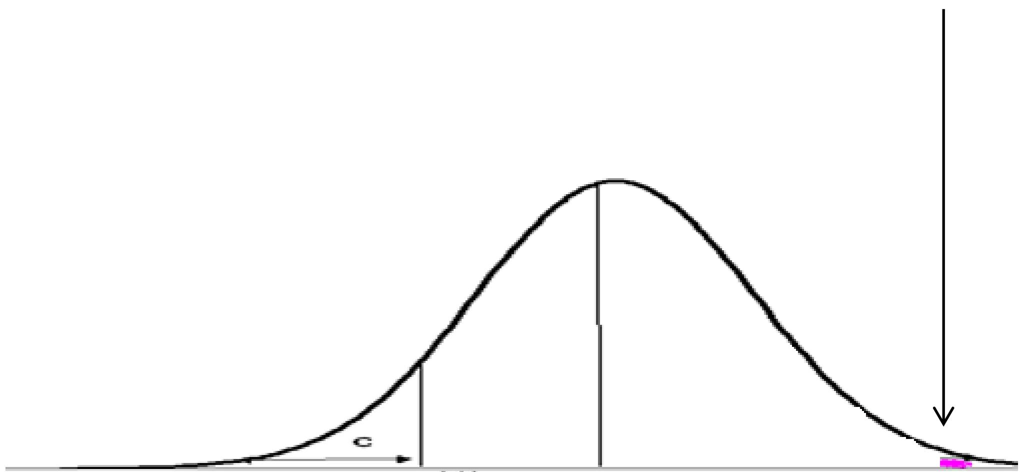


A mean of 3.0 is estimated to be fully 4.6 standard errors (Z scores) above the mean of our sampling distribution (assuming that the true population parameter is 2.7)
We can estimate the probability of scoring this high on this sampling distribution

Our sampling distribution

Z(obtained)
in significance tests

4.6



Column © gives us
The area here in this tail..
This gives us the probability of
getting a Z score of at least 4.6
on a normal distribution

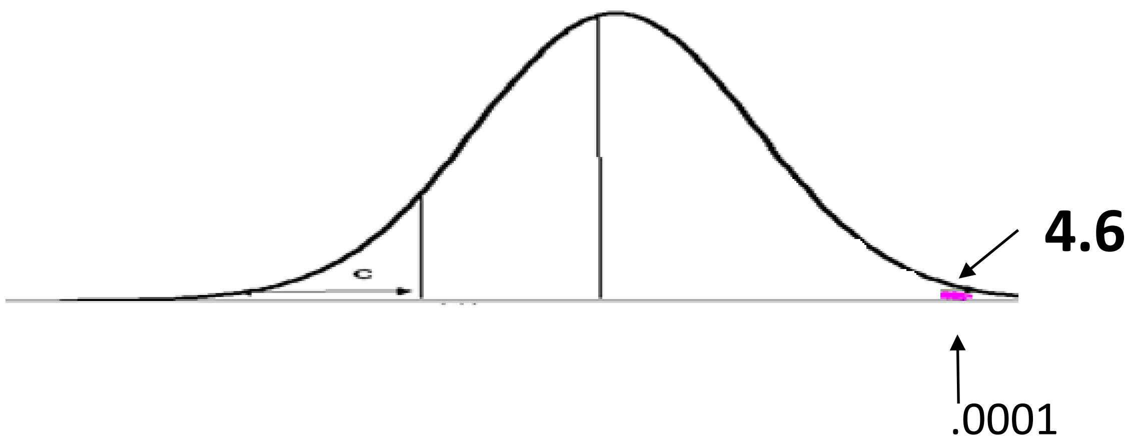
(a) z	(b) Area between Mean and Z	(c) Area beyond Z
2.61	0.4955	0.0045
2.62	0.4956	0.0044
2.63	0.4957	0.0043
2.64	0.4959	0.0041
2.65	0.4960	0.0040
2.66	0.4961	0.0039
2.67	0.4962	0.0038
2.68	0.4963	0.0037
2.69	0.4964	0.0036
2.70	0.4965	0.0035
2.71	0.4966	0.0034
2.72	0.4967	0.0033
2.73	0.4968	0.0032
2.74	0.4969	0.0031
2.75	0.4970	0.0030
2.76	0.4971	0.0029
2.77	0.4972	0.0028
2.78	0.4973	0.0027
2.79	0.4974	0.0026
2.80	0.4974	0.0026
2.81	0.4975	0.0025
2.82	0.4976	0.0024
2.83	0.4977	0.0023
2.84	0.4977	0.0023
2.85	0.4978	0.0022
2.86	0.4979	0.0021
2.87	0.4979	0.0021
2.88	0.4980	0.0020
2.89	0.4981	0.0019
2.90	0.4981	0.0019
2.91	0.4982	0.0018
2.92	0.4982	0.0018
2.93	0.4983	0.0017
2.94	0.4984	0.0016
2.95	0.4984	0.0016
2.96	0.4985	0.0015
2.97	0.4985	0.0015
2.98	0.4986	0.0014
2.99	0.4986	0.0014
3.00	0.4986	0.0014
3.01	0.4987	0.0013
3.02	0.4987	0.0013
3.03	0.4988	0.0012
3.04	0.4988	0.0012
3.05	0.4989	0.0011
3.06	0.4989	0.0011
3.07	0.4989	0.0011
3.08	0.4990	0.0010
3.09	0.4990	0.0010
3.10	0.4990	0.0010
3.11	0.4991	0.0009
3.12	0.4991	0.0009
3.13	0.4991	0.0009
3.14	0.4992	0.0008
3.15	0.4992	0.0008
3.16	0.4992	0.0008
3.17	0.4992	0.0008
3.18	0.4993	0.0007
3.19	0.4993	0.0007
3.20	0.4993	0.0007
3.21	0.4993	0.0007
3.22	0.4994	0.0006
3.23	0.4994	0.0006
3.24	0.4994	0.0006
3.25	0.4994	0.0006
3.26	0.4994	0.0006
3.27	0.4995	0.0005
3.28	0.4995	0.0005
3.29	0.4995	0.0005
3.30	0.4995	0.0005
3.31	0.4995	0.0005
3.32	0.4995	0.0005
3.33	0.4996	0.0004
3.34	0.4996	0.0004
3.35	0.4996	0.0004
3.36	0.4996	0.0004
3.37	0.4996	0.0004
3.38	0.4996	0.0004
3.39	0.4997	0.0003
3.40	0.4997	0.0003
3.41	0.4997	0.0003
3.42	0.4997	0.0003
3.43	0.4997	0.0003
3.44	0.4997	0.0003
3.45	0.4997	0.0003
3.46	0.4997	0.0003
3.47	0.4997	0.0003
3.48	0.4997	0.0003
3.49	0.4998	0.0002
3.50	0.4998	0.0002
3.60	0.4998	0.0002
3.70	0.4999	0.0001
3.80	0.4999	0.0001
3.90	0.499[illegible]	<0.0001
4.00	[illegible]	<0.0001

4.60 .4999 <.0001

The probability of getting
Such an outcome is
<.0001!

Our sampling distribution

Z(obtained)
in significance tests



The probability of this sort of outcome (a mean of 3.0 in a sample of Brock students) when in reality the mean is expected to be 2.7 is extremely slim… less than .0001 chance

So, logically, here, can we can be safe in “rejecting our null hypothesis”

Accept our H1 research hypothesis.. A significant difference appears to be documented..
Brock students are “significantly” different!!

Hypothesis Testing: The Five Step Model

- All the elements used in the example above can be formally organized into a five-step model:

1. Making assumptions and meeting test requirements.
2. Stating the null hypothesis.
3. Selecting the sampling distribution and establishing the critical region.
4. Computing the test statistic.
5. Making a decision and interpreting the results of the test.

Step 1: Make Assumptions and Meet Test Requirements

- Random sampling
  - Hypothesis testing assumes samples were selected with an equal probability of selection method: random

  The sample of 117 was randomly selected from all Brock majors.

- Level of Measurement is Interval-Ratio
  - GPA is Interval-Ratio so the mean is an appropriate statistic.

- Sampling Distribution is normal in shape
  - This assumption is satisfied by using a large enough sample (n>100).

Step 2: State the Null Hypothesis

H_0: $\mu = 2.7$

- The sample of 117 comes from a population that has a GPA of 2.7.
- The difference between 2.7 and 3.0 is trivial and caused by random chance.

- H_1: $\mu \neq 2.7$
  - The sample of 117 comes from a population that *does not* have a GPA of 2.7.
  - The difference between 2.7 and 3.0 reflects an actual difference between Brock majors and other students

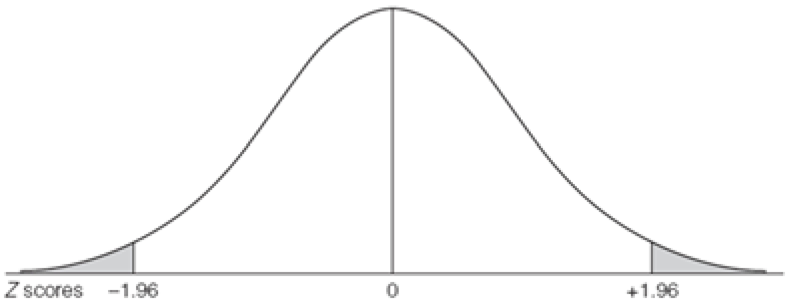
Step 3: Select Sampling Distribution and Establish the Critical Region

- Sampling Distribution= *Z*

Critical Region:

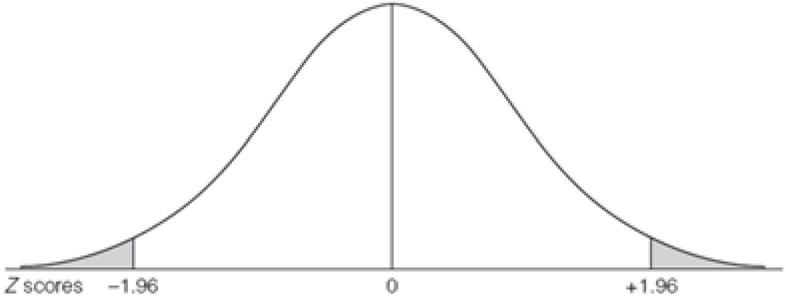
That segment of the "sampling distribution" whereby we consider a sample statistic to be "a rare event", hence if our test statistic falls into it, we "reject the null hypothesis"..

- **By convention, we typically set the .05 value as a guideline to identify differences that would be rare if H_0 is true.**
- In a normal distribution, we know that 5% of all samples in a sampling distribution fall outside of +/- 1.96 standard errors



- So we could set our "critical region" to include 5% of all sample outcomes.. outside of +/- 1.96 Z scores
- 5% of all samples fall within the critical region,..
- the remaining 95% are considered non-significant

Hypothesis Testing for Single Sample Means



Shaded tails here
Represent our critical region

Step 3: Select Sampling Distribution and Establish the Critical Region

- Sampling Distribution= Z
  - Alpha (α) = .05
  - α is an indication of a “rare” event

Step 3: Select Sampling Distribution and Establish the Critical Region

- Sampling Distribution= *Z*
  - Alpha (α) = .05
  - α is the indicator of “rare” events.

- Critical Region (C.R) begins at $\pm$ 1.96
  - This is the critical *Z* score associated with α = .05, two-tailed test.
  - If the obtained *Z* score falls in the C.R., reject the H_0.

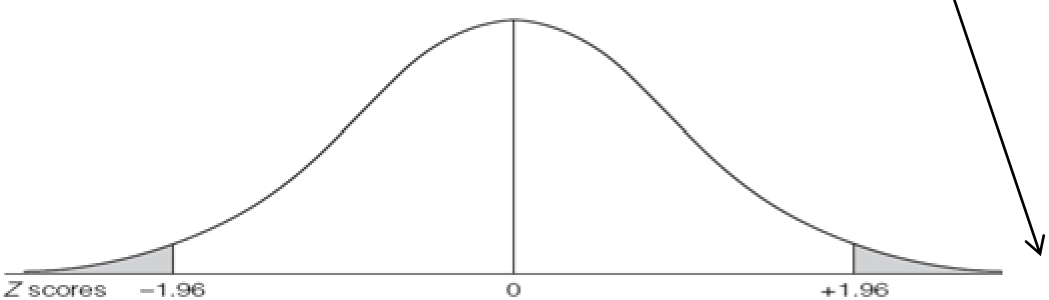
Step 4: Compute the Test Statistic & draw the Z distribution

With our example,
The sample mean = 3.00 (GPA for sample of Brock students)
With: μ = 2.70; σ*=0.70 (population mean & standard deviation)*

$$Z(\text{obtained}) = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

- Z (obtained) = $\dfrac{3.0 - 2.7}{.70/\sqrt{117}}$ = 4.6

Draw it:



Z (critical) Z (critical)

Step 5: Make Decision and Interpret Results

MAKING A DECISION IN STEP 5 AND INTERPRETING THE RESULTS OF THE TEST

Situation	Decision	Interpretation
The test statistic is *in* the critical region	Reject the null hypothesis (H_0)	The difference is statistically significant

OR:

Situation	Decision	Interpretation
The test statistic is *not* the critical region	Fail to reject the null hypothesis (H_0)	The difference is not in statistically significant

Step 5: Make Decision and Interpret Results *(continued)*

- The obtained Z score fell in the C.R., so we *reject* the H_0.
  - If the H_0 were true, a sample outcome of 3.00 would be highly unlikely.
  - Therefore, the H_0 is false and must be rejected.
- Brock majors have a GPA that is significantly different from the general student body.

Hypothesis Testing for Single Sample Means: Example with one tailed test

- The rumor is: Montreal Habs fans have a different level of intelligence than other Canadians?
- Sample: n=101 Habs Fans (SRS)
- Mean IQ = 95
- We know: Population mean is 100, with a standard deviation of 15 on this standard IQ test.
- Are they “significantly” different?

Step 1: Make Assumptions and Meet Test Requirements

- Random sampling
  - The sample of 101 Habs fans was randomly selected from all Habs fans.
- Level of Measurement is Interval-Ratio
  - IQ test score is Interval-Ratio so the mean is an appropriate statistic.
- Sampling Distribution is normal in shape
  - This assumption is satisfied by using a large enough sample (n>100).

Step 2: State the Null Hypothesis

H_0: $\mu = 100$

- The sample of 101 comes from a population that has an IQ of 100 (i.e. all Habs fans).
- The difference between 95 and 100 is trivial and caused by random chance.

- H_1: $\mu \neq 100$
  - The sample of 101 comes from a population that *has* an IQ different than 100.
  - The difference between 95 and 100 is real!

Step 3: Select Sampling Distribution and Establish the Critical Region

- Sampling Distribution= Z
  - Alpha (α) = .05
  - α is the indicator of “rare” events.
  - Any difference with a probability less than α is rare and will cause us to reject the H_0.
- Critical Region (C.R) begins at +/- 1.96
- This is the critical Z score associated with α = .05, two-tailed test.
  - If the obtained Z score falls in the C.R., reject the H_0.

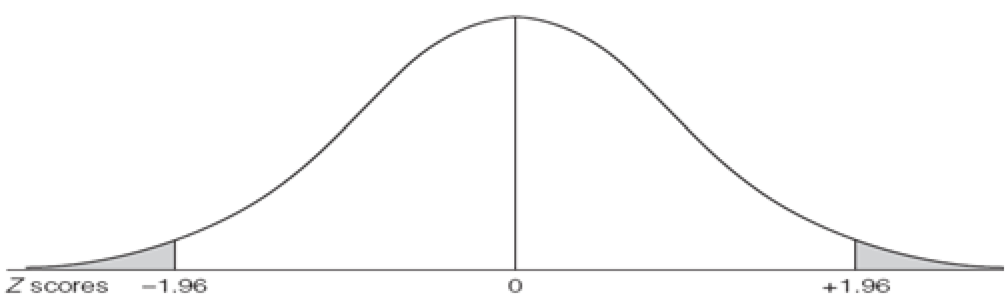
Step 4: Compute the Test Statistic & draw

So, substituting the values into Formula 7.1,
With: μ = 100; *σ=15 and N=101*

$$Z(\text{obtained}) = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

- $$Z\ (\text{obtained}) = \frac{95-100}{15/\sqrt{101}} = -3.35$$

Draw it:



-3.35

Z (critical)

Z (critical)

Step 5: Make Decision and Interpret Results

MAKING A DECISION IN STEP 5 AND INTERPRETING THE RESULTS OF THE TEST

Situation	Decision	Interpretation
The test statistic is *in* the critical region	Reject the null hypothesis (H_0)	The difference is statistically significant

OR:

Situation	Decision	Interpretation
The test statistic is *not* the critical region	Fail to reject the null hypothesis (H_0)	The difference is not in statistically significant

Step 5: Make Decision and Interpret Results *(continued)*

- The obtained *Z* score fell in the C.R., so we *reject* the H_0.
  - If the H_0 were true, a sample outcome of 95 would be highly unlikely.
  - Therefore, the H_0 is false and must be rejected.
- Habs fans are significantly different than other Canadians

Crucial Choices in the Five Step Model

- Model is fairly rigid, but there are two crucial choices:

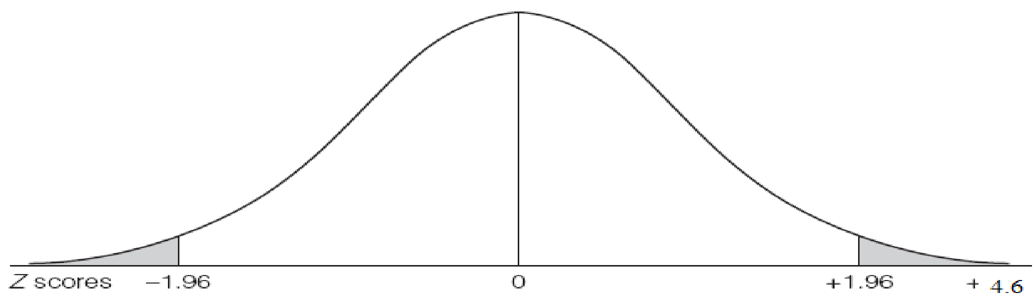
1. One-tailed or two-tailed test (Section 7.4)
2. Alpha (α) level (Section 7.5)

One- and Two-Tailed Hypothesis Tests

- Two-tailed: States that population mean is “not equal” to value stated in null hypothesis.

Example:

H_1: μ≠2.7, where ≠ means “not equal to”. Note: the GPA example illustrated above was a two-tailed test, with two critical regions



Alternatively: one-tailed tests are possible: Differences in a specific direction.

Example:

H_1: μ>2.7, where > signifies “greater than”

The Curve for Two- vs. One-tailed Tests *at α = .05*:

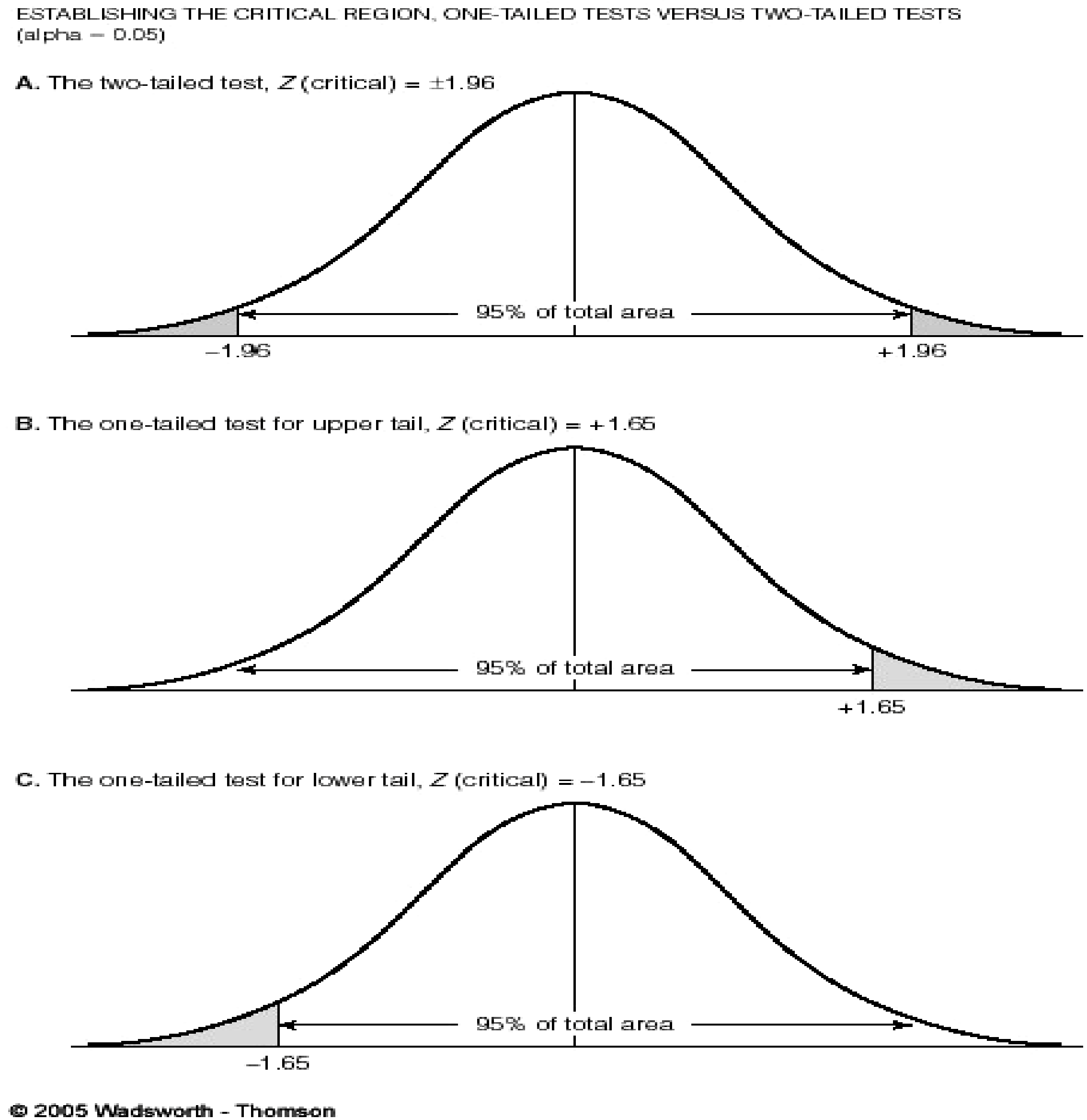
Two-tailed test:

"is there a **significant difference**?"

One-tailed tests:

"is the sample mean **greater** than μ?"

"is the sample mean **less** than μ?"



- Tails affect Critical Region in Step 3:

TABLE 8.1 ONE- VS. TWO-TAILED TESTS, $\alpha = .05$

If the Research Hypothesis Uses	The Test Is	And Concern Is with	Z(critical) =
$\neq$	Two-tailed	Both tails	± 1.96
$>$	One-tailed	Upper tail	$+1.65$
$<$	One-tailed	Lower tail	-1.65

Hypothesis Testing for Single Sample Means: Example with one tailed test

- Previous research has established that immigrants earn less than other Canadians:

- Sample: n=101 recent immigrants
- Mean employment earnings = $36,000
- We know: Population mean is $40,000, with a standard deviation of $15,000
- Are they earning “significantly” less?

Step 1: Make Assumptions and Meet Test Requirements

- Random sampling
  - The sample of 101 immigrants was randomly selected from all immigrants.

- Level of Measurement is Interval-Ratio
- Earnings is Interval-Ratio so the mean is an appropriate statistic.

- Sampling Distribution is normal in shape
  - This assumption is satisfied by using a large enough sample (n>100).

Step 2: State the Null Hypothesis

H_0: μ = \$40,000
  - The sample of 101 immigrants comes from a population that has a mean income of \$40,000 (i.e. all Canadians).
  - The difference between \$36,000 and \$40,000 is caused by random chance.

- H_1: $\mu < \$40,000$
  - The sample of 101 comes from a population that *has* a mean income less than \$40,000.
  - The difference between \$36,000 and \$40,000 reflects lower earnings among new Canadians relative to all Canadians.

Step 3: Select Sampling Distribution and Establish the Critical Region

- Sampling Distribution= *Z*
  - Alpha (α) = .05
  - α is the indicator of "rare" events.
  - Any difference with a probability less than α is rare and will cause us to reject the H_0.
- Critical Region (C.R) begins at - 1.65
  - This is the critical *Z* score associated with α = .05, one-tailed test.
  - If the obtained *Z* score falls in the C.R., reject the H_0.

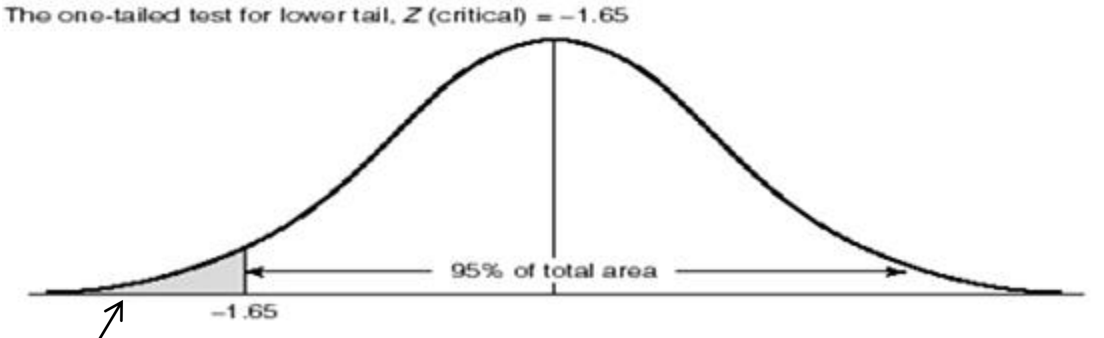
Step 4: Compute the Test Statistic

So, substituting the values into Formula 7.1,
With: μ = 40,000; *σ=15,000 and N=101*

$$Z(\text{obtained}) = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

- Z (obtained) = $$\frac{36000-40000}{15000/\sqrt{101}}$$ = -2.68

Draw it:



-2.68

Z (critical)

Step 5: Make Decision and Interpret Results

MAKING A DECISION IN STEP 5 AND INTERPRETING THE RESULTS OF THE TEST

Situation	Decision	Interpretation
The test statistic is *in* the critical region	Reject the null hypothesis (H_0)	The difference is statistically significant

OR:

Situation	Decision	Interpretation
The test statistic is *not* the critical region	Fail to reject the null hypothesis (H_0)	The difference is not in statistically significant

Step 5: Make Decision and Interpret Results *(continued)*

- The obtained Z score fell in the C.R., so we *reject* the H_0.
  - If the H_0 were true, a sample outcome of $36,000 would be highly unlikely.
  - Therefore, the H_0 is false and must be rejected.
- New Canadians are earning less than other Canadians

Note: In previous examples, I have used the formula:

$$Z(\text{obtained}) = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

What if our “population standard deviation is unknown?

- What if our population standard deviation is unknown?
- How can we test a hypothesis if σ is not known, as is usually the case?
  - –For large samples (N>100), we use *s* as an estimator of σ and use standard normal distribution (*Z* scores), suitably corrected for the bias (*n* is replaced by *n-1* to correct for the fact that *s* is a biased estimator of σ.

$$Z(\text{obtained}) = \frac{\overline{X} - \overline{\mu}}{s/\sqrt{n-1}}$$

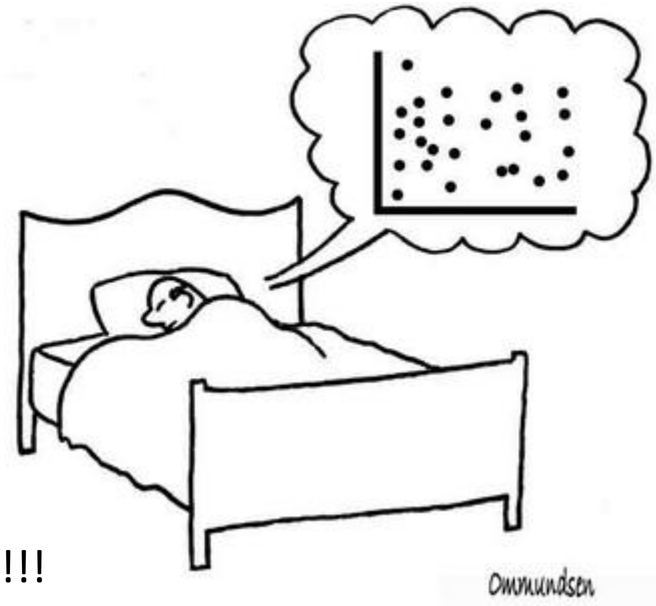
What if our sample size is small? Less than 100?

The statistician's nightmare..

Really small samples…

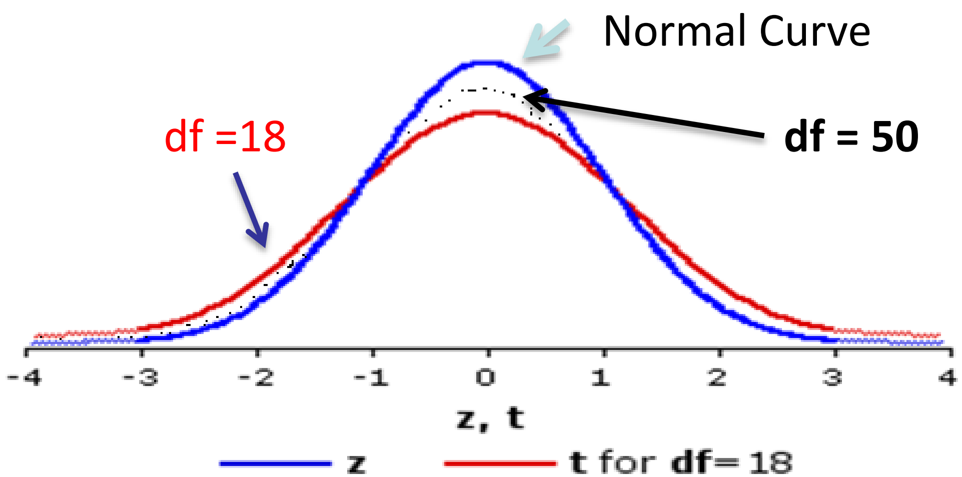
Much more sampling error!!!

Harder to document significant differences!!!



- When the sample size is small (< 100) then the **<u>Student's *t* distribution</u>** should be used

- The test statistic is known as “*t*”.
- The curve of the *t* distribution is flatter than that of the Z distribution
- but as the sample size increases, the *t*-curve starts to resemble the normal curve



The Student's *t* distribution

- A similar formula as for *Z* (obtained) is used in hypothesis testing:

$$t(\text{obtained}) = \frac{\overline{X} - \mu}{s/\sqrt{n-1}}$$

- The logic of the five-step model for hypothesis testing is followed.
- However in testing hypothesis we use the *t* table (Appendix B), not the *Z* table (Appendix A).

- NOTE:
- The *t* table differs from the *Z* table in the following ways:

Degrees of Freedom (df)	Level of Significance for One-tailed Test					
	.10	.05	.025	.01	.005	.0005
	Level of Significance for Two-tailed Test					
	.20	.10	.05	.02	.01	.001
1	3.078	6.314	12.706	31.821	63.657	636.619
2	1.886	2.920	4.303	6.965	9.925	31.598
3	1.638	2.353	3.182	4.541	5.841	12.941
4	1.533	2.132	2.776	3.747	4.604	8.610
5	1.476	2.015	2.571	3.365	4.032	6.859
6	1.440	1.943	2.447	3.143	3.707	5.959
7	1.415	1.895	2.365	2.998	3.499	5.405
8	1.397	1.860	2.306	2.896	3.355	5.041
9	1.383	1.833	2.262	2.821	3.250	4.781
10	1.372	1.812	2.228	2.764	3.169	4.587
11	1.363	1.796	2.201	2.718	3.106	4.437
12	1.356	1.782	2.179	2.681	3.055	4.318
13	1.350	1.771	2.160	2.650	3.012	4.221
14	1.345	1.761	2.145	2.624	2.977	4.140
15	1.341	1.753	2.131	2.602	2.947	4.073
16	1.337	1.746	[illegible]	[illegible]	[illegible]	[illegible]

1. Column at left for degrees of freedom (df) (df = n – 1)
   i.e. the smaller the sample, the flatter the distribution
1. Alpha levels along top two rows: one- and two-tailed
2. Entries in table are actual scores: t (critical)
   i.e. they mark beginning of critical region, not areas under the curve

Degrees of Freedom (df)	Level of Significance for One-tailed Test					
	.10	.05	.025	.01	.005	.0005
	Level of Significance for Two-tailed Test					
	.20	.10	.05	.02	.01	.001
1	3.078	6.314	12.706	31.821	63.657	636.619
2	1.886	2.920	4.303	6.965	9.925	31.598
3	1.638	2.353	3.182	4.541	5.841	12.941
4	1.533	2.132	2.776	3.747	4.604	8.610
5	1.476	2.015	2.571	3.365	4.032	6.859
6	1.440	1.943	2.447	3.143	3.707	5.959
7	1.415	1.895	2.365	2.998	3.499	5.405
8	1.397	1.860	2.306	2.896	3.355	5.041
9	1.383	1.833	2.262	2.821	3.250	4.781
10	1.372	1.812	2.228	2.764	3.169	4.587
11	1.363	1.796	2.201	2.718	3.106	4.437
12	1.356	1.782	2.179	2.681	3.055	4.318
13	1.350	1.771	2.160	2.650	3.012	4.221
14	1.345	1.761	2.145	2.624	2.977	4.140
15	1.341	1.753	2.131	2.602	2.947	4.073
16	1.337	1.746	[illegible]	[illegible]	[illegible]	[illegible]

Example

- A random sample of 30 students at Kings reported drinking on average 4 bottles of beer per week, with a standard deviation of 5.
- Is this significantly different from the population average ($\mu = 6$ bottles)?

Solution (using five step model)

- Step 1: Make Assumptions and Meet Test Requirements:

- Random sample
- Level of measurement is interval-ratio
- The sample is small (<100), so we can't work with a normal distribution

Solution (cont.)

Step 2: State the null and alternate hypotheses.

H_0: $\mu = 6$ Kings students are no different from the population overall

H_1: $\mu \neq 6$ Kings students differ from other Canadians in their beer consumption

Solution (cont.)

- Step 3: Select Sampling Distribution and Establish the Critical Region

1. Small sample, I-R level, so use *t* distribution.
2. Alpha (α) = .05
3. Degrees of Freedom = N-1 = 30-1 = 29
4. Critical t = ?

Degrees of Freedom (df)	Level of Significance for One-tailed Test					
	.10	.05	.025	.01	.005	.0005
	Level of Significance for Two-tailed Test					
	.20		.05	.02	.01	.001
1	3.078	6.314	12.706	31.821	63.657	636.619
2	1.886	2.920	4.303	6.965	9.925	31.598
3	1.638	2.353	3.182	4.541	5.841	12.941
4	1.533	2.132	2.776	3.747	4.604	8.610
5	1.476	2.015	2.571	3.365	4.032	6.859
6	1.440	1.943	2.447	3.143	3.707	5.959
7	1.415	1.895	2.365	2.998	3.499	5.405
8	1.397	1.860	2.306	2.896	3.355	5.041
9	1.383	1.833	2.262	2.821	3.250	4.781
10	1.372	1.812	2.228	2.764	3.169	4.587
11	1.363	1.796	2.201	2.718	3.106	4.437
12	1.356	1.782	2.179	2.681	3.055	4.318
13	1.350	1.771	2.160	2.650	3.012	4.221
14	1.345	1.761	2.145	2.624	2.977	4.140
15	1.341	1.753	2.131	2.602	2.947	4.073
16	1.337	1.746	2.120	2.583	2.921	4.015
17	1.333	1.740	2.110	2.567	2.898	3.965
18	1.330	1.734	2.101	2.552	2.878	3.922
19	1.328	1.729	2.093	2.539	2.861	3.883
20	1.325	1.725	2.086	2.528	2.845	3.850
21	1.323	1.721	2.080	2.518	2.831	3.819
22	1.321	1.717	2.074	2.508	2.819	3.792
23	1.319	1.714	2.069	2.500	2.807	3.767
24	1.318	1.711	2.064	2.492	2.797	3.745
25	1.316	1.708	2.060	2.485	2.787	3.725
26	1.315	1.706	2.056	2.479	2.779	3.707
27	1.314	1.703	2.052	2.473	2.771	3.690
28	1.313	1.701	2.048	2.467	2.763	3.674
29	1.311	[illegible]	2.045	2.462	2.756	3.659
30	1.310	[illegible]	2.042	2.457	2.750	3.646

Solution (cont.)

- Step 3: Select Sampling Distribution and Establish the Critical Region

1. Small sample, I-R level, so use *t* distribution.
2. Alpha (α) = .05
3. Degrees of Freedom = N-1 = 30-1 = 29
4. t (critical)= ?

   = 2.045

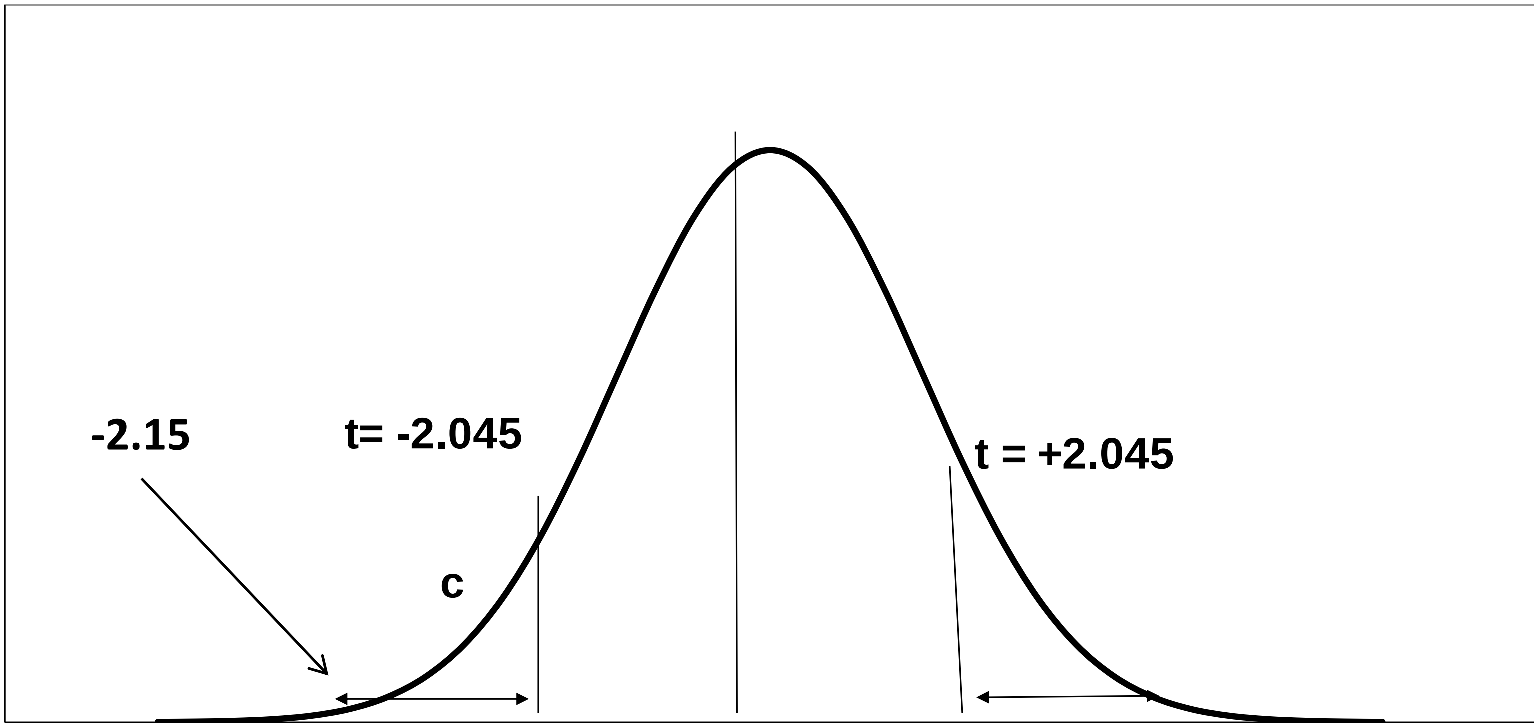
Solution (cont.)

- Step 4: Use Formula to Compute the Test Statistic

t (obtained) $$= \frac{\overline{\mathrm{X}} - \mu}{S \big/ \sqrt{n-1}} = \frac{4-6}{5 \big/ \sqrt{30-1}} = -2.15$$

Looking at the curve for the t distribution

Alpha (a) = .05



Step 5 Make a Decision and Interpret Results

- The obtained t score fell in the Critical Region, so we *reject* the H_0 (t (obtained) > t (critical))
  - If the H_0 were true, a sample outcome of 4 would be unlikely.
  - Therefore, we consider H_0 to be false and must be rejected.

- Kings students likely have drinking habits (beer consumption) that are significant different from other Canadians ($t = -2.15$, $df = 29$, $\alpha = .05$).

Testing Sample Proportions (n>100):

Formula 7.3

$$Z(\text{obtained}) = \frac{P_s - P_u}{\sqrt{P_u(1 - P_u)/n}}$$

Beyond this, if working with a large sample, same essential 5 steps as with "testing sample means"…

Testing Sample Proportions:

- Example: 48% of a sample of 250 Kings students report being Catholic. Data is available from the Census indicate that 43% of all residents of Ontario report being Catholic. Are Kings students significantly more likely to be Catholic than other Ontario residents?

- If the data are in % format, convert to a proportion first (48% -> .48)
- The method is virtually identical as the one sample Z-test for means, except we work with proportions and a slightly different formula for Z (obtained) -> Follow the 5 step model.

Solution (using five step model)

- Step 1: Make Assumptions and Meet Test Requirements:

- Random sample
- Level of measurement is nominal -> use proportions
- The sample is large (>100), so we can work with a normal distribution as our sampling distribution

Solution (cont.)

Step 2: State the null and alternate hypotheses.

H_0: $P_\mu = .43$ The full population of Kings students are no different from the Ontario population overall

H_1: $P_\mu > .43$ Kings students are more likely to be Catholic than other Ontario residents

Solution (cont.)

- Step 3: Select Sampling Distribution and Establish the Critical Region

1. Sampling distribution = Z distribution
2. Alpha (α) = .05
3. Critical Z = ?

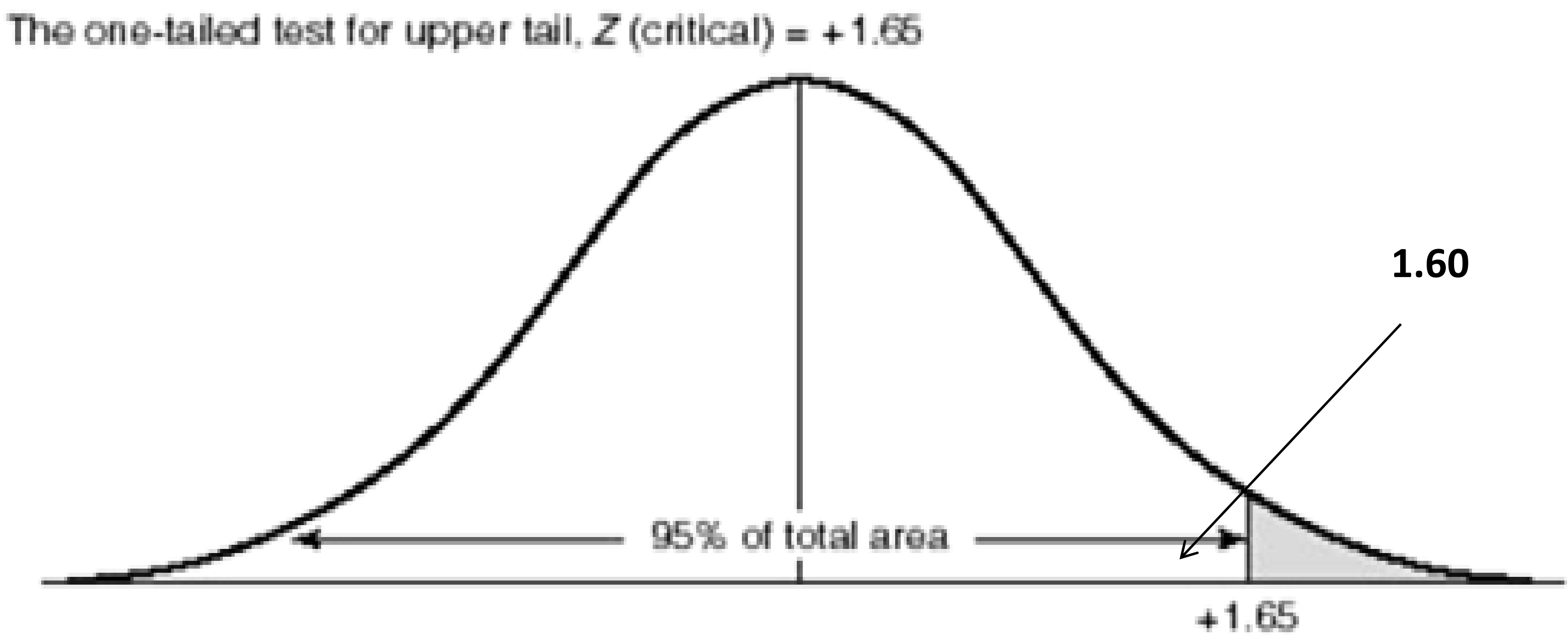
   = +1.65 (why? one tailed test)

Solution (cont.)

- Step 4: Use Formula to Compute the Test Statistic

Z (obtained) $= \dfrac{Ps - P\mu}{\sqrt{P\mu(1-P\mu)/n}} = \dfrac{.48 - .43}{\sqrt{.43(1-.43)/250}} = 1.60$

Looking at the curve for the Z distribution
Alpha (a) = .05



Step 5 Make a Decision and Interpret Results

- The obtained Z score did not fall in the Critical Region, so we fail to *reject* the H_0

- Kings students do not appear to be significantly more likely to be Catholic than the population over all ($Z = 1.60$, $\alpha = .05$, $n=250$).

- NOTE: What if we had the exact same data, but this time with N=1000, rather than 250?

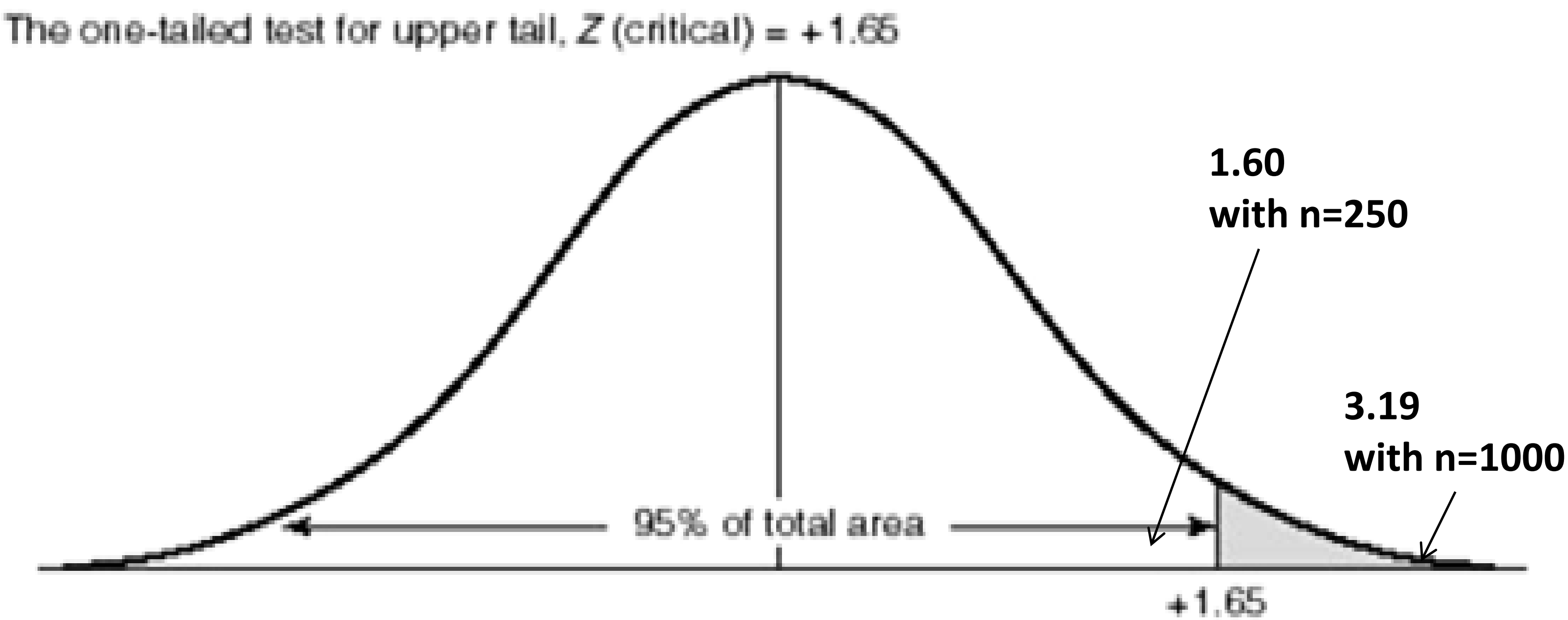
- With original sample (N=250)

Z (obtained) $$= \frac{Ps - P\mu}{\sqrt{P\mu(1 - P\mu)/n}} = \frac{.48 - .43}{\sqrt{.43(1 - .43)/250}} = 1.60$$

- With a larger sample (N=1000)

Z (obtained) $$= \frac{Ps - P\mu}{\sqrt{P\mu(1 - P\mu)/n}} = \frac{.48 - .43}{\sqrt{.43(1 - .43)/1000}} = 3.19$$

Looking at the curve for the t distribution
Alpha (a) = .05



WITH A MUCH LARGER SAMPLE, THE SAME PROPORTIONAL DIFFERENCE CAN BECOME SIGNIFICANT!!! (WE KNOW FOR SURE IT IS A REAL DIFFERENCE)

Alpha levels affect Critical Region in Step 3:

If Alpha Equals	The Two-Tailed Critical Region Will Begin at Z(critical) Equal to
0.10	± 1.65
0.05	± 1.96
0.01	± 2.58
0.001	± 3.29

Alpha Levels

- By assigning an alpha level, *α,* one defines an “unlikely” sample outcome.
- Alpha level is the probability that the decision to reject the null hypothesis, H_0, is incorrect.
  - Incorrectly rejecting a true null hypothesis: **Type I** or **alpha** error

There is a very small chance that I just committed a **Type I** error,.. i.e. incorrectly concluded that recent immigrants are earning “significantly” less than other Canadians when in reality they aren’t!!

Alpha Levels *(continued)*

- Type I, or alpha error:
  - Rejecting a true null hypothesis
- Type II, or beta error:
  - Failing to reject a false null hypothesis

    (i.e. in reality there is a significant difference, but I don't conclude this to be the case)

- There is a relationship between decision making and error:

TABLE 8.3 DECISION MAKING AND THE NULL HYPOTHESIS

The H_0 Is Actually:	Decision: Reject	Decision: Fail to Reject
True	Type I or α error	OK
False	OK	Type II or β error

Relative to alpha = .05:
- an alpha = .10 leads to a greater chance of a Type 1 error
- an alpha = .01 leads to a greater chance of a Type II error

Setting our alpha at .05 is usually considered a relatively good compromise in light of the possibility of both types of errors

Is there a “significant” difference?