Econometrics

Course Calendar	
Week	Main Content
Week 1	Introduction to Simple Regression
Week 2	Simple Regression
Week 3	Simple Regression: r^2 & Hands-on-Exercise
Week 4	Central Limit Theorem, Probability and Probability Density Function (PDF)
Week 5	Hypothesis Testing: Basics
Week 6	Simple Regression: Testing of Hypothesis

Econometrics
Lecture 5: Hypothesis Testing

GEETHA RANI PRAKASAM, PH.D.

Recap

- Central Limit Theorem
- Probability
- Probability distribution
- Probability density function

Outline

- Some important probability distribution
- Normal distribution
- Z distribution
- Student 't' distribution
- Chi square distribution
- Hypothesis Testing Methodology
- Z Test for the Mean (σ Known)
- p-Value Approach to Hypothesis Testing
- Connection to Confidence Interval Estimation

Some Important Theoretical Probability Distributions

- Many probability distributions that are important in theory or applications have been given specific names. They are divided in 2 main groups: discrete distributions and continuous distributions.
- **Discrete distribution**
- **Bernoulli distribution**
- In probability theory and statistics, the Bernoulli distribution, named after Swiss mathematician Jacob Bernoulli, is the discrete probability distribution of a random variable which takes the value 1 with probability p and the value 0 with probability $q = 1 - p$.

Discrete Distribution: Bernoulli

- That is, the probability distribution of any single experiment that asks a yes–no question; the question results in a boolean-valued outcome, a single bit whose value is success with probability *p* and failure with probability *q*.
- It can be used to represent a (possibly biased) coin toss where 1 and 0 would represent "heads" and "tails" (or vice versa), respectively, and *p* would be the probability of the coin landing on heads or tails, respectively.

Discrete Distribution: Binomial

- In probability theory and statistics, the binomial distribution with parameters *n* and *p* is the discrete probability distribution of the number of successes in a sequence of *n* independent experiments, each asking a yes–no question, and each with its own boolean-valued outcome: success (with probability *p*) or failure (with probability $q = 1 - p$).
- A single success/failure experiment is also called a Bernoulli trial or Bernoulli experiment and a sequence of outcomes is called a Bernoulli process; for a single trial, i.e., $n = 1$, the binomial distribution is a Bernoulli distribution.
- The binomial distribution is the basis for the popular binomial test of statistical significance.

Continuous Distribution

- **Normal distribution**
- In probability theory, the normal (or Gaussian or Gauss or Laplace–Gauss) distribution is a very common continuous probability distribution.
- Normal distributions are important in statistics and are often used in the natural and social sciences to represent real-valued random variables whose distributions are not known.
- A random variable with a Gaussian distribution is said to be normally distributed and is called a normal deviate.

Normal distribution

- The normal distribution is useful because of the central limit theorem. In its most general form, under some conditions (which include finite variance), it states that averages of samples of observations of random variables independently drawn from the same distribution converge in distribution to the normal, that is, they become normally distributed when the number of observations is sufficiently large.

Normal distribution

- The normal distribution is the most important distribution in statistics, since it arises naturally in numerous applications. The key reason is that large sums of (small) random variables often turn out to be normally distributed;
- A random variable X is said to have the normal distribution with parameters μ and σ if its density function is given by:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}$$

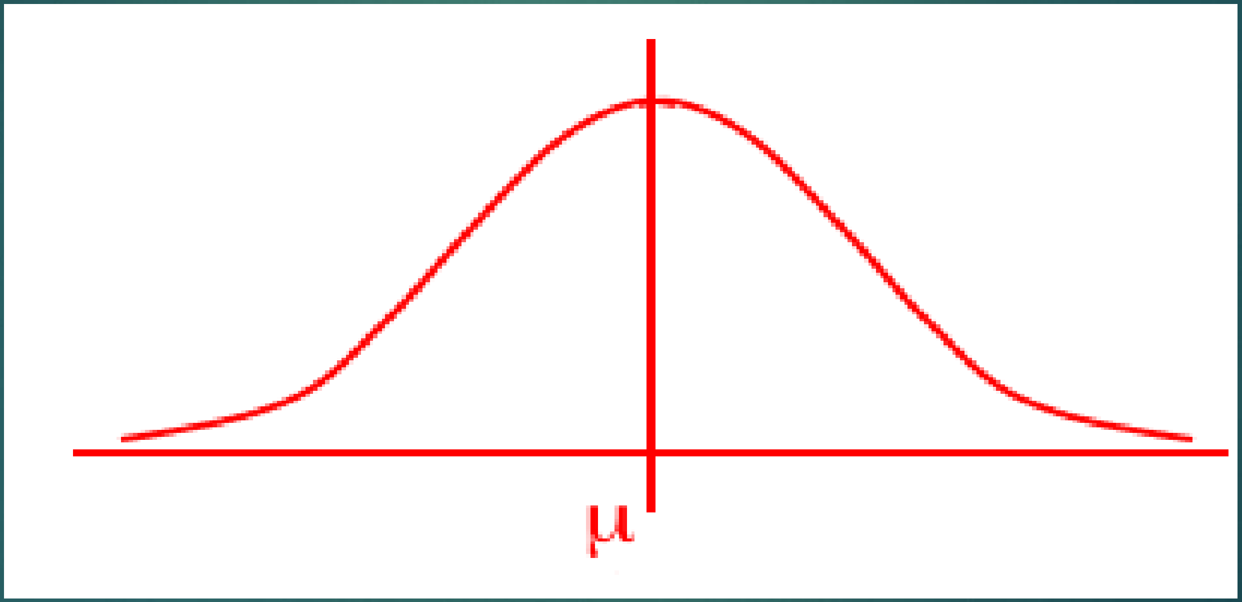
for $-\infty < x < \infty$.

Normal distribution

- It can be shown that
  $E(X) = \mu$ and $V(X) = \sigma 2$.
- Thus, the normal distribution is characterized by a mean μ and a standard deviation σ .
- A typical normal density curve looks like this:

Thus, the curve is bell shaped and is symmetric around the mean μ.

The standard deviation σ controls the "flatness" of the curve.



https://michaelminn.net/tutorials/normal-curve-grading/index.html

Normal distribution

- A normal distribution whose mean is 0 and standard deviation is 1 is called the standard normal distribution.
- In this case, the density function assumes the simpler form:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

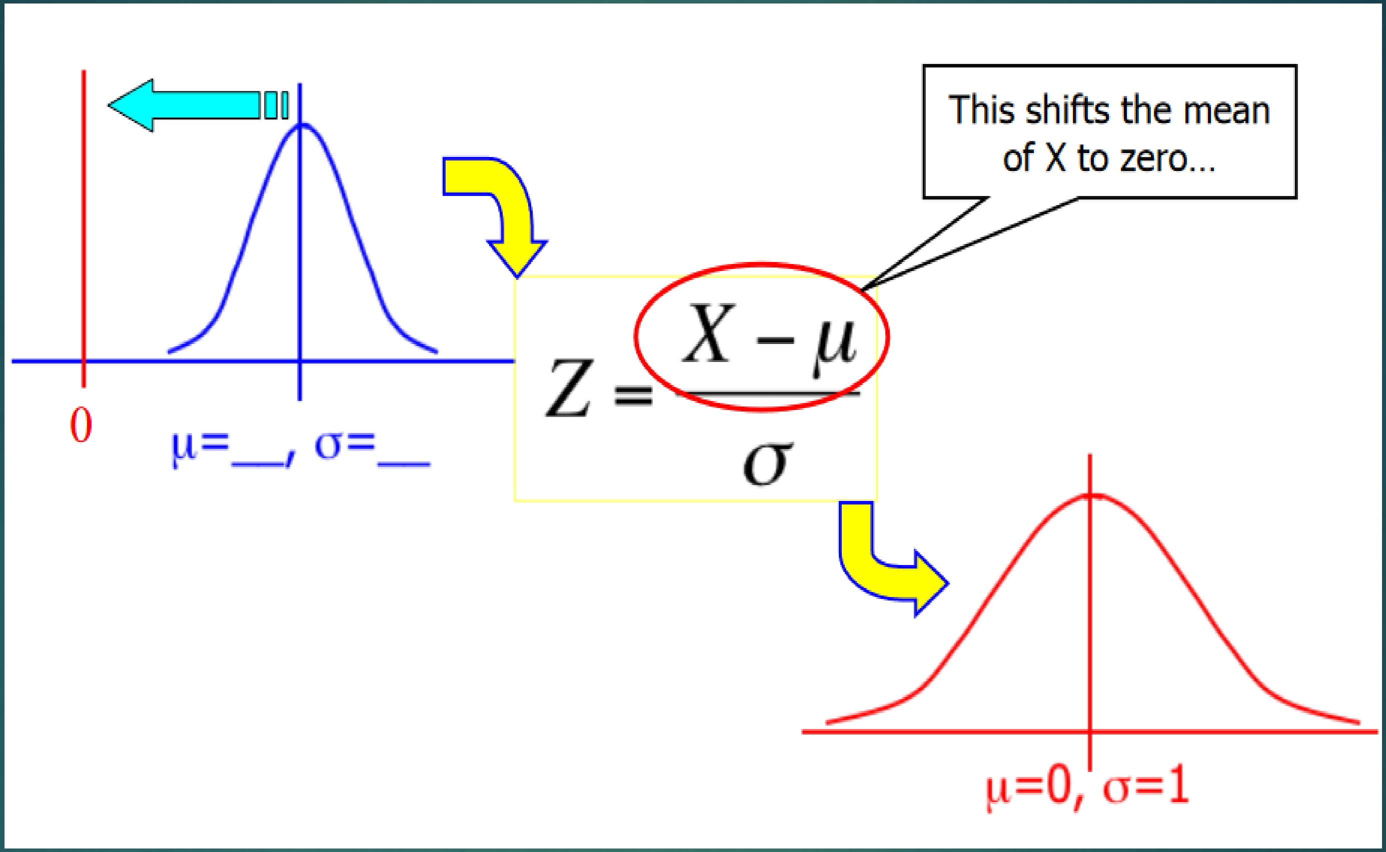
for $-\infty < x < \infty$.

Normal distribution

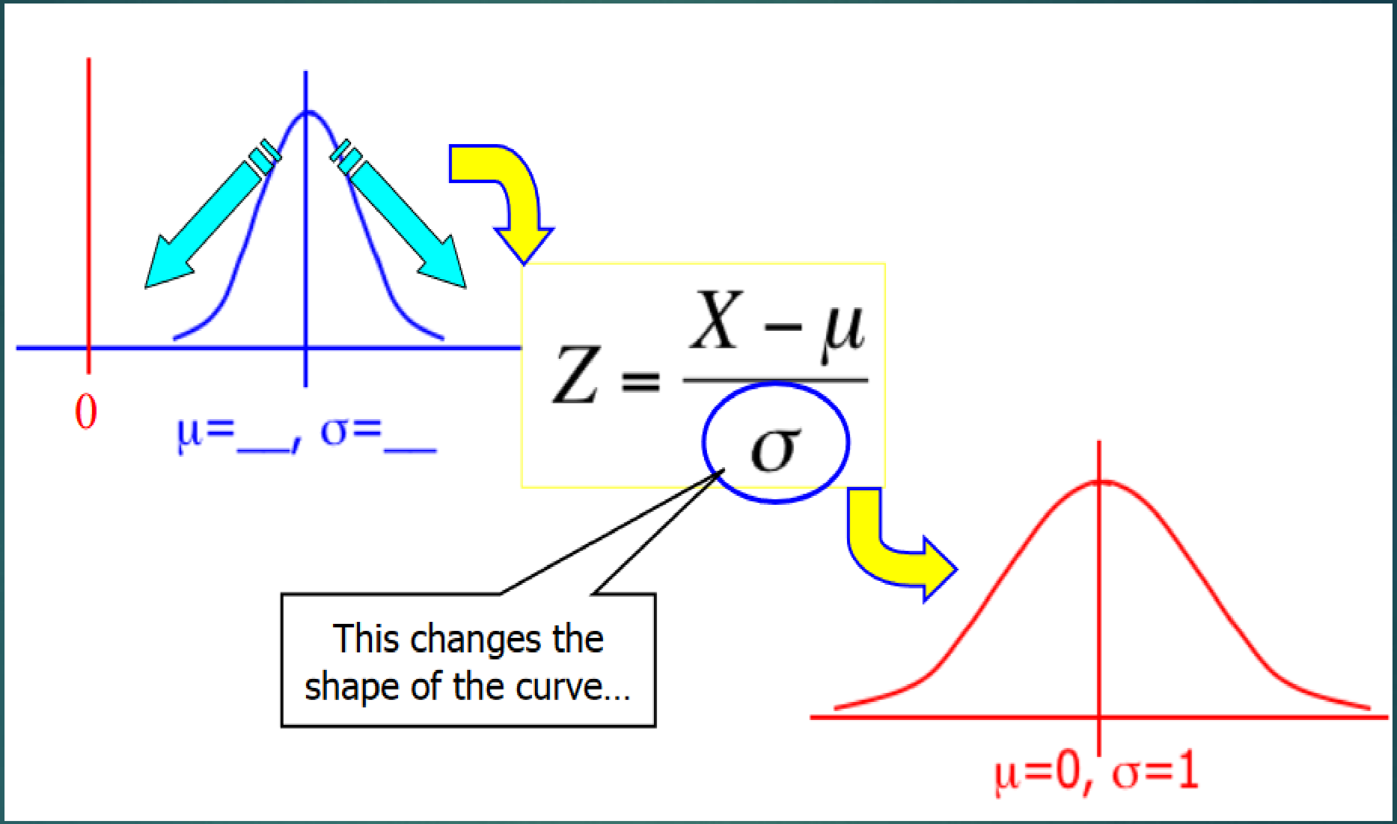
- Denote by Z a random variable that follows the standard normal distribution.
- A standard practice is to convert a normal random variable X with arbitrary parameters μ and σ into a standardized normal random variable Z with parameters 0 and 1 via the transformation:

$$Z = \frac{X - \mu}{\sigma};$$

- this is illustrated in:



Source: Important Probability Distributions, OPRE 6301, page. 27, available at https://personal.utdallas.edu/~scniu/OPRE-6301/documents/Important_Probability_Distributions.pdf



Source: Important Probability Distributions, OPRE 6301, page. 27, available at https://personal.utdallas.edu/~scniu/OPRE-6301/documents/Important_Probability_Distributions.pdf,

What is a Hypothesis?

- A hypothesis is an assumption about the population parameter.
  - A parameter is a Population mean or proportion
  - The parameter must be identified before analysis.

It is assumed the mean GPA of this class is 3.5!

What is Hypotesis Testing?

The hypothesis test is used to evaluate the results from a research study in which

1. A sample is selected from the population.
2. The treatment is administered to the sample.
3. After treatment, the individuals in the sample are measured.

Why Hypothesis Testing ?

- The general goal of a hypothesis test is to rule out chance (sampling error) as a plausible explanation for the results from a research study.
- Hypothesis testing is a technique to help determine whether a specific treatment has an effect on the individuals in a population.

The Null Hypothesis, H_0

- States the Assumption (numerical) to be tested
- E.g. The average # Radio sets in PNG homes is at least 3 (H_0: $\mu \geq 3$)
- Begin with the assumption that the null hypothesis H_0 is TRUE.

(Similar to the notion of innocent until proven guilty)

- Refers to the Status Quo
- Always contains the ' = ' sign
- The Null Hypothesis may or may not be rejected.

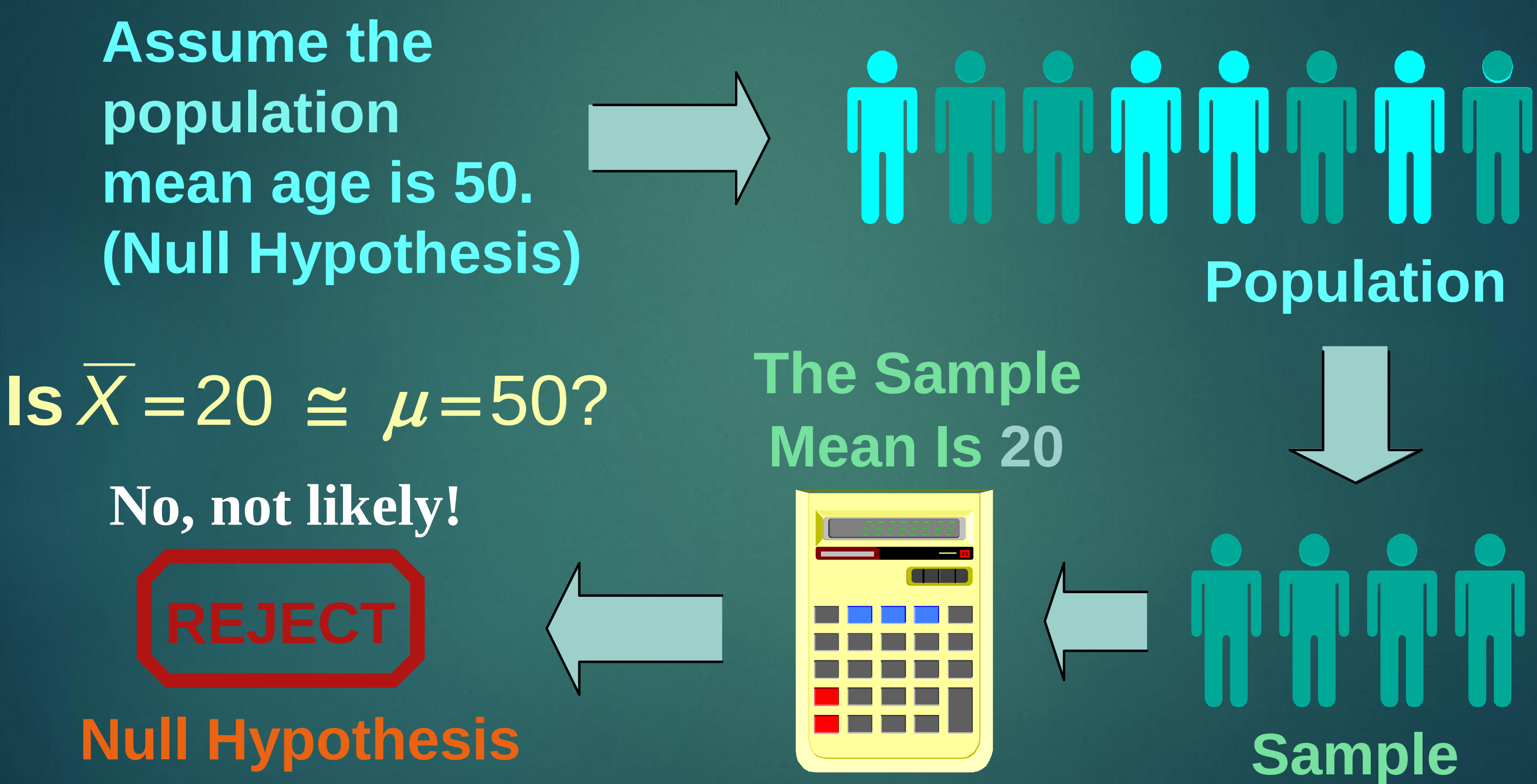
The Alternative Hypothesis, H_1

- **Is the opposite of the null hypothesis H_1 ? e.g. The average # Radio sets in PNG homes is less than 3**
- **(H_1: μ < 3)**
- **Challenges the Status Quo**
- **Never contains the '=' sign**
- **Alternative Hypothesis may or may not be accepted.**

Identify the Problem

- **Steps:**
  - State the Null Hypothesis (H_0: $\mu \geq 3$)
  - State its opposite, the Alternative Hypothesis (H_1: $\mu < 3$)
    - **Hypotheses are mutually exclusive & exhaustive**
    - **Sometimes it is easier to form the alternative hypothesis first.**

Hypothesis Testing Process

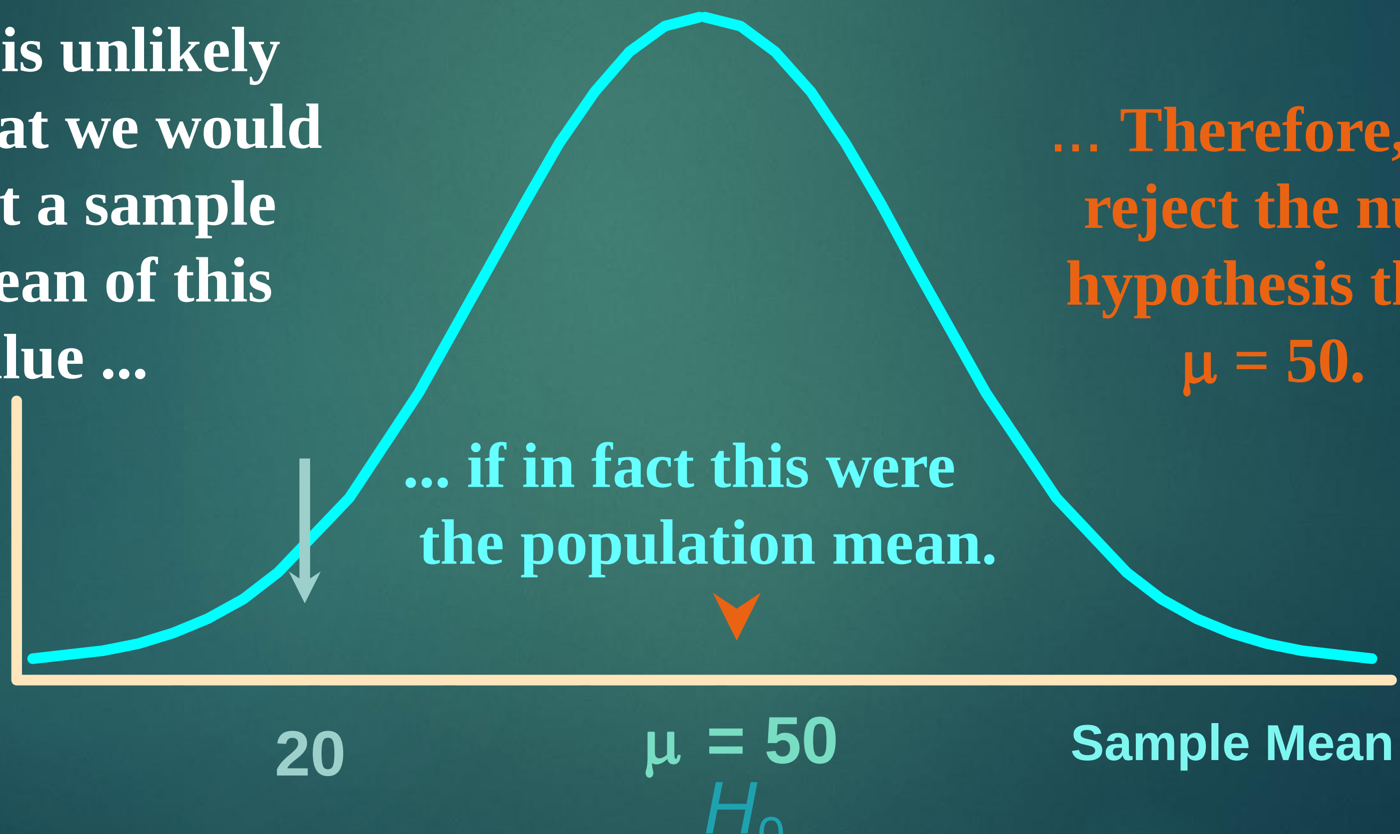


Chapter 11 Sampling and Sampling Distributions, available at *http://users.hist.umn.edu › ~ruggles › hist5011 ›*

Reason for Rejecting H_0

Sampling Distribution

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



… Therefore, we reject the null hypothesis that μ = 50.

Level of Significance, α

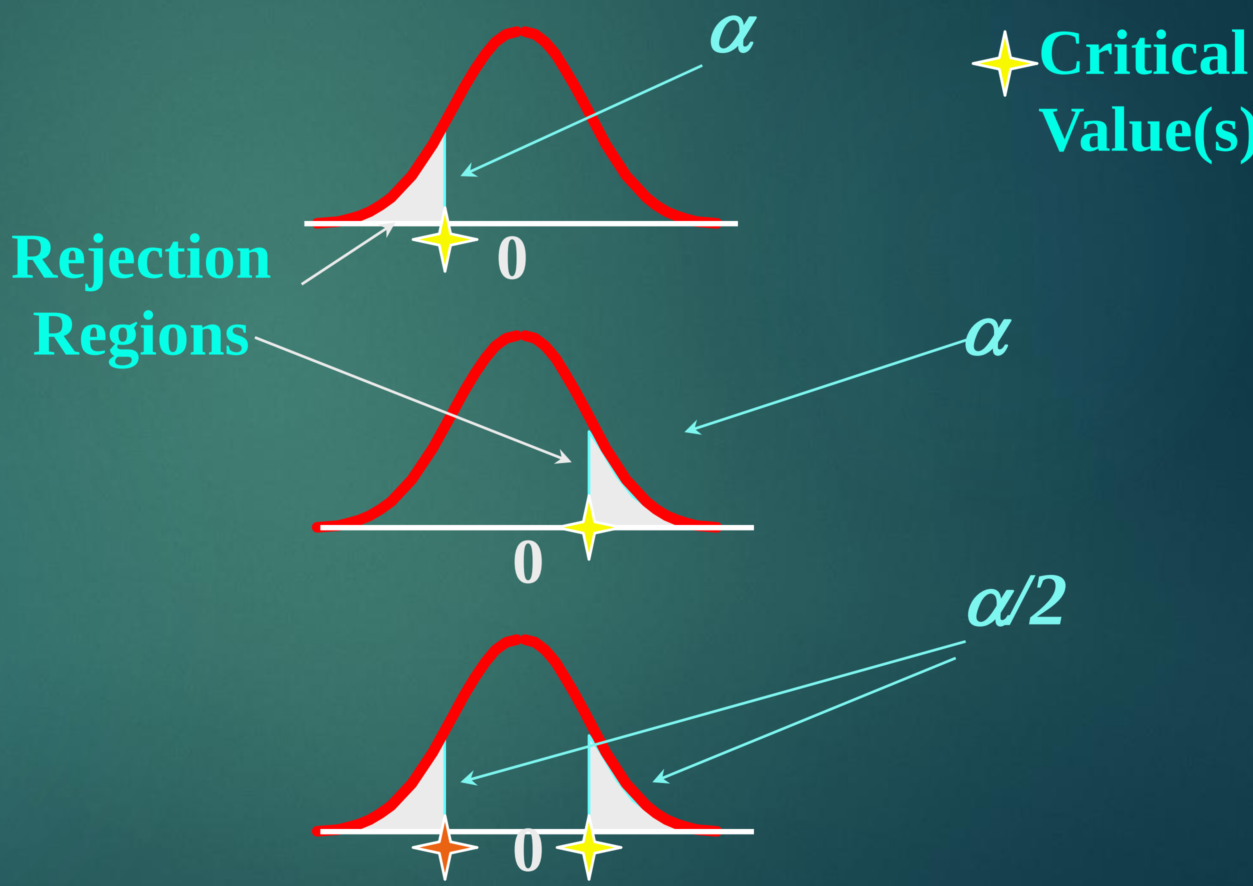
- **Defines Unlikely Values of Sample Statistic if Null Hypothesis Is True**
  - **Called Rejection Region of Sampling Distribution**
- **Designated α (alpha)**
  - **Typical values are 0.01, 0.05, 0.10**
- **Selected by the Researcher at the Start**
- **Provides the Critical Value(s) of the Test**

Level of Significance, α and the Rejection Region

H_0: $\mu \geq 3$
H_1: $\mu < 3$

H_0: $\mu \leq 3$
H_1: $\mu > 3$

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



Chapter 11 Sampling and Sampling Distributions, available at http://users.hist.umn.edu › ~ruggles › hist5011 ›

Errors in Making Decisions

- **Type I Error**
  - **Reject True Null Hypothesis**
  - **Has Serious Consequences**
  - **Probability of Type I Error Is α**
    - **Called Level of Significance**
- **Type II Error**
  - **Do Not Reject False Null Hypothesis**
  - **Probability of Type II Error Is β (Beta)**

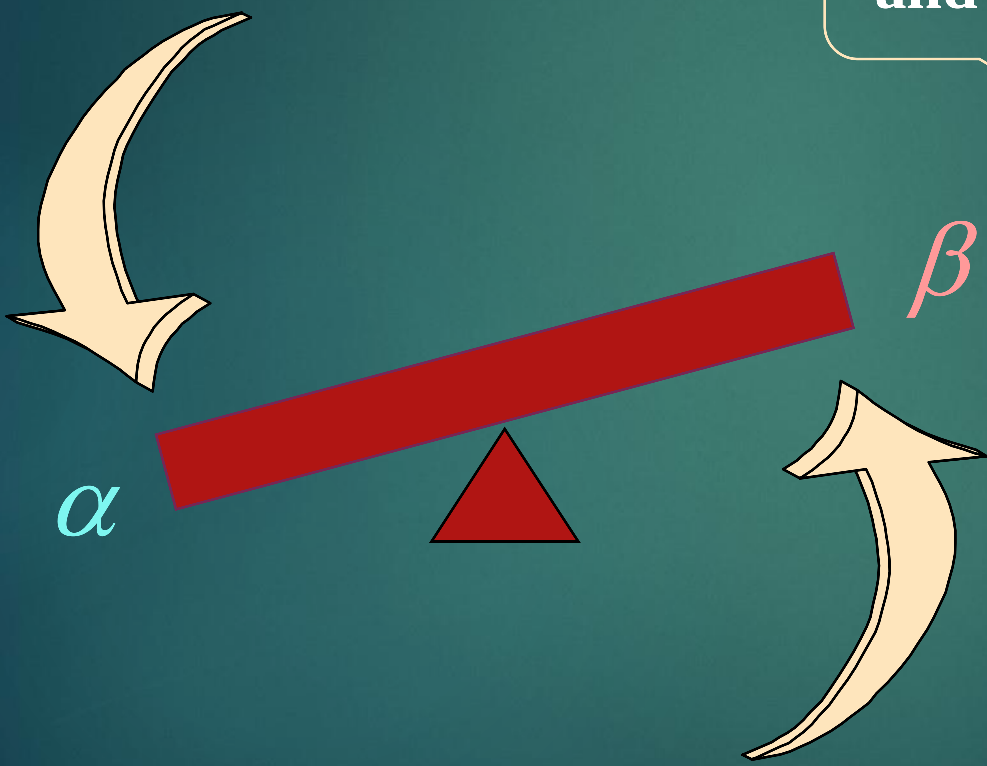
Result Possibilities

H_0: Innocent

Jury Trial			**Hypothesis Test**		
	Actual Situation			Actual Situation	
Verdict	Innocent	Guilty	**Decision**	H_0 True	H_0 False
Innocent	**Correct**	**Error**	Do Not Reject H_0	**$1 - \alpha$**	**Type II Error (β)**
Guilty	**Error**	**Correct**	Reject H_0	**Type I Error (α)**	**Power ($1 - \beta$)**

α & β Have an Inverse Relationship

Reduce probability of one error and the other one goes up.



Z-Test Statistics (σ Known)

- **Convert Sample Statistic (e.g., $\bar{X}$) to Standardized *Z* Variable**

$$Z = \frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

Test Statistic

- **Compare to Critical *Z* Value(s)**
  - **If *Z* test Statistic falls in Critical Region, Reject H_0; Otherwise Do Not Reject H_0**

p Value Test

- **Probability of Obtaining a Test Statistic More Extreme (≤ or ≥) than Actual Sample Value Given H_0 Is True**
- **Called Observed Level of Significance**
  - **Smallest Value of a H_0 Can Be Rejected**
- **Used to Make Rejection Decision**
  - **If *p* value ≥ α, Do Not Reject H_0**
  - **If *p* value < α, Reject H_0**

Hypothesis Testing: Steps

Test the Assumption that the true mean # of Radio sets in PNG homes is at least 3.

- **1. State H_0** $H_0: \mu \geq 3$
- **2. State H_1** $H_1: \mu < 3$
- **3. Choose α** α = .05
- **4. Choose *n*** n = 100
- **5. Choose Test:** *Z Test (or p Value)*

Hypothesis Testing: Steps
(continued)

Test the Assumption that the average # of Radio sets in PNG homes is at least 3.

- **6. Set Up Critical Value(s)** ***Z = -1.645***
- **7. Collect Data** ***100 households surveyed***
- **8. Compute Test Statistic** ***Computed Test Stat.= -2***
- **9. Make Statistical Decision** ***Reject Null Hypothesis***
- **10. Express Decision** ***The true mean # of Radio set is less than 3 in the PNG households.***

Decision Rule 1

1. The decision rule depends on whether an upper-tailed, lower-tailed, or two-tailed test is proposed.
2. In an upper-tailed test the decision rule has investigators reject H_0 if the test statistic is larger than the critical value.
3. In a lower-tailed test the decision rule has investigators reject H_0 if the test statistic is smaller than the critical value.
4. In a two-tailed test the decision rule has investigators reject H_0 if the test statistic is extreme, either larger than an upper critical value or smaller than a lower critical value.

Decision Rules 2 & 3

2. The exact form of the test statistic is also important in determining the decision rule. If the test statistic follows the standard normal distribution (Z), then the decision rule will be based on the standard normal distribution.

3. The third factor is the level of significance. The level of significance which is selected in Step 1 (e.g., a =0.05) dictates the critical value. For example, in an upper tailed Z test, if a =0.05 then the critical value is Z=1.645.

Example: Two Tail Test

- Does an average box of cereal contain **368** grams of cereal? A random sample of **25** boxes showed $\bar{X}$ = **372.5**. The company has specified σ to be **15** grams. Test at the α=**0.05** level.

$$H_0: \mu = 368$$
$$H_1: \mu \neq 368$$

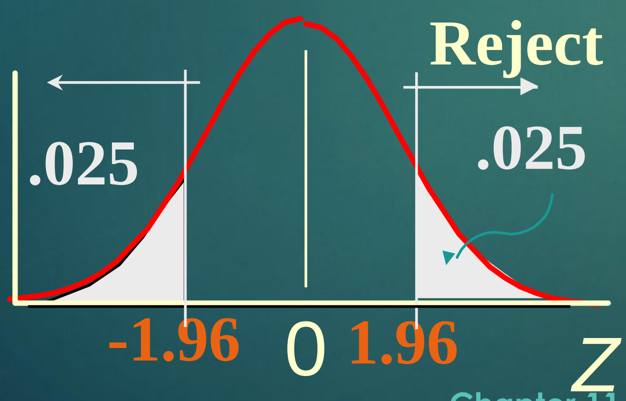
Chapter 11 Sampling and Sampling Distributions, available at http://users.hist.umn.edu › ~ruggles › hist5011 ›

Example Solution: Two Tail

$H_0: \mu = 368$

$H_1: \mu \neq 368$

- $\alpha = 0.05$
- $n = 25$
- **Critical Value: ±1.96**



Test Statistic:

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} = \frac{372.5 - 368}{15 / \sqrt{25}} = 1.50$$

Decision:

Do Not Reject at α = .05

Conclusion:

No Evidence that True Mean Is Not 368

Chapter 11 Sampling and Sampling Distributions, available at http://users.hist.umn.edu › ~ruggles › hist5011 ›

Connection to Confidence Intervals

For $\bar{X} = 372.5$, $\sigma = 15$ and $n = 25$,

The 95% Confidence Interval is:

$372.5 - (1.96)\sqrt{15}/25$ to $372.5 + (1.96)\sqrt{15}/25$

or

$$366.62 \leq \mu \leq 378.38$$

If this interval contains the Hypothesized mean (368), we do not reject the null hypothesis.

It does. Do not reject.

Student ‘t’ distribution

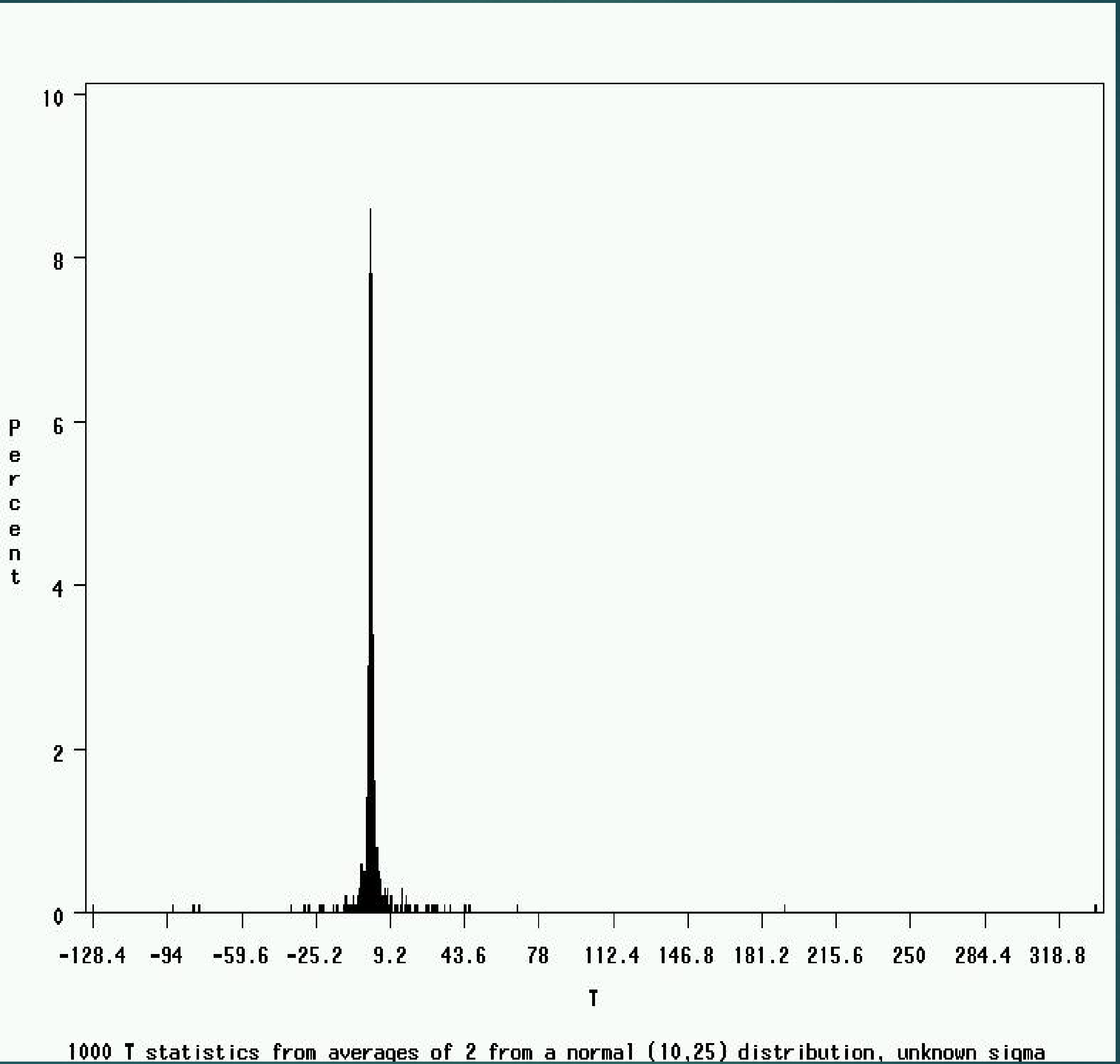
What we learn?

- What is a t test ?
- How is the distribution of *t* related to the unit normal?
- When would we use a t-test instead of a z-test? Why might we prefer one to the other?
- What are the chief varieties or forms of the *t*-test?
- What is the standard error of the difference between means? What are the factors that influence its size?

What is a T-distribution?

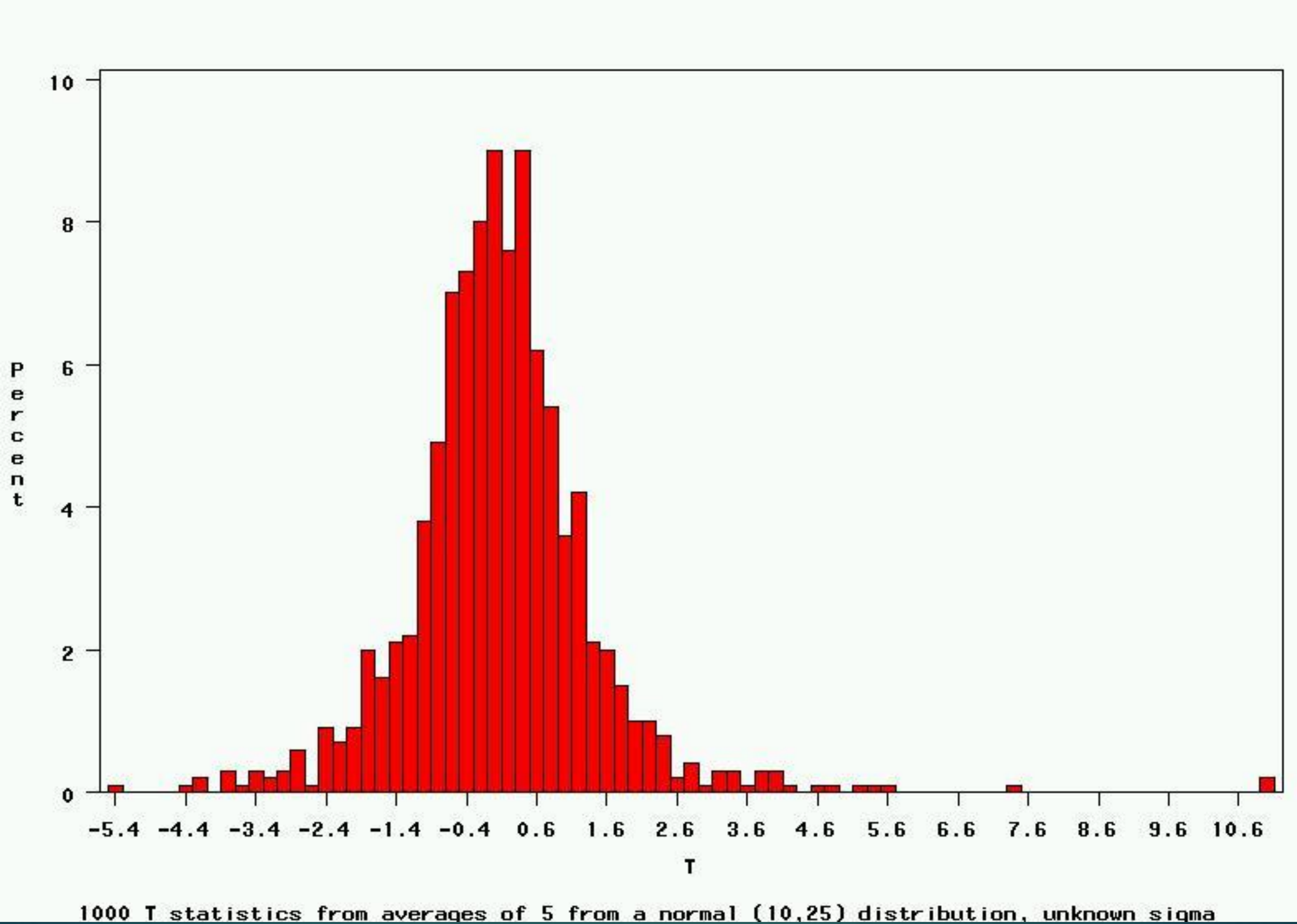
- A t-distribution is like a Z distribution, except has slightly fatter tails to reflect the uncertainty added by estimating σ.
- The bigger the sample size (i.e., the bigger the sample size used to estimate σ), then the closer t becomes to Z.
- If n>100, t approaches Z.

T-distribution with only 1 degree of freedom.



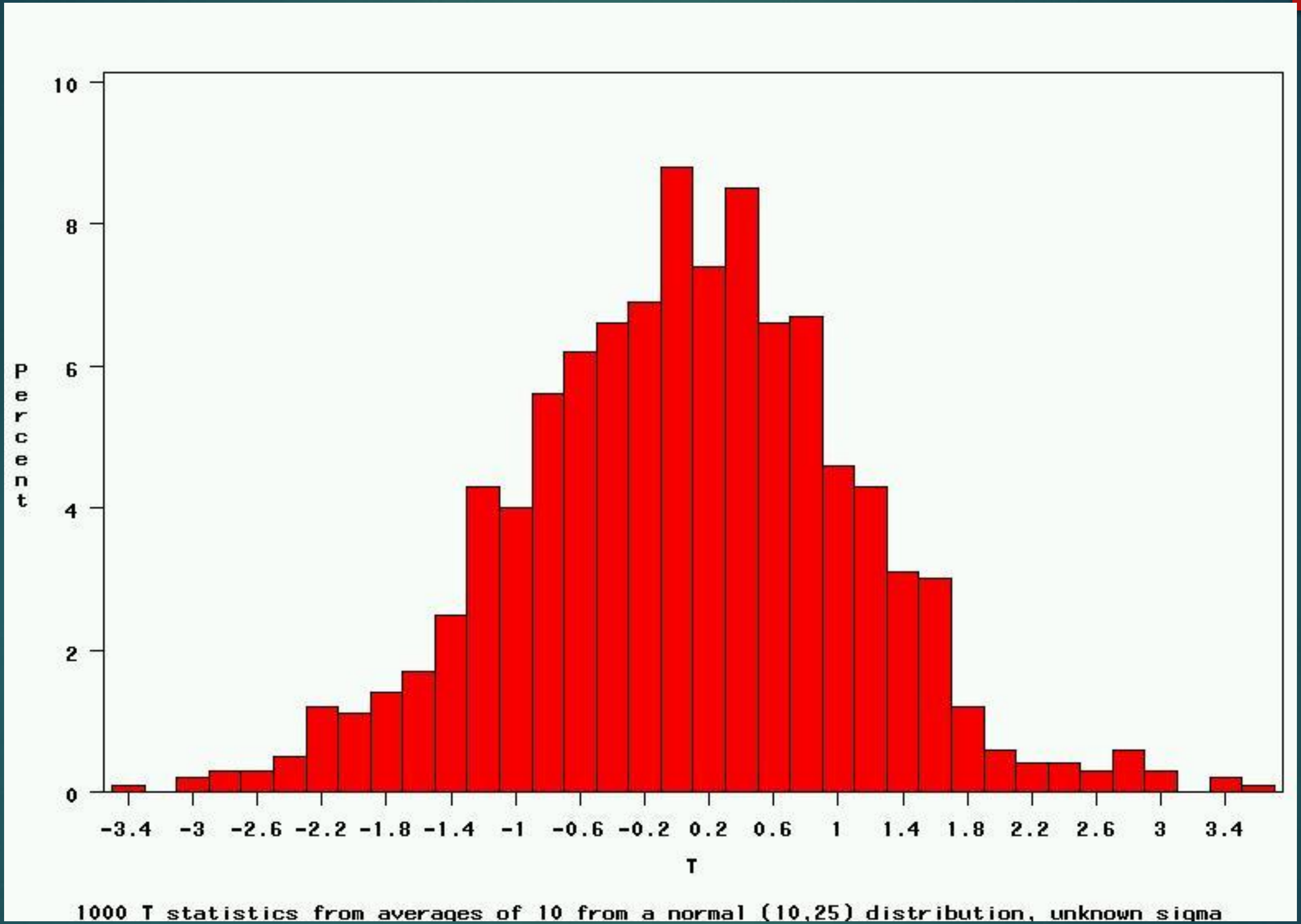
Chapter 11 Sampling and Sampling Distributions, available at *http://users.hist.umn.edu › ~ruggles › hist5011 ›*

T-distribution with 4 degrees of freedom.



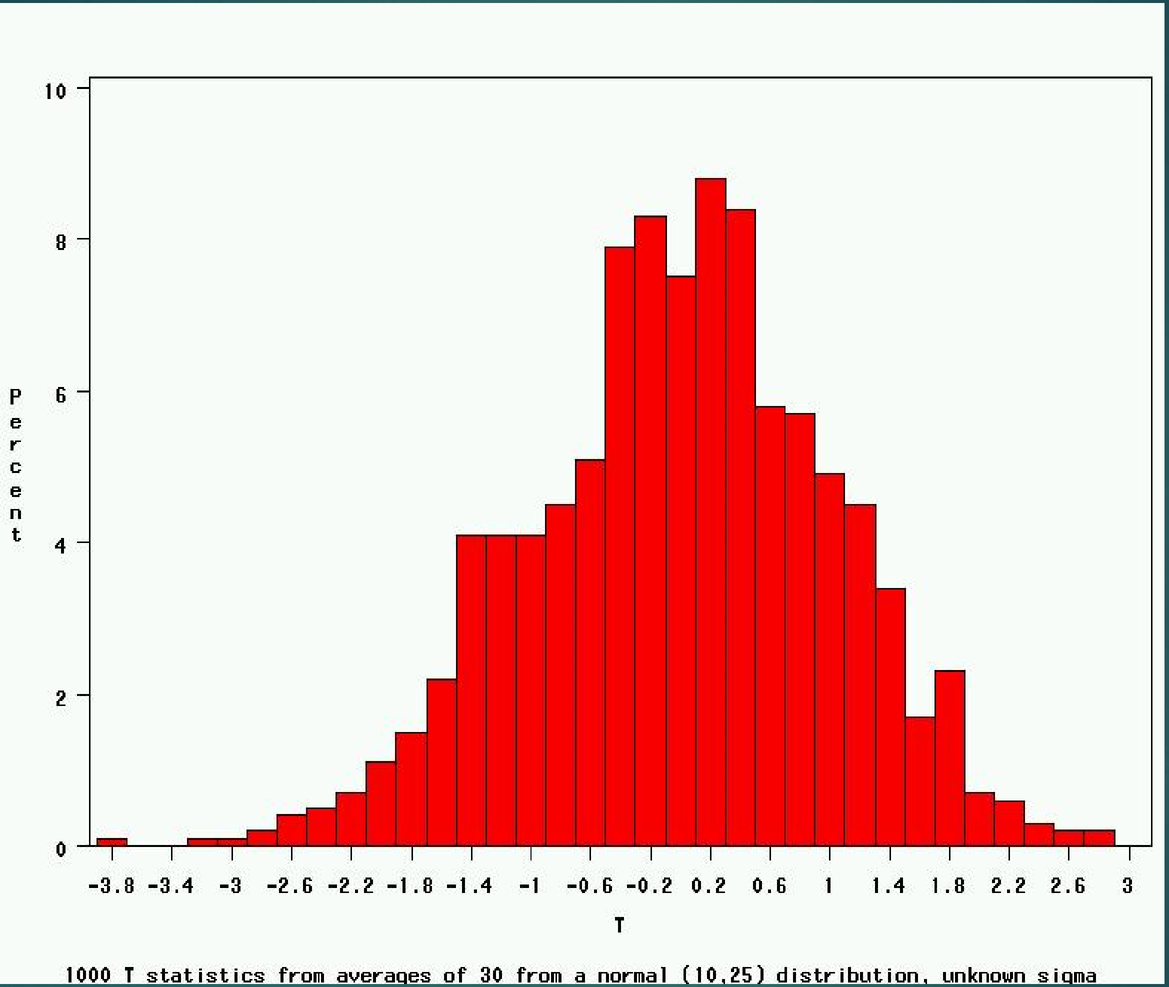
[Chapter 11 Sampling and Sampling Distributions, available at](http://users.hist.umn.edu) *http://users.hist.umn.edu › ~ruggles › hist5011 ›*

T-distribution with 9 degrees of freedom.



[Chapter 11 Sampling and Sampling Distributions, available at](http://users.hist.umn.edu/~ruggles/hist5011/)
http://users.hist.umn.edu › ~ruggles › hist5011 ›

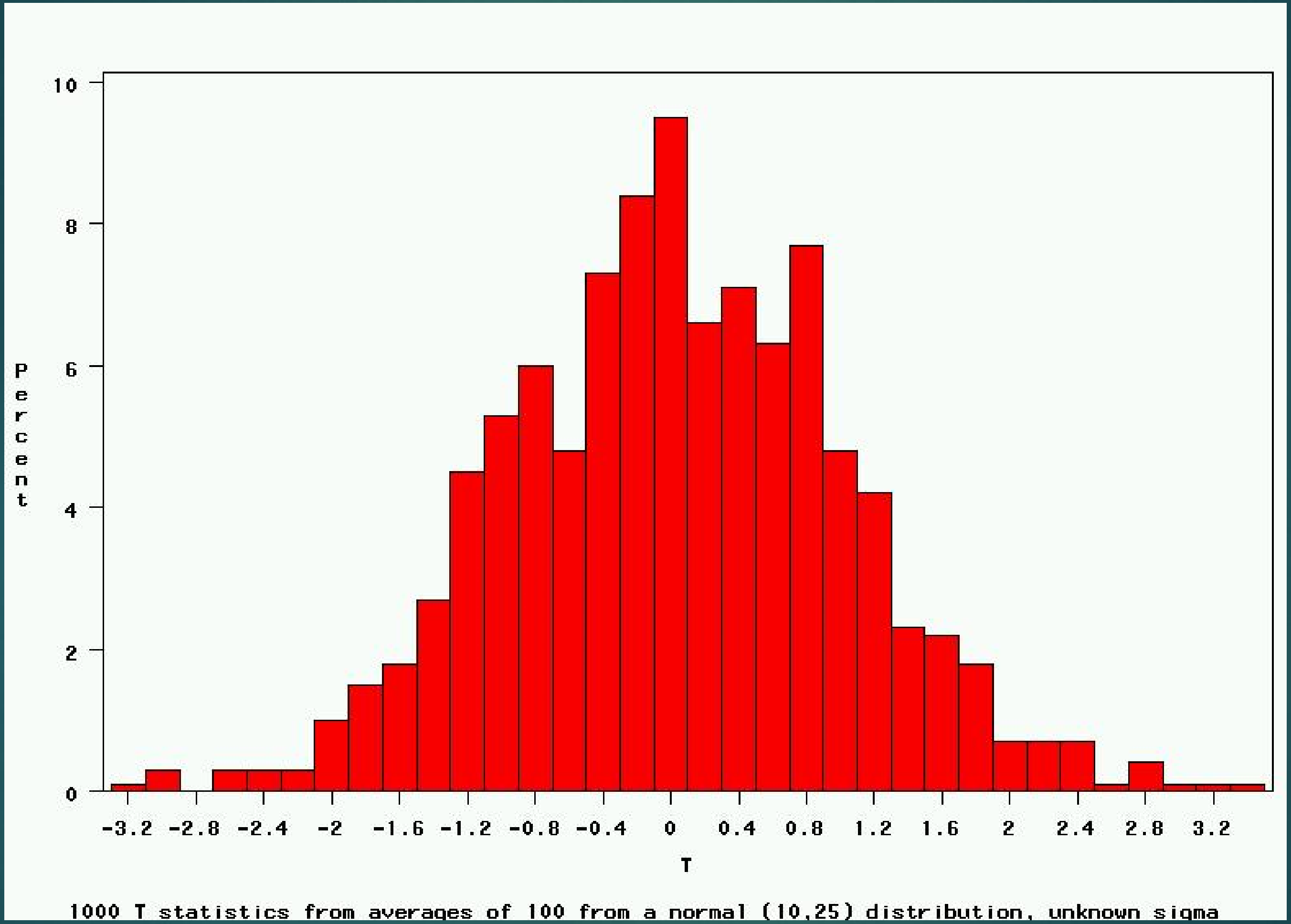
T-distribution with 29 degrees of freedom.



Chapter 11 Sampling and Sampling Distributions, available at *http://users.hist.umn.edu › ~ruggles › hist5011 ›*

T-distribution with 99 degrees of freedom.

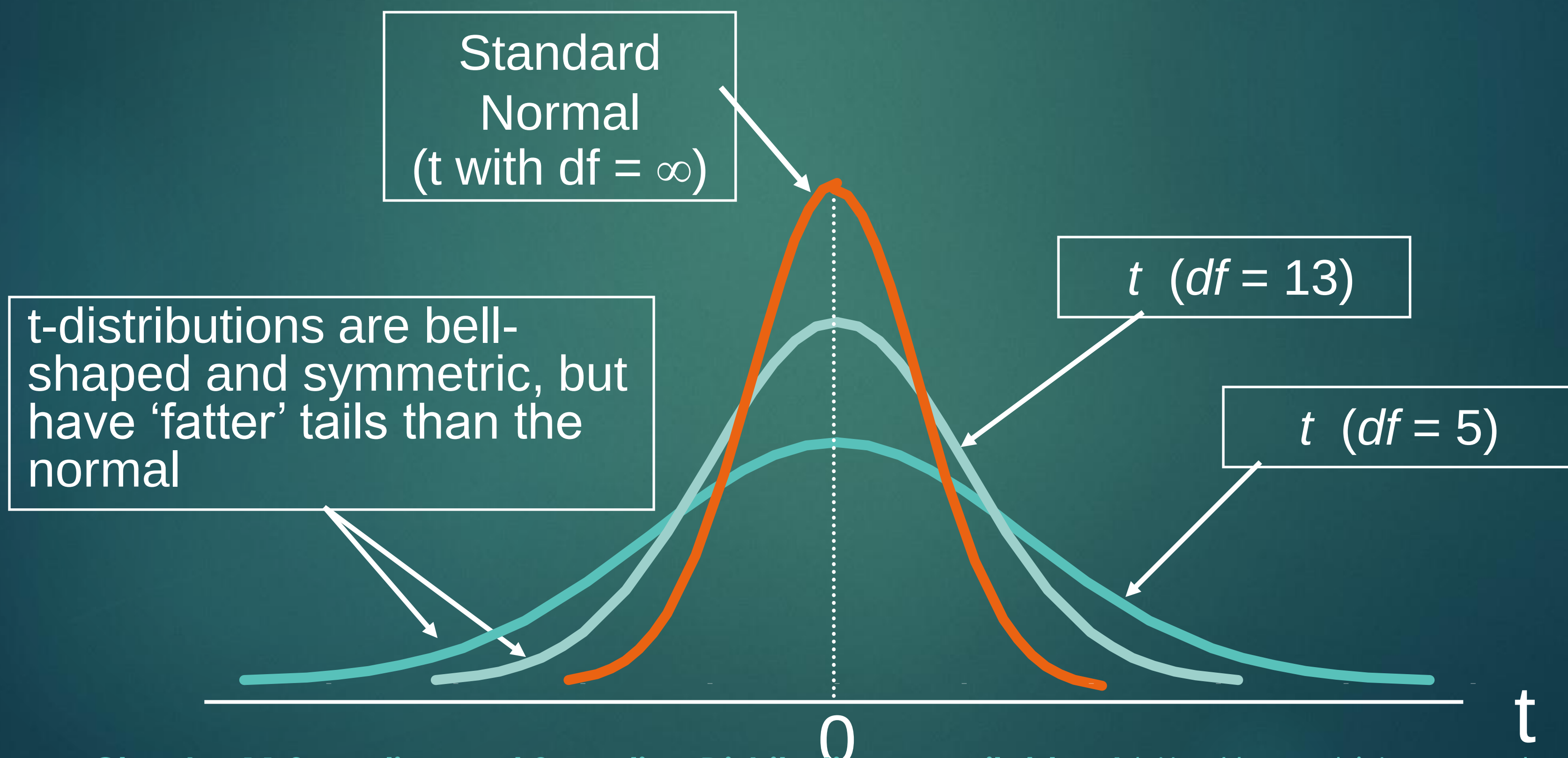
Looks a lot like Z!!



[Chapter 11 Sampling and Sampling Distributions, available at](http://users.hist.umn.edu/~ruggles/hist5011/)
http://users.hist.umn.edu › ~ruggles › hist5011 ›

Student’s t Distribution

Note: t ⟶ Z as n increases

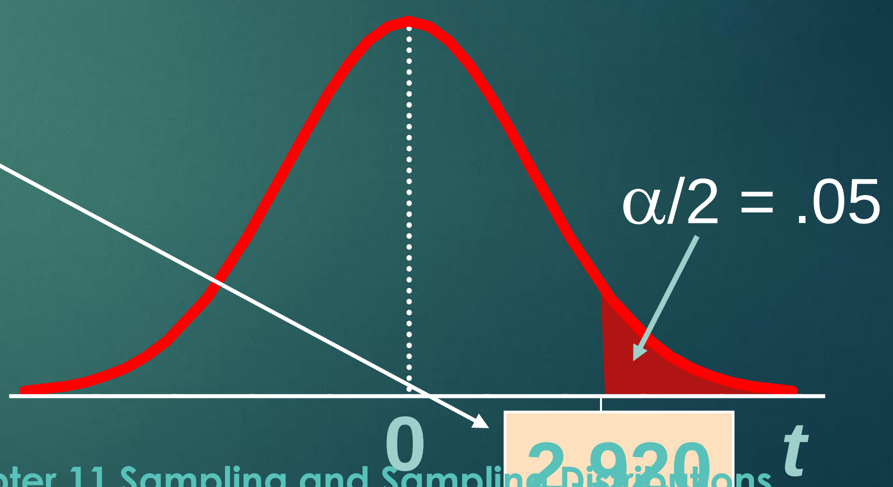


Student's t Table

	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



Chapter 11 Sampling and Sampling Distributions, available at *http://users.hist.umn.edu › ~ruggles › hist5011 ›*

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 → Z as n increases

The normality assumption…

- T tests (and all linear models, in fact) have a "normality assumption":
  - If the outcome variable is not normally distributed and the sample size is small, a ttest is inappropriate
  - it takes longer for the CLT to kick in and the sample means do not immediately follow a t-distribution…
- This is the source of the "normality assumption" of the ttest…

Assumptions

- The *t*-test is based on assumptions of normality and homogeneity of variance.
- As long as the samples in each group are large and nearly equal, the *t*-test is robust, that is, still good, even though assumptions are not met.

The *t* Distribution

We use *t* when the population variance is unknown (usual case) and sample size is small (N<100, the usual case).

The *t* distribution is a short, fat relative of the normal. The shape of *t* depends on its *df*. As *N* becomes infinitely large, *t* becomes normal.

The *t* Distribution

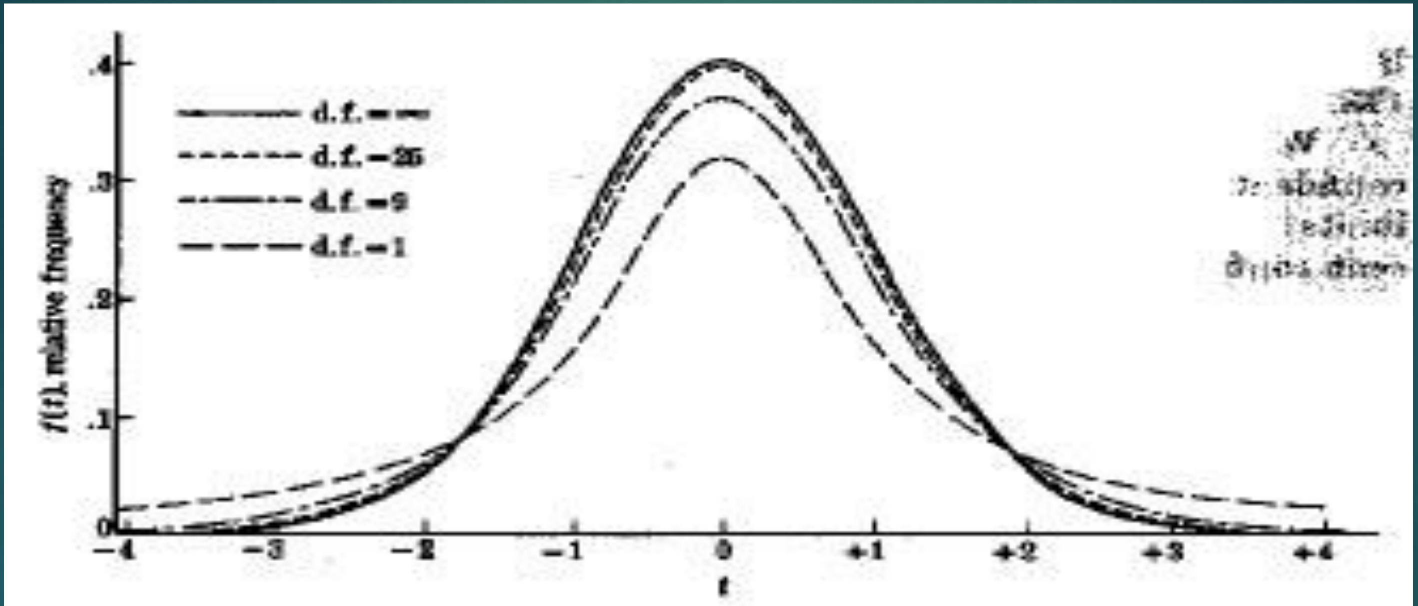


fig. 10.1 Distribution of *t* for various degrees of freedom. (*From D. Lewis, quantitative methods in psychology, McGraw-Hill Book Company, New York, 1960.*)

Chapter 11 Sampling and Sampling Distributions, available at *http://users.hist.umn.edu › ~ruggles › hist5011 ›*

Degrees of Freedom

For the *t* distribution, degrees of freedom are always a simple function of the sample size, e.g., (N-1).

One way of explaining *df* is that if we know the total or mean, and all but one score, the last (N-1) score is not free to vary. It is fixed by the other scores. $4+3+2+X = 10$. $X=1$.

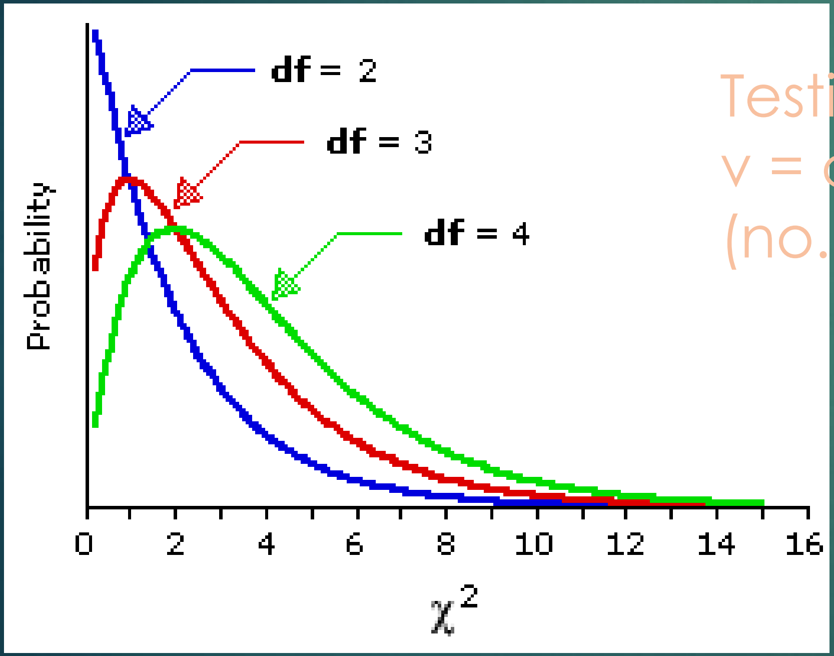
Summing up t distribution

- If the underlying data are not normally distributed AND n is small**, the means do not follow a t-distribution (so using a t test will result in erroneous inferences).
- Data transformation or non-parametric tests should be used instead.
- **How small is too small? No hard and fast rule—depends on the true shape of the underlying distribution. Here N>30 (closer to 100) is needed.

- Chi squared distribution

Chi squared distribution

- The p-value is calculated using the Chi-squared distribution for this test
- Chi-squared is a skewed distribution which varies depending on the degrees of freedom



Testing relationships between 2:
v = degrees of freedom
(no. of rows – 1) x (no. of columns – 1)

Note: One sample test:
v = df = outcomes – 1

Chapter 11 Sampling and Sampling Distributions, available at *http://users.hist.umn.edu › ~ruggles › hist5011 ›*

Chi squared distribution

- The previous slide illustrates what the Chi-squared distribution looks like and that as the degrees of freedom increases (i.e. the amount of independent pieces of information we have from a study) then the more the wider the distribution is spread and becomes flatter.

5-1. Statistical Prerequisites

- See Appendix A with key concepts such as probability, probability distributions, Type I Error, Type II Error, level of significance, power of a statistic test, and confidence interval

The 68–95–99.7 Rule

- In statistics, the 68–95–99.7 rule, also known as the empirical rule, is **a shorthand used to remember the percentage of values that lie within an interval estimate in a normal distribution.**
- **Video from Learning Pro from youtube** **https://www.youtube.com/watch?v=mtbJbDwqWLE**

Before we conclude

- The one-sample z test checks if there is a difference in the sample and population mean,
- The two sample z test checks if the means of two different groups are equal.
- **Z Test vs T-Test**
- Both z test and t-test are univariate tests used on the means of two datasets. The differences between both tests are outlined in the table given below:

Z Test	T-Test
A z test is a statistical test that is used to check if the means of two data sets are different when the population variance is known.	A t-test is used to check if the means of two data sets are different when the population variance is not known.
The sample size is greater than or equal to 30.	The sample size is lesser than 30.
The data follows a normal distribution.	The data follows a student-t distribution.
The one-sample z test statistic is given by $\frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$	The t test statistic is given as – where s is the sample standard deviation $\frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}}$

Example 1:Z test

- An online medicine shop claims that the mean delivery time for medicines is less than 120 minutes with a standard deviation of 30 minutes. Is there enough evidence to support this claim at a 0.05 significance level if 49 orders were examined with a mean of 100 minutes?
- **Solution:** As the sample size is 49 and population standard deviation is known, this is an example of a left-tailed one-sample z test.
- H0: $\mu = 120$; H1: $\mu < 120$
- From the z table the critical value at $\alpha = -1.645$. A negative sign is used as this is a left tailed test.

Example 1: **Z test**

- Z = $\dfrac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$
- = 100, μ = 120, n = 49, σ= 30
- z = -4.66
- As -4.66 < -1.645 thus, the null hypothesis is rejected and it is concluded that there is enough evidence to support the medicine shop's claim.
- **Answer:** Reject the null hypothesis

Single-sample *t*-test

With a small sample size, we compute the same numbers as we did for *z*, but we compare them to the *t* distribution instead of the *z* distribution.

$$H_0 : \mu = 10; \quad H_1 : \mu \neq 10; \quad s_X = 5; \quad N = 25$$

$$est.\sigma_M = \frac{s_X}{\sqrt{N}} = \frac{5}{\sqrt{25}} = 1$$

$$\bar{X} = 11 \rightarrow t = \frac{(11-10)}{1} = 1$$

$t(.05,24) = 2.064$ (c.f. z=1.96) 1<2.064, n.s.

Interval = $\bar{X} \pm t\hat{\sigma}_M$

$$11 \pm 2.064(1) = [8.936, 13.064]$$

Interval is about 9 to 13 and contains 10, so n.s.

Reference

CHAPTER 11 SAMPLING AND SAMPLING DISTRIBUTIONS, AVAILABLE AT *HTTP://USERS.HIST.UMN.EDU › ~RUGGLES › HIST5011 ›*

HTTPS://WWW.GOOGLE.COM/URL?SA=T&RCT=J&Q=&ESRC=S&SOURCE=WEB&CD=&CAD=RJA&UACT=8&VED=2AHUKEWJ_6ZVPY4D3AHWDIBCAHA5WDFKQFNOECBEQAQ&URL=HTTPS%3A%2F%2FWEB.STANFORD.EDU%2F~KCOBB%2FHRP259%2FLECTURE8.PPT&USG=AOVVAW1BVTCGH0JJMB1EK7DLQB9W

VIDEO FROM LEARNING PRO FROM YOUTUBE HTTPS://WWW.YOUTUBE.COM/WATCH?V=MTBJBDWQWLE

What Next ?

- Simple Regression: Testing of Hypothesis