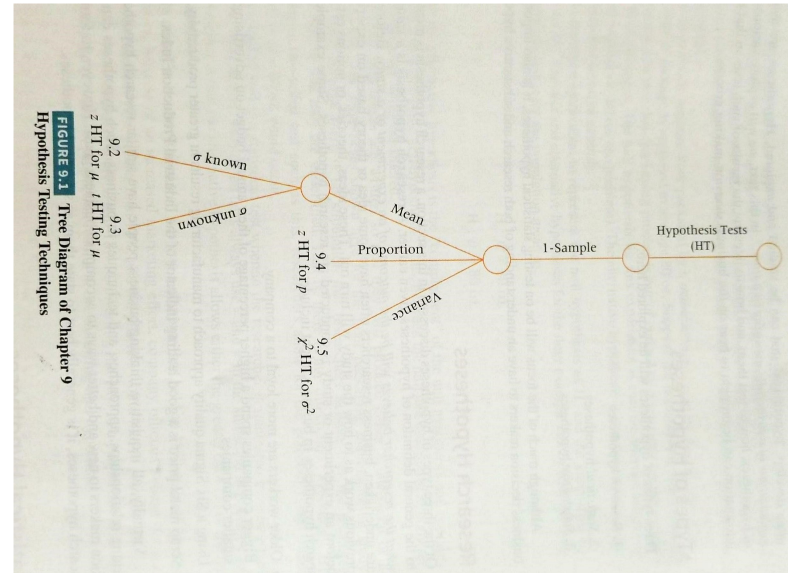
Week 6 - Review

Friday, February 16, 2018 3:43 AM

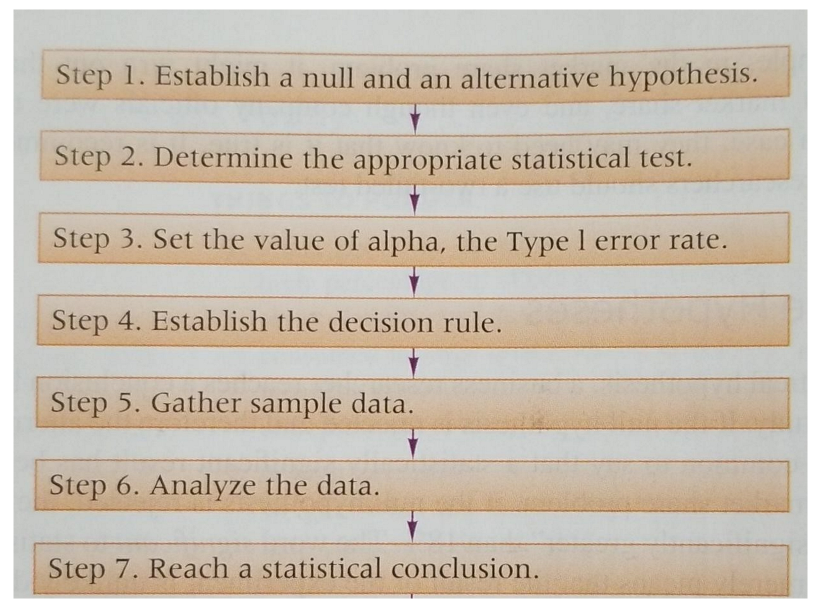
Types of Hypotheses

- **Research hypotheses**
  - It is a statement of what the research believes will be the outcome of an experiment of study.
    - Example: Older workers are more loyal to a company. Bigger companies spend a higher percentage of their annual budget on advertising than smaller companies.
- **Statistical hypotheses**
  - Consist of two parts: null hypothesis and alternative hypothesis
    - ***Null Hypothesis***
      - States that the null condition exists; that is, there is nothing new happening, the old theory is still true, the old standard is correct, and the system is in control.
    - ***Alternative Hypothesis***
      - States that the new theory is true, there are new standards, the systems is out of control and/or something is happening.
  - Types of tests
    - **Two-tailed test** is when the researcher is interested in testing both side of the distribution.
      - Always use = and != in the statistical hypotheses.
    - **One-tailed test** is when the researcher is interested in testing only one side of the distribution.
      - One-tailed tests are always directional and use either greater than (>) or less than (<) signs.
- **Substantive hypotheses**
  - It is when the outcome of a statistical study produces results that are important to the decision maker.

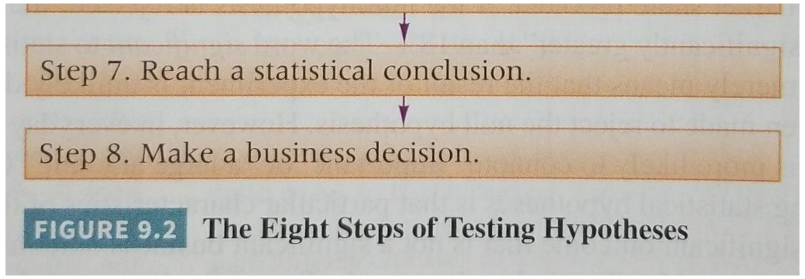


Eight-Step Process of Testing Hypotheses

1. Establish a null and an alternative hypothesis
2. Determine the appropriate statistical test
3. Set the value of alpha, the Type 1 error rate
4. Establish the decision rule
5. Gather sample data
6. Analyze the data
7. Reach a statistical conclusion
8. Make a business decision



7. Reach a statistical conclusion
8. Make a business decision


Step 7. Reach a statistical conclusion.

Step 8. Make a business decision.

FIGURE 9.2 The Eight Steps of Testing Hypotheses

Type I and Type II Errors

- **Type I Error** -committed by rejecting a true null hypothesis
  - **Alpha** or **level of significance** - the probability of committing a Type 1 Error.
- **Type II Error** - committed when fail to reject a false null hypothesis.
  - **Beta** - the probability of committing a Type II error.
    - Beta is typically not stated at the beginning of hypothesis testing procedure.
- **Relationship between Alpha and Beta**
  - Alpha and beta are inversely related. If alpha is reduced, beta is increased and vice versa.
- **Power *(1 - Beta)*** - The probability of a statistical test rejecting the null hypothesis when the null hypothesis is false
- **Operating Characteristics (OC) Curve**
  - This is constructed by plotting Beta values against the various values of the alternative hypothesis.

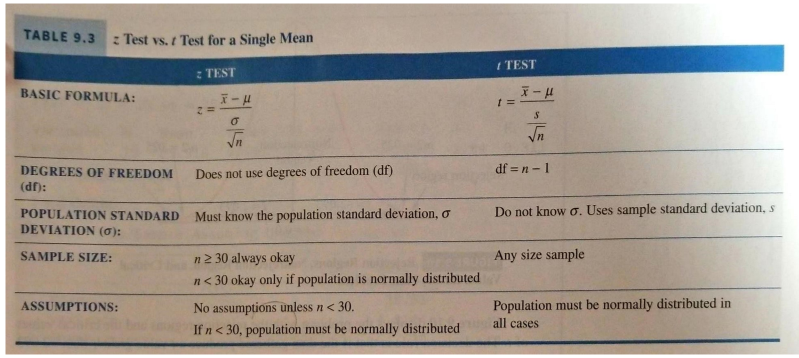


TABLE 9.3 z Test vs. t Test for a Single Mean

	z TEST	t TEST
BASIC FORMULA:	$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$	$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$
DEGREES OF FREEDOM (df):	Does not use degrees of freedom (df)	df = $n - 1$
POPULATION STANDARD DEVIATION (σ):	Must know the population standard deviation, σ	Do not know σ. Uses sample standard deviation, s
SAMPLE SIZE:	$n \geq 30$ always okay $n < 30$ okay only if population is normally distributed	Any size sample
ASSUMPTIONS:	No assumptions unless $n < 30$. If $n < 30$, population must be normally distributed	Population must be normally distributed in all cases

Testing Hypothesis

- **P-Value Method (aka observed significance level)**
  - The p-value defines the smallest value of alpha for which the null hypothesis can be rejected.
- **Critical Value Method**
  - The critical value method determines the critical mean value required for z to be in the rejection region and uses it to test the hypothesis.
    - **Rejection region** - the area where statistical outcomes that results in the rejection of the null hypothesis
    - **Non-rejection region** - the area where statistical outcomes that fail to result in the rejection of the null hypothesis lie.

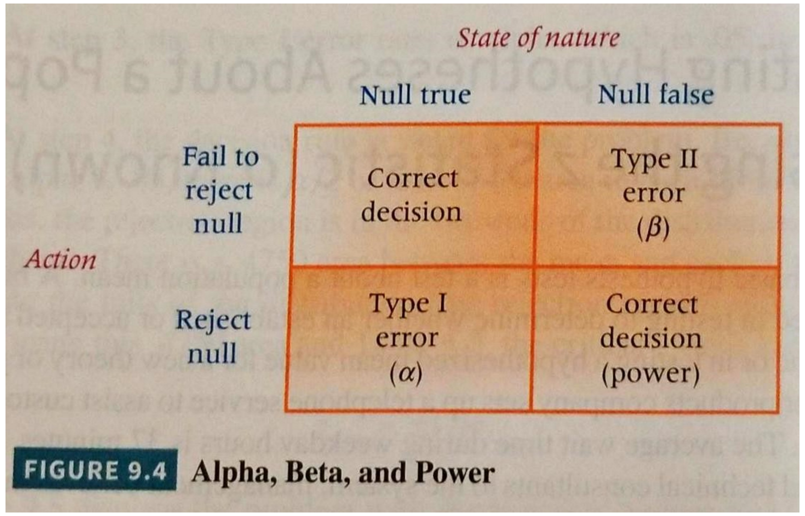
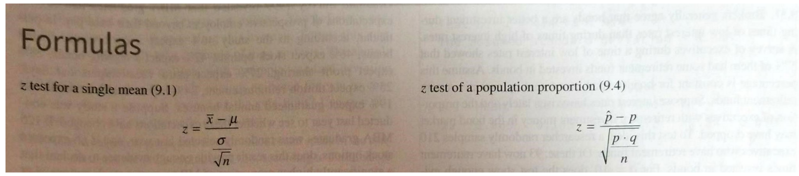


FIGURE 9.4 Alpha, Beta, and Power

Chi-Square

- **Chi-Square Goodness-of-Fit Test**
  - This is used to analyze the probabilities of multinomial distribution trails along a single dimensions
  - This compares the expected, or theoretical, frequencies of categories from a population distribution to the observed, or actual, frequencies from a distribution to determine whether there is a difference between what was expected and what was observed.
- **Chi-Square Distribution**
  - It is based on a sum of squared terms, where the individual terms are approximately normal when the sample size is sufficiently large.
  - More formal definition, is that it is based on the sum of squared standard normal random variables.
  - It can never be less than zero
  - It extends indefinitely in the positive direction


Formulas

z test for a single mean (9.1)

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

z test of a population proportion (9.4)

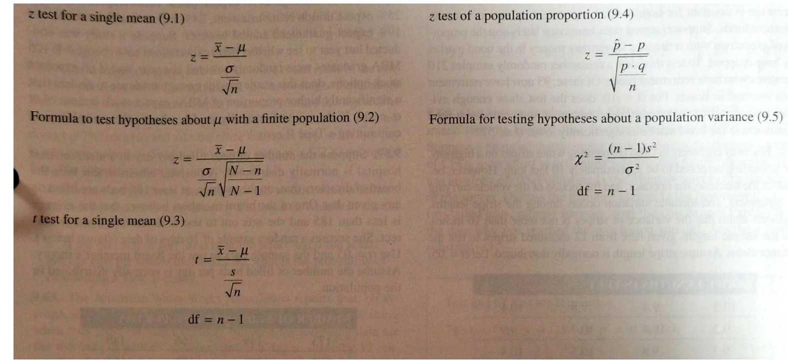
$$z = \frac{\hat{p} - p}{\sqrt{\frac{p \cdot q}{n}}}$$

more formal definition, is that it is based on the sum of squared standard normal random variables.
    - It can never be less than zero
    - It extends indefinitely in the positive direction
- **Chi-Square Test of Independence**
    - The Chi-squared test of independence performs reasonably well in terms of Type I errors, but difficulties can arise, particularly when the number of observations is relatively small.
    - This is used to analyze the frequencies of two variables with multiple categories to determine whether the two variance are independent.
        - Example: understand whether the type of soft drink preferred by a consumer is independent of the consumer's age. Whether type of preferred stock investment in independent of the region where the investor resides.
    - Can be used to analyze any level of data measurement but is particularly useful in analyzing nominal data.

Statistical and Practical Significance

- **Statistical Significance**
    - Could we have obtained this result by chance?
    - Any time the calculated value of t is greater than the critical value of t (from a table) than our results are statistically significance
        - "The probability of us obtaining this results by chance alone if the real correlation coefficient was equal to zero is less than five out of 100.
        - If we wanted a higher level of confidence we could choose a level of significance of 0.01 which would mean that the probability would be less than one out of 100.
    - Important because it allows us to make judgement about how likely it is that a result could have4 been obtained by chance alone.

- **Practical Significance**
    - MORE important because it give us information about how USEFUL a result is.
    - Does the results "matter"?
- ***The larger the sample size, the smaller the correlation that will be statistically significant.***
    - *"*The odds of us getting a result like this from chance alone IF the TRUE correlation was zero, would be less than 5 out of 100. Therefore, we are justified in saying that there IS in fact a relationship."
    - Larger the sample size, smaller t value required for statistical significance
- **Coefficient of Determination: r^2**
    - Another way in which to interpret the correlation coefficient in terms of practical significance is to calculate the "r^2"
    - The coefficient of determination tells you how much of the variability in one variable is explained by the other variable.
    - If r = 0.5, then r^2 = .25



z test for a single mean (9.1)

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

Formula to test hypotheses about μ with a finite population (9.2)

$$z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}}$$

t test for a single mean (9.3)

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

$$df = n - 1$$

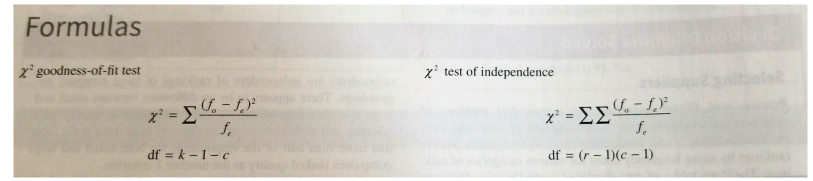
z test of a population proportion (9.4)

$$z = \frac{\hat{p}-p}{\sqrt{\frac{p \cdot q}{n}}}$$

Formula for testing hypotheses about a population variance (9.5)

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

$$df = n - 1$$



Formulas

χ^2 goodness-of-fit test

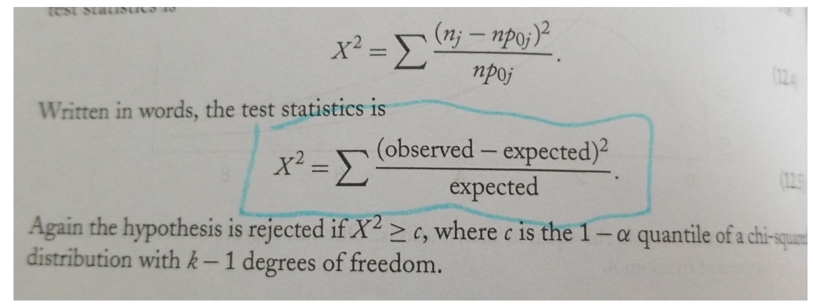
$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$

$$df = k - 1 - c$$

χ^2 test of independence

$$\chi^2 = \sum\sum \frac{(f_o - f_e)^2}{f_e}$$

$$df = (r-1)(c-1)$$



test statistics is

$$X^2 = \sum \frac{(n_j - np_{0j})^2}{np_{0j}}.$$

Written in words, the test statistics is

$$X^2 = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}.$$

Again the hypothesis is rejected if $X^2 \geq c$, where c is the $1-\alpha$ quantile of a chi-squ distribution with $k-1$ degrees of freedom.

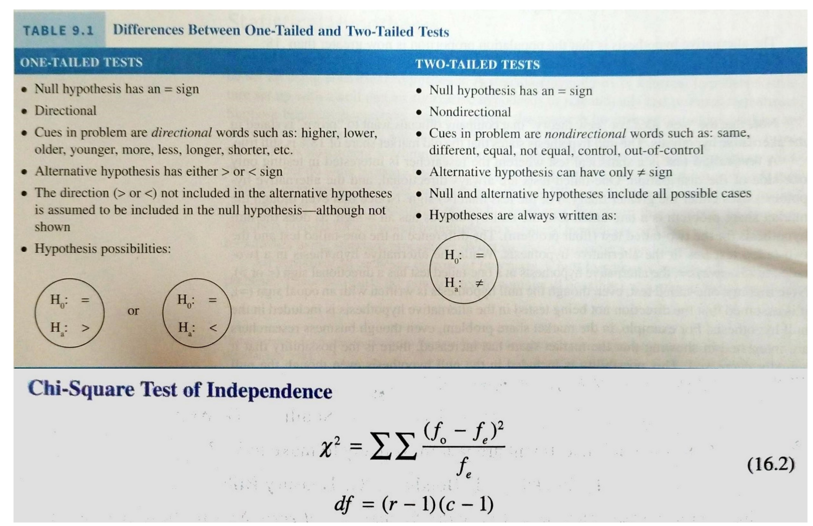


TABLE 9.1 Differences Between One-Tailed and Two-Tailed Tests

ONE-TAILED TESTS	TWO-TAILED TESTS
• Null hypothesis has an = sign	• Null hypothesis has an = sign
• Directional	• Nondirectional
• Cues in problem are *directional* words such as: higher, lower, older, younger, more, less, longer, shorter, etc.	• Cues in problem are *nondirectional* words such as: same, different, equal, not equal, control, out-of-control
• Alternative hypothesis has either > or < sign	• Alternative hypothesis can have only ≠ sign
• The direction (> or <) not included in the alternative hypotheses is assumed to be included in the null hypothesis—although not shown	• Null and alternative hypotheses include all possible cases
• Hypothesis possibilities: H_0: = H_a: > or H_0: = H_a: <	• Hypotheses are always written as: H_0: = H_a: ≠

Chi-Square Test of Independence

$$\chi^2 = \sum\sum \frac{(f_o - f_e)^2}{f_e} \quad (16.2)$$

$$df = (r-1)(c-1)$$

- The coefficient of determination tells you how much of the variability in one variable is explained by the other variable.
- If r = 0.5, then $r^2 = .25$
    - We would conclude that 25% of job performance is related to cognitive ability

P-Values

1. **P-values can indicate how incompatible the data are with a specified statistical model.**

2. **P-values do not measure the probability that the studied hypothesis is true, or the probability that the data were produced by random chance**

3. **Scientific conclusions and business or policy decisions should not be based solely on whether a p-value passes a specific threshold.**

4. **Proper inference requires full reporting and transparency.**

5. **A p-value, or statistical significance, does not measure the size of an effect or the importance of a result.**

6. **By itself, a p-value does not provide a good measure of evidence regarding a model or hypothesis.**

Ronald L. Wasserstein & Nicole A. Lazar (2016): The ASA's statement on p-values: context, process, and purpose, The American Statistician, DOI: 10.1080/00031305.2016.1154108

To link to this article: http://dx.doi.org/10.1080/00031305.2016.1154108

$$\chi^2 = \sum\sum\frac{(f_o - f_e)^2}{f_e} \quad (16.2)$$

$$df = (r-1)(c-1)$$

where:

r = number of rows
c = number of columns
f_o = frequency of observed values
f_e = frequency of expected values

Chi-Square Goodness-of-Fit Test

$$\chi^2 = \sum\frac{(f_o - f_e)^2}{f_e} \quad (16.1)$$

$$\text{df} = k - 1 - c$$

where

f_o = frequency of observed values
f_e = frequency of expected values
k = number of categories
c = number of parameters being estimated from the sample data