


Inferential Statistics and Probability
a Holistic Approach

Chapter 10
Two Population Inference



This Course Material by Maurice Geraghty is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License. Conditions for use are shown here: <https://creativecommons.org/licenses/by-sa/4.0/>

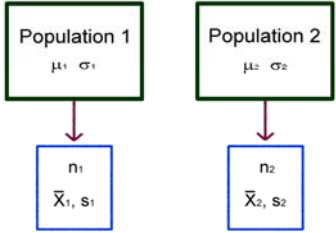
1

Comparing two population means

- Four models
 - Independent Sampling
 - Large Sample or known variances
 - **Z - test**
 - The 2 population variances are equal
 - **Pooled variance t-test**
 - The 2 population variances are unequal
 - **t-test for unequal variances**
 - Dependent Sampling
 - **Matched Pairs t-test**

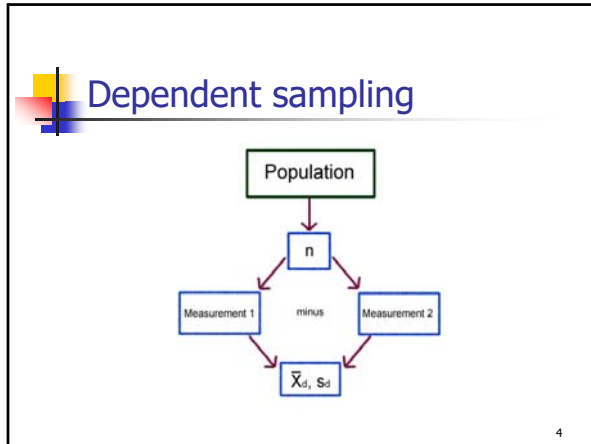
2

Independent Sampling



```
graph TD; P1[Population 1  
μ₁, σ₁] --> S1[n₁  
X̄₁, s₁]; P2[Population 2  
μ₂, σ₂] --> S2[n₂  
X̄₂, s₂];
```

3



Difference of Two Population means

- $\bar{X}_1 - \bar{X}_2$ is Random Variable
- $\bar{X}_1 - \bar{X}_2$ is a point estimator for $\mu_1 - \mu_2$
- The standard deviation is given by the formula $\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$
- If n_1 and n_2 are sufficiently large, $\bar{X}_1 - \bar{X}_2$ follows a normal distribution.


5

Difference between two means – large sample Z test

- If both n_1 and n_2 are over 30 and the two populations are independently selected, this test can be run.
- Test Statistic:

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$


6



Example 1

- Are larger houses more likely to have pools?
- The housing data square footage (size) was split into two groups by pool (Y/N).
- Test the hypothesis that the homes with pools have more square feet than the homes without pools. Let $\alpha = .01$

7



EXAMPLE 1 - Design


$$H_o : \mu_1 \leq \mu_2 \quad H_a : \mu_1 > \mu_2$$
$$H_o : \mu_1 - \mu_2 \leq 0 \quad H_a : \mu_1 - \mu_2 > 0$$

$\alpha = .01$

$$Z = (\bar{X}_1 - \bar{X}_2) / (\sqrt{\sigma_1^2 / n_1 + \sigma_2^2 / n_2})$$

H_0 is rejected if $Z > 2.326$

8



EXAMPLE 1 Data

■ Population 1 Size with pool	■ Population 2 Size without pool
■ Sample size = 130	■ Sample size = 95
■ Sample mean = 26.25	■ Sample mean = 23.04
■ Standard Dev = 6.93	■ Standard Dev = 4.55

9

EXAMPLE 1 DATA

$$Z = \frac{(26.25 - 23.04) - 0}{\sqrt{\frac{6.93^2}{130} + \frac{4.55^2}{95}}} = 4.19$$

- Decision: Reject Ho
- Conclusion: Homes with pools have more mean square footage.

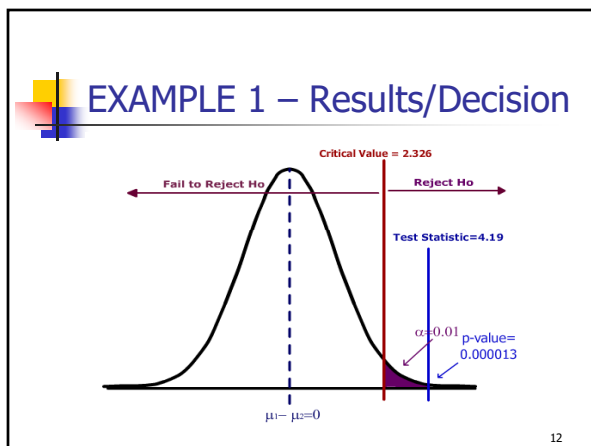
10


EXAMPLE 1 *p-value method*

- Using Technology
Reject Ho if the p-value < α .

	Sq ft with pool	Sq ft no pool
Mean	26.25	23.04
Std Dev	6.93	4.55
Observations	130	95
Hypothesized Mean Difference	0	
Z	4.19	
p-value	0.0000137	

11






Pooled variance t-test

- To conduct this test, three assumptions are required:
 - The populations must be normally or approximately normally distributed (or central limit theorem must apply).
 - The sampling of populations must be **independent**.
 - The **population variances** must be **equal**.

13




Pooled Sample Variance and Test Statistic

- Pooled Sample Variance:
$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$
- Test Statistic:
$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$df = n_1 + n_2 - 2$$

14



EXAMPLE 2

- A recent EPA study compared the highway fuel economy of domestic and imported passenger cars.
 - A sample of 12 imported cars revealed a mean of 35.76 mpg with a standard deviation of 3.86.
 - A sample of 15 domestic cars revealed a mean of 33.59 mpg with a standard deviation of 2.16 mpg.
 - At the .05 significance level can the EPA conclude that the mpg is higher on the imported cars? *(Let subscript 2 be associated with domestic cars.)*

15

EXAMPLE 2

- $H_o : \mu_1 \leq \mu_2 \quad H_a : \mu_1 > \mu_2$
- $\alpha = .05$
- $t = (\bar{X}_1 - \bar{X}_2) / (s_p \sqrt{1/n_1 + 1/n_2})$
- H_0 is rejected if $t > 1.708$, $df = 25$
- $t = 1.85$ H_0 is rejected. Imports have a higher mean mpg than domestic cars.

16

t-test when variances are not equal.


- Test statistic:
$$t' = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$
- Degrees of freedom:
$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\left[\frac{(s_1^2/n_1)^2}{(n_1-1)} + \frac{(s_2^2/n_2)^2}{(n_2-1)}\right]}$$
- This test (also known as the Welch-Aspin Test) has **less power** than the prior test and should only be used when it is clear the population variances are different.

17

EXAMPLE 2

- $H_o : \mu_1 \leq \mu_2 \quad H_a : \mu_1 > \mu_2$
- $\alpha = .05$
- t' test
- H_0 is rejected if $t > 1.746$, $df = 16$
- $t' = 1.74$ H_0 is not rejected. There is insufficient sample evidence to claim a higher mpg on the imported cars.

18




Using Technology

- Decision Rule: Reject H_0 if $p\text{-value} < \alpha$
- Megastat: Compare Two Independent Groups
- Use Equal Variance or Unequal Variance Test
- Use Original Data or Summarized Data

domestic 29.8 33.3 34.7 37.4 34.4 32.7 30.2 36.2 35.5 34.6 33.2 35.1 33.6 31.3 31.9

import 39.0 35.1 39.1 32.2 35.6 35.5 40.8 34.7 33.2 29.4 42.3 32.2

19




Megastat Result – Equal Variances

import	domestic	
35.76	33.59	mean
3.86	2.16	std. dev.
12	15	n

25 df
 2.17000 difference (import - domestic)
 9.16856 pooled variance
 3.02796 pooled std. dev.
 1.17273 standard error of difference
 0 hypothesized difference

1.85 t
 .0381 p-value (one-tailed, upper)

20




Megastat Result – Unequal Variances

import	domestic	
35.76	33.59	mean
3.86	2.16	std. dev.
12	15	n

16 df
 2.17000 difference (import - domestic)
 1.24606 standard error of difference
 0 hypothesized difference

1.74 t
 .0504 p-value (one-tailed, upper)


21



Hypothesis Testing - Paired Observations

- **Independent samples** are samples that are not related in any way.
- **Dependent samples** are samples that are paired or related in some fashion.
 - For example, if you wished to buy a car you would look at the *same* car at two (or more) *different* dealerships and compare the prices.
- Use the following test when the samples are **dependent**:

22




Hypothesis Testing Involving Paired Observations

$$t = \frac{\bar{X}_d - \mu_d}{s_d / \sqrt{n}}$$

- where \bar{X}_d is the average of the differences
- s_d is the standard deviation of the differences
- n is the number of pairs (differences)

23



EXAMPLE 3

- An independent testing agency is comparing the daily rental cost for renting a compact car from Hertz and Avis.
- A random sample of 15 cities is obtained and the following rental information obtained.
- At the .05 significance level can the testing agency conclude that there is a difference in the rental charged?

24

Example 3 – continued

- Data for Hertz
 - $\bar{X}_1 = 46.67$
 - $s_1 = 5.23$
- Data for Avis
 - $\bar{X}_2 = 44.87$
 - $s_2 = 5.62$

City	Hertz	Avis
Atlanta	42	40
Baltimore	51	47
Boston	46	42
Chicago	56	52
Cleveland	45	43
Denver	48	48
Dallas	56	54
Honolulu	37	32
Los Angeles	51	48
Kansas City	45	48
Miami	41	39
New York	44	42
San Francisco	48	45
Seattle	46	50
Washington DC	44	43

25

Example 3 - continued

By taking the difference of each pair, variability (measured by standard deviation) is reduced.

- $\bar{X}_d = 1.80$
- $s_d = 2.513$
- $n = 15$

City	Hertz	Avis	Difference
Atlanta	42	40	2
Baltimore	51	47	4
Boston	46	42	4
Chicago	56	52	4
Cleveland	45	43	2
Denver	48	48	0
Dallas	56	54	2
Honolulu	37	32	5
Los Angeles	51	48	3
Kansas City	45	48	-3
Miami	41	39	2
New York	44	42	2
San Francisco	48	45	3
Seattle	46	50	-4
Washington DC	44	43	1

26

EXAMPLE 3 *continued*

- $H_0: \mu_d = 0$ $H_1: \mu_d \neq 0$
- $\alpha = .05$
- Matched pairs t test, $df = 14$
- H_0 is rejected if $t < -2.145$ or $t > 2.145$
- $t = (1.80) / [2.513 / \sqrt{15}] = 2.77$
- Reject H_0 .
- There is a difference in mean price for compact cars between Hertz and Avis. Avis has lower mean prices.

27

Megastat Output – Example 3
 Hypothesis Test: Paired Observations

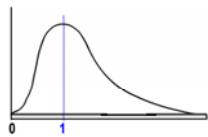
0.000 hypothesized value
 46.667 mean Hertz
 44.867 mean Avis
 1.800 mean difference (Hertz - Avis)
 2.513 std. dev.
 0.649 std. error
 15 n
 14 df

2.77 t
 .0149 p-value (two-tailed)

28

Characteristics of F-Distribution

- There is a "family" of F Distributions.
- Each member of the family is determined by two parameters: the numerator degrees of freedom and the denominator degrees of freedom.
- F cannot be negative, and it is a continuous distribution.
- The F distribution is positively skewed.
- Its values range from 0 to ∞ . As $F \rightarrow \infty$ the curve approaches the X -axis.




29

Test for Equal Variances

- For the two tail test, the test statistic is given by:

$$F = \frac{S_i^2}{S_j^2}$$
- s_i^2 and s_j^2 are the sample variances for the two populations.
- There are 2 sets of degrees of freedom: n_i-1 for the numerator, n_j-1 for the denominator


30



EXAMPLE 4

- A stockbroker at brokerage firm, reported that the mean rate of return on a sample of 10 software stocks was 12.6 percent with a standard deviation of 4.9 percent.
- The mean rate of return on a sample of 8 utility stocks was 10.9 percent with a standard deviation of 3.5 percent.
- At the .05 significance level, can the broker conclude that there is more variation in the software stocks?


31



Test Statistic depends on Hypotheses

Hypotheses	Test Statistic
$H_o : \sigma_1 \geq \sigma_2$ $H_a : \sigma_1 < \sigma_2$	$F = \frac{s_2^2}{s_1^2}$ use α table
$H_o : \sigma_1 \leq \sigma_2$ $H_a : \sigma_1 > \sigma_2$	$F = \frac{s_1^2}{s_2^2}$ use α table
$H_o : \sigma_1 = \sigma_2$ $H_a : \sigma_1 \neq \sigma_2$	$F = \frac{\max(s_1^2, s_2^2)}{\min(s_1^2, s_2^2)}$ use $\alpha / 2$ table

32



EXAMPLE 4 *continued*

- $H_o : \sigma_1 \leq \sigma_2$ $H_a : \sigma_1 > \sigma_2$
- $\alpha = .05$
- F-test
- H_0 is rejected if $F > 3.68$, $df=(9,7)$
- $F = 4.9^2 / 3.5^2 = 1.96 \rightarrow$ Fail to Reject H_0 .
- There is insufficient evidence to claim more variation in the software stock.

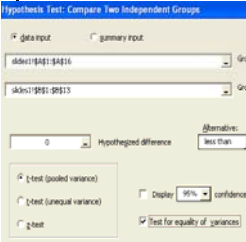
33

Excel Example

- Using Megastat – Test for equal variances under two population independent samples test and click the box to test for equality of variances
- The default p-value is a two-tailed test, so take one-half reported p-value for one-tailed tests
- Example – Domestic vs Import Data
- $H_o : \sigma_1 = \sigma_2$ $H_a : \sigma_1 \neq \sigma_2$
- $\alpha = .10$
- Reject H_o means use unequal variance t-test
- FTR H_o means use pooled variance t-test

34

Excel Output



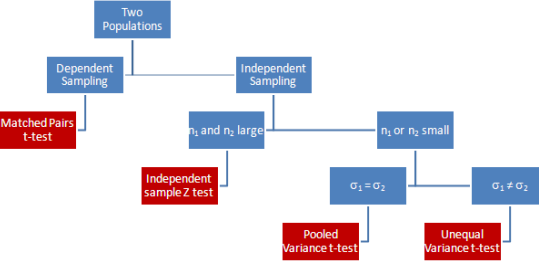
F-test for equality of variance
 14.894 variance: import
 4.654 variance: domestic
 3.20 F
 .0438 p-value

pvalue < .10, Reject H_o

Use unequal variance t-test to compare means.

35


Compare Two Means Flowchart



```

    graph TD
      A[Two Populations] --> B[Dependent Sampling]
      A --> C[Independent Sampling]
      B --> D[Matched Pairs t-test]
      C --> E[n1 and n2 large]
      C --> F[n1 or n2 small]
      E --> G[Independent sample Z test]
      F --> H[sigma1 = sigma2]
      F --> I[sigma1 != sigma2]
      H --> J[Pooled Variance t-test]
      I --> K[Unequal Variance t-test]
    
```

36




Comparing two proportions

- Suppose we take a sample of n_1 from population 1 and n_2 from population 2.
- Let X_1 be the number of success in sample 1 and X_2 be the number of success in sample 2.
- The sample proportions are then calculated for each group.

$$\hat{p}_1 = \frac{X_1}{n_1} \quad \hat{p}_2 = \frac{X_2}{n_2}$$

37




Hypothesis testing for 2 Proportions

- In conducting a Hypothesis test where the Null hypothesis assumes equal proportions, it is best practice to pool or combine the sample proportions into a single estimated proportion, and use an estimated standard error.

$$\bar{p} = \frac{X_1 + X_2}{n_1 + n_2} \quad s_{\hat{p}_1 - \hat{p}_2} = \sqrt{\frac{\bar{p}(1-\bar{p})}{n_1} + \frac{\bar{p}(1-\bar{p})}{n_2}}$$

38




Hypothesis testing for 2 Proportions

- The test statistic will have a Normal Distribution as long as there are at least 10 successes and 10 failures in both samples.

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\bar{p}(1-\bar{p})}{n_1} + \frac{\bar{p}(1-\bar{p})}{n_2}}}$$


39



Example

- In an August 2016 Study, Pew Research asked the sampled Americans if background checks required at gun stores should be made universal extended to all sales of guns between private owners or at gun shows.
- 772 out 990 men said yes, while 857 out of 1020 women said yes.
- Is there a difference in the proportion of men and women who support universal background checks for purchasing guns? Design and conduct the test with a significance level of 1%.


40



Example (Design)

- Ho:** $p_m = p_w$ (There is no difference in the proportion of support for background checks by gender)
- Ha:** $p_m \neq p_w$ (There is a difference in the proportion of support for background checks by gender)
- Model:**
Two proportion Z test. This is a two-tailed test with $\alpha = 0.01$.
- Model Assumptions:** for men there are 772 yes and 218 no. For women there are 857 yes and 16 no. Since all these numbers exceed 10, the model is appropriate.
- Decision Rules:**
Critical Value Method - Reject Ho if $Z > 2.58$ or $Z < -2.58$.
P-value method - Reject Ho if p-value < 0.01

41



Example (Results)

$$\hat{p}_m = \frac{772}{990} = 0.780 \quad \hat{p}_w = \frac{857}{1020} = 0.840$$

$$\bar{p} = \frac{772+857}{990+1020} = 0.810 \quad Z = \frac{(0.780 - 0.840) - 0}{\sqrt{\frac{0.810(1-0.810)}{990} + \frac{0.810(1-0.810)}{1020}}} = -3.45$$

p-value = 0.0005 $< \alpha$
Reject Ho Under both methods.

Conclusion: There is a difference in the proportion of support for background checks by gender. Women are more likely to support background checks.

42
