Making Comparisons

Inferences Based on Two Samples:
Confidence Intervals & Tests of Hypotheses
PBAF 527 Winter 2005

1

Today

1. Scallops, Sampling and the Law
   - Confidence Intervals, Hypothesis Testing, and Sampling
2. Hypothesis Testing
   - Special Cases: Small samples, proportions
3. Making Comparisons
   - Solve Hypothesis Testing Problems for Two Populations
     - Mean
     - Proportion
   - Distinguish Independent & Related Populations
   - Create Confidence Intervals for the Differences

2

Scallops, Sampling, and the Law

- Read Case

a. Can a reliable estimate of the mean weight of all the scallops be obtained from a sample size of 18?
b. Do you see any flaws in the rule to confiscate a scallop catch if the sample mean weight is less than 1/36 of a pound?
c. Develop your own procedure for determining whether a ship is in violation of the weight restriction using the data provided.
d. Apply your procedure to the data provided.

3

Today

1. Scallops, Sampling and the Law
   - Confidence Intervals, Hypothesis Testing, and Sampling
2. Hypothesis Testing
   - Special Cases: Small samples, proportions
3. Making Comparisons
   - Solve Hypothesis Testing Problems for Two Populations
     - Mean
     - Proportion
   - Distinguish Independent & Related Populations
   - Create Confidence Intervals for the Differences

4

Hypothesis Testing When n is Small and s Unknown

Because the sample is small
- Cannot assume normality
- Cannot assume s is a good approximation for s

So, use t-distribution:

$$t = \frac{\bar{x} - m}{\frac{s}{\sqrt{n}}}$$ with n-1 degrees of freedom

5

Small Sample t-test Example 1 (1)

Most water treatment facilities monitor the quality of their drinking water on hourly basis. One variable monitored it is pH, which measures the degree of alkalinity or acidity in the water. A pH below 7.0 is acidic, one above 7.0 is alkaline, and a pH of 7.0 is neutral. One water treatment plant has a target pH of 8.5 (most try to maintain a slightly alkaline level). The mean and standard deviation of 1 hour's test results, based on 17 water samples at this plant are:

$\bar{x} = 8.24$ s=.16

Does this sample provide sufficient evidence that the mean pH level in the water differs from 8.5?

6

Small Sample t-test Example 1 (2)

1. Establish hypotheses
2. Set the decision rule for the test:

pick α
find t_α at n-1 df

3. Find test statistic $t = \frac{\bar{x} - m}{\frac{s}{\sqrt{n}}} = \frac{8.42 - 8.5}{\frac{.16}{\sqrt{17}}} = \frac{-.08}{.039} = -2.05$
4. Compare test statistic to critical value.

7

Small Sample t-test Example 2 (1)

A major car manufacturer wants to test a new engine to determine whether it meets new air-pollution standards. The mean emission μ for all engines of this type must be less than 20 parts per million of carbon. 10 engines are manufactured for testing purposes, and the emission level for each is determined. The mean and standard deviation for the tests are:

$\bar{x} = 17.17$ s=2.98

Do the data supply enough evidence to allow the manufacturer to conclude that this type of engine meets the pollution standard? Assume the manufacturer is willing to risk a Type I error with probability α=.01.

8

Small Sample t-test Example 2 (2)

1. Establish hypotheses
2. Set the decision rule for the test:

pick α
find t_α

3. Find test statistic $t = \frac{\bar{x} - m}{\frac{s}{\sqrt{n}}} = \frac{17.17 - 20}{\frac{2.98}{\sqrt{10}}} = -3.00$
4. Compare test statistic to critical value.

9

Large Sample Test for the Population Proportion

When the sample size is large (np and nq are greater than 5)

- Assume $\hat{p}$ is distributed normally with mean p and standard deviation $\sqrt{\frac{pq}{n}}$ where q=1-p
- Test statistic: $z = \frac{\hat{p} - p_0}{\sqrt{p_0 q_0 / n}}$
- 2- or 1-tailed tests

10

Large Sample Tests for Proportion Example (1)

In screening women for breast cancer, doctors use a method that fails to detect cancer in 20% of the women who actually have the disease. Suppose a new method has been developed that researchers hope will detect cancer more accurately. This new method was used to screen a random sample of 140 women known to have breast cancer. Of these, the new method failed to detect cancer in 12 women.

Does this sample provide evidence that the failure rate of the new method differs from the one currently in use?

11

Large Sample Tests for Proportion Example (2)

1. Establish hypotheses
2. Set the decision rule for the test:

pick α

find t_α

3. Find test statistic $z = \frac{\hat{p} - p_0}{\sqrt{p_0 q_0 / n}} = \frac{.086 - .2}{\sqrt{(.2)(.8)/140}} = \frac{-.114}{-.034} = -3.36$
4. Compare test statistic to critical value.

12

Today

1. Scallops, Sampling and the Law
   - Confidence Intervals, Hypothesis Testing, and Sampling
2. Hypothesis Testing
   - Special Cases: Small samples, proportions
3. Making Comparisons
   - Solve Hypothesis Testing Problems for Two Populations
     - Mean
     - Proportion
   - Distinguish Independent & Related Populations
   - Create Confidence Intervals for the Differences

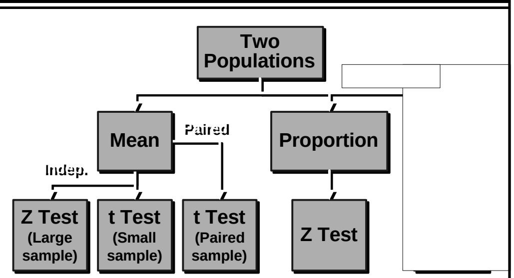
13

How Would You Try to Answer These Questions?

1. Do house prices in two Seattle neighborhoods differ?
   - By how much?
2. Does one method of teaching reading produce better results than another?
   - Can I still have a reliable result with a small sample size?
   - How much better are the results of the method?
3. Do energy conservation efforts really reduce consumption over time?
   - How much of a reduction?
4. Is the proportion of subprime mortgages to low-income households greater than that for moderate-income households?
   - How much greater?

14

Two Population Tests



15

Comparison of Means for Independent Subsamples

Three scenarios:

1. H_0: μ_1-μ_2=0 Ha: μ_1-$\mu_2 \neq 0$
2. H_0: μ_1-$\mu_2 \leq 0$ Ha: μ_1-μ_2>0
3. H_0: μ_1-$\mu_2 \leq$D Ha: μ_1-μ_2>D

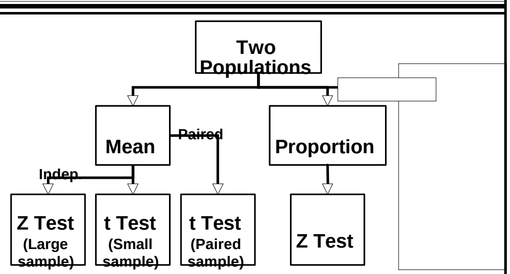
(not common, nor is 2-tailed test of D)

Could be:

- Separate (unequal) Variances (Large Samples)
- Equal Population Variances (Small Samples)

16

Two Population Tests



17

Comparison of Means Separate (Unequal) Variances for 2 Independent Subsamples

1. Assumptions
   - Independent, Random Samples
   - Populations Are Normally Distributed
   - If Not Normal, Can Be Approximated by Normal Distribution ($n_1 \geq 30$ & $n_2 \geq 30$)
   - For n's<30, use t with the smaller of n_1-1, n_2-1 df
2. Two Independent Sample Z-Test Statistic

$$Z = \frac{\left(\bar{X}_1 - \bar{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \approx \frac{\left(\bar{X}_1 - \bar{X}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

18

Example
Separate (unequal) Variances

Do house prices in two Seattle Neighborhoods differ? By how much?

Is the average price of a home of a certain size equal in Sandpoint and Ravenna?

You gather data on property value for a random sample of 32 properties in Sandpoint and find that =$345,650 and s=$48,500. Then you gather data on the value of a random sample of 35 properties in Ravenna and find that =$289,440 and s=$87,090. Is the average property value of all properties in both locations equal or not?

19

Separate (unequal) Variances Solution

H0:

Ha:

a =

***n*$_1$ = , *n*$_2$ =**

Critical Value(s) ? or ? /2?:

Test Statistic:

Decision:

Conclusion:

20

Separate (unequal) Variances Solution

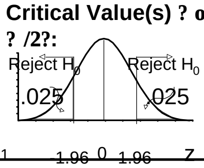
H0: $\mu_1 - \mu_2 = 0$ ($\mu_1 = \mu_2$)

Ha: $\mu_1 - \mu_2$? 0 (μ_1 ? μ_2)

a = **.05**

***n*$_1$ = 32 , *n*$_2$ = 35**

Critical Value(s) ? or ? /2?:



Test Statistic:

$$z = \frac{(\overline{x}_1 - \overline{x}_2) - (m_1 - m_2)_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(345{,}650 - 289{,}440) - (0)}{\sqrt{\frac{48{,}650^2}{32} + \frac{87{,}090^2}{35}}} = 3.3$$

Decision: $|z| > z_{?/2}$

Reject at a **= .05**

Conclusion:

There is Evidence of a Difference in Means

21

Confidence Interval for the Difference

Do house prices in two Seattle Neighborhoods differ?

By how much?

A (1-α)100% confidence interval for the difference between two population means μ_1-μ_2 using independent random sampling:

$$\left(\bar{x}_1 - \bar{x}_2\right) \pm z_{a/2}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$$

for n's<30 use $t_{\alpha/2}$ with the lesser of n_1-1, n_2-1 df

NB: are the variances of each of the two populations; when σ_1^2 and σ^2_2 are unknown, use s_1^2 and s_2^2.

22

Confidence Interval for the Difference

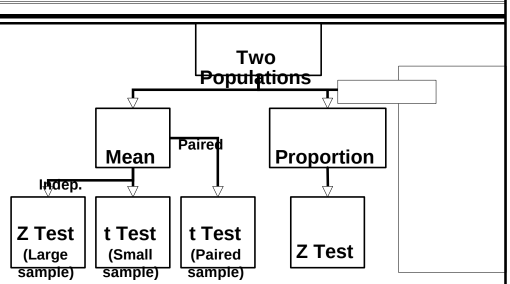
Do house prices in two Seattle Neighborhoods differ?

By how much?

Construct a 95% confidence interval around the difference and interpret it in words

23

Two Population Tests



24

Comparison of Means Equal Population Variances

1. Tests Means of 2 Independent Populations Having **Equal** Variances
2. Assumptions
   - Independent, Random Samples
   - Both Populations Are Normally Distributed
   - Population Variances Are **Unknown** But Assumed **Equal**
3. Usually small samples

25

Comparison of Means Equal Population Variances

We select two independent random samples:
from population 1 of size n_1 with mean $\bar{x}_1$ and variance s_1^2
from population 2 of size n_2 with mean $\bar{x}_2$ and variance s_2^2

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (m_1 - m_2)}{\sqrt{s_P^2 \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Estimate of Standard Error

Pooled Estimate of Variance

$$s_P^2 = \frac{(n_1 - 1) \cdot s_1^2 + (n_2 - 1) \cdot s_2^2}{n_1 + n_2 - 2}$$

$$df = n_1 + n_2 - 2$$

For large n's, use z

26

Comparison of Means Equal Population Variances

Example:

Does one method of teaching reading produce better results than another?

- Can I still have a reliable result with a small sample size?
- How much better are the results of the method?

27

Comparison of Means Equal Population Variances Example

Compare a new method of teaching reading to "slow learners" to the current standard method.

You decide to base this comparison on the results of a reading test given at the end of a learning period of 6 months.

Of a random sample of 22 slow learners, 10 are taught by the new method and 12 are taught by the standard method. Qualified instructors under similar conditions teach all 22 children for a 6-month period. The results of the reading test at the end of 6 months are as follows:

New Method: $\bar{x}_1$=76.4; s_1^2=34.04

Standard Method: $\bar{x}_2$=72.33; s_2^2=40.24

Are the reading scores of children using the new method greater than those of children using the standard method with alpha=.05?

28

Comparison of Means Equal Population Variances Solution

H_0:

H_a:

a =

df =

Critical Value(s):

Decision Rule:

Test Statistic:

Decision:

Conclusion:

29

Small-Sample t Test Solution

$10+12-2$

30

Comparison of Means Equal Population Variances Solution

H_0:

H_a:

a =

df =

Critical Value(s):

Decision Rule:

Test Statistic:

Decision:

Conclusion:

31

Confidence Interval for the Difference (equal population variances)

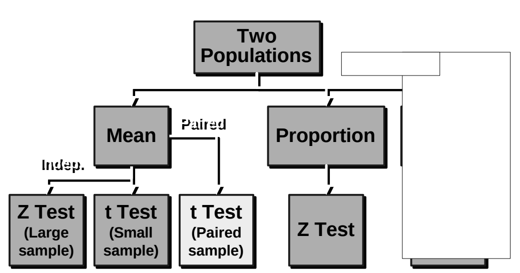
Does one method of teaching reading produce better results than another?

- How much better are the results of the method?

Construct a 95% confidence interval for the difference between the two means and interpret it.

32

Two Population Tests



33

Paired-Sample t Test for Mean Difference

Is a population parameter different over time or between groups?

1. Tests Means of 2 Related Populations
   - Paired or Matched
   - Repeated Measures (Before/After)
2. Eliminates Variation Among Subjects
3. Assumptions
   - Both Population Are Normally Distributed
   - If Not Normal, Can Be Approximated by Normal Distribution ($n_1 \geq 30$ & $n_2 \geq 30$)

34

Paired-Sample t Test Hypotheses

	Research Questions		
Hypothesis	No Difference Any Difference	Pop 1 ≥ Pop 2 Pop 1 < Pop 2	Pop 1 ≤ Pop 2 Pop 1 > Pop 2
H_0	$\mu_D = 0$	$\mu_D \geq 0$	$\mu_D \leq 0$
H_1	$\mu_D \neq 0$	$\mu_D < 0$	$\mu_D > 0$

Note: $D_i = X_{1i} - X_{2i}$ for ith observation

Be Careful!

35

Paired-Sample t Test Data Collection Table

Observation	Group 1	Group 2	Difference
1	x_{11}	x_{21}	$D_1 = x_{11}-x_{21}$
2	x_{12}	x_{22}	$D_2 = x_{12}-x_{22}$
:	:	:	:
i	x_{1i}	x_{2i}	$D_i = x_{1i} - x_{2i}$
:	:	:	:
n	x_{1n}	x_{2n}	$D_n = x_{1n} - x_{2n}$

36

Paired-Sample t Test
Test Statistic

$$z \text{ or } t = \frac{\bar{x}_D - D_0}{\frac{s_D}{\sqrt{n_D}}} \qquad df = n_D - 1$$

When n>30,use z
When n<30 use t(n-1) df

D_0=0 when testing whether there is any difference or not.

Sample Mean

$$\bar{x}_D = \frac{\sum_{i=1}^{n} D_i}{n_D}$$

Sample Standard Deviation

$$s_D = \sqrt{\frac{\sum_{i=1}^{n} (D_i - \bar{x}_D)^2}{n_D - 1}}$$

37

Paired-Sample t Test
Example

Do energy conservation efforts really reduce consumption over time? By how much?

A study is undertaken to determine how consumers react to energy conservation efforts. A random group of 60 families is chosen. Each family's rate of consumption of electricity is monitored in equal length time periods before and after they are offered financial incentives to reduce their energy consumption rate. The difference in electric consumption between the periods is recorded for each family. The average reduction in consumption is 0.2 kW and the standard deviation of the differences s_D=1.0 kW. At α=0.01, is there evidence to conclude that the incentives reduce consumption?

38

Paired-Sample t Test
Solution

H_0:

H_a:

α =

Decision Rule:

Critical Value(s):

-2.326 0 Z

Test Statistic:

n_D 60

Decision:

Conclusion:

39

Confidence Interval for Paired Observations

Do energy conservation efforts really reduce consumption over time? By how much?

A (1-α)100% confidence interval for the mean difference is constructed using the t distribution for small sample sizes and z distribution for large sample sizes.

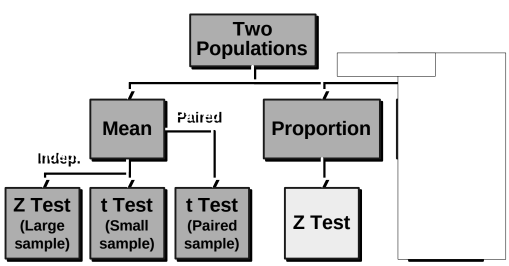
$$\bar{X}_D \pm z_{\alpha/2} \frac{S_D}{\sqrt{n}}$$

When n>30,use z
When n<30 use t(n-1) df

40

Portions of these notes are adapted from Statistics 7e © 1997 Prentice-Hall, Inc.

Two Population Tests



41

Portions of these notes are adapted from Statistics 7e © 1997 Prentice-Hall, Inc.

Z Test for Difference in Two Proportions

1. Can test
   - H_0: p_1-p_2=0 H_a: p_1-p_2≠0
   - H_0: p_1-p_2≤0 H_a: p_1-p_2>0
2. Assumptions
   - Populations Are Independent
   - Normal Approximation Can Be Used
     - $n\hat{p} \pm 3\sqrt{n\hat{p}(1-\hat{p})}$ Does Not Contain 0 or n
3. Z-Test Statistic for Two Proportions

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\hat{p}\cdot(1-\hat{p})\cdot\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \quad \text{where } \hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

42

Z Test for Difference in Two Proportions

Is the proportion of sub-prime mortgages to low-income households greater than than for moderate-income households? How much greater?

In 1998, a sample of mortgages was taken from the over 1 million mortgages disclosed nationally under HMDA. Here is a decription of the sample:

Income Group	Percent Subprime	n
Low-income	26%	400
Moderate-income	11%	600

Is there sufficient evidence to claim that the proportion of sub-prime mortgages to low-income households exceeds that among moderate income households? Test using ?=.01

43

Z Test for Two Proportions Solution

H0:

Ha:

a =

n_1 = n_2 =

Decision Rule:

Critical Value(s):

Test Statistic:

Decision:

Conclusion:

44

Z Test for Two Proportions Solution

45

Z Test for Two Proportions Solution

H_0:

H_a:

a =

n_1 = n_2 =

Decision Rule:

Critical Value(s):

Test Statistic:

Decision:

Conclusion:

46

Confidence Interval for the Difference in Proportions

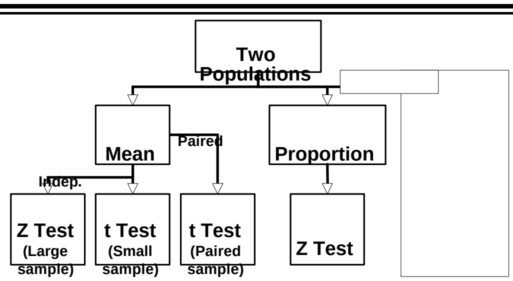
Is the proportion of sub-prime mortgages to low-income households greater than than for moderate-income households? How much greater?

A large sample (1-α)100% confidence interval for the difference between two population proportions:

How much greater is the proportion of subprime mortgages to low-income buyers compared to moderate-income buyers?

47

Two Population Tests



48

Today

1. Scallops, Sampling and the Law
   - Confidence Intervals, Hypothesis Testing, and Sampling
2. Making Comparisons
   - Solve Hypothesis Testing Problems for Two Populations
     - Mean
     - Proportion
   - Distinguish Independent & Related Populations
   - Create Confidence Intervals for the Differences

49

End of Chapter

Any blank slides that follow are blank intentionally.