HYPOTHESIS TESTING (TESTS FOR TWO SAMPLES)

Anthony S. Morfe_CAS_Math Cluster

OBJECTIVES:

At the end of this presentation, students should be able to:

- Test the difference between sample means, using the *z* test.
- Test the difference between two means for independent samples, using the *t* test.
- Test the difference between two means for dependent samples, using the *t* test.
- Test the difference between two proportions.
- Test the difference between two variances or standard deviations.

TESTING THE DIFFERENCE BETWEEN TWO MEANS

Anthony S. Morfe_CAS_Math Cluster

TESTING THE DIFFERENCE BETWEEN TWO MEANS: USING THE Z-TEST

When to use?

If researchers wish to compare two sample means, using experimental and control groups or two experimental groups.

Examples:

Two different brands of fertilizer might be tested to see whether one is better than the other for growing plants.

Two brands of cough syrup might be tested to see whether one brand is more effective than the other.

TESTING THE DIFFERENCE BETWEEN TWO MEANS: USING THE Z-TEST

Suppose a researcher wishes to determine whether there is a difference in the average age of nursing students who enroll in a nursing program at a community college and those who enroll in a nursing program at a university. His research question is, Does the mean age of nursing students who enroll at a community college differ from the mean age of nursing students who enroll at a university? Here, the hypotheses are:

$$H_o: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

TESTING THE DIFFERENCE BETWEEN TWO MEANS: USING THE Z-TEST

Assumptions for the *z* Test to Determine the Difference Between Two Means

1. Both samples are random samples.
2. The samples must be independent of each other. That is, there can be no relationship between the subjects in each sample.
3. The standard deviations of both populations must be known, and if the sample sizes are less than 30, the populations must be normally or approximately normally distributed.

TESTING THE DIFFERENCE BETWEEN TWO MEANS: USING THE Z-TEST

Formula for the z Test for Comparing Two Means from Independent Populations

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

TESTING THE DIFFERENCE BETWEEN TWO MEANS: USING THE Z-TEST –EXAMPLE 1

A survey found that the average hotel room rate in New Orleans is \$88.42 and the average room rate in Phoenix is \$80.61. Assume that the data were obtained from two samples of 50 hotels each and that the standard deviations of the populations are \$5.62 and \$4.83, respectively. At $\alpha=0.05$, can it be concluded that there is a significant difference in the rates?

TESTING THE DIFFERENCE BETWEEN TWO MEANS: USING THE Z-TEST –EXAMPLE 1

Two-tailed test	Right-tailed test	Left-tailed test
H_0: $\mu_1=\mu_2$	H_0: $\mu_1\leq\mu_2$	H_0: $\mu_1\geq\mu_2$
H_1: $\mu_1\neq\mu_2$	H_1: $\mu_1>\mu_2$	H_1: $\mu_1<\mu_2$

the populations are \$5.62 and \$4.83, respectively. At $\alpha=0.05$, can it be concluded that there is a significant difference in the rates?

Values needed: $\alpha=0.05$

Group 1 (New Orleans) – $\bar{x}_1=88.42$, $n_1=50$, $\sigma_1=5.62$

Group 2 (Phoenix) – $\bar{x}_2=80.61$, $n_2=50$, $\sigma_2=4.83$

TESTING THE DIFFERENCE BETWEEN TWO MEANS: USING THE Z-TEST –EXAMPLE 1

Solution

Step 1 State the hypotheses and identify the claim.

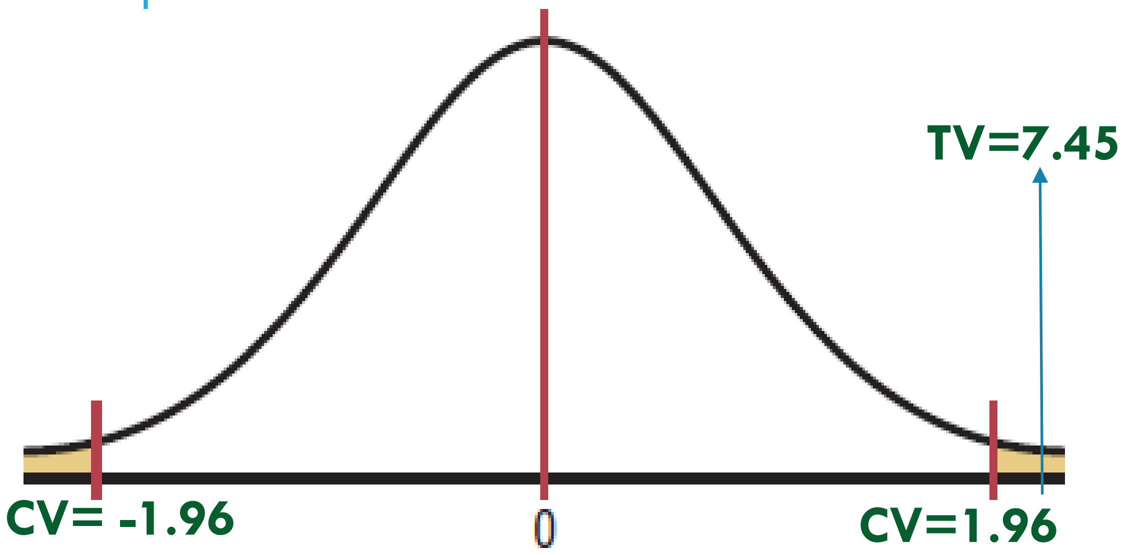
H_0: $\mu_1 = \mu_2$ and H_1: $\mu_1 \neq \mu_2$ (claim) Two-tailed test

Step 2 Find the critical values. Since $\alpha = 0.05$, the critical values are $+1.96$ and -1.96.

Step 3 Compute the test value.

$$z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} = \frac{(88.42 - 80.61) - 0}{\sqrt{\dfrac{5.62^2}{50} + \dfrac{4.83^2}{50}}} = 7.45$$

TESTING THE DIFFERENCE BETWEEN TWO MEANS: USING THE Z-TEST –EXAMPLE 1



Step 4: Make the decision.
H_o: Reject and H_1: Do not reject (claim) since the test value falls in the critical region.

- **Step 5:** Summarize the results. The evidence is enough to support the claim that the means are not equal. Hence, there is a significant difference in the rates.

INDEPENDENT VS DEPENDENT SAMPLE GROUPS: USING T-TEST

Independent Sample Groups:

Independent groups, also called unpaired groups or unrelated groups, are groups in which the cases (e.g., participants) in each group are different. Often we are investigating differences in individuals, which means that when comparing two groups, an individual in one group cannot also be a member of the other group and vice versa.

Dependent Sample Groups:

Dependent sample groups, also called related groups have that same participants present in both groups. This indicates that the same participants are tested more than once. The reason that it is possible to have the same participants in each group is because each participant has been measured on the same dependent variable under two different conditions.

TESTING THE DIFFERENCE BETWEEN TWO MEANS OF INDEPENDENT SAMPLES: USING THE T-TEST

Assumptions:

- Population standard deviations are unknown and the sample sizes are less than 30.
- Variables are normally or approximately normally distributed.
- Samples are random samples and **independent samples.**

TESTING THE DIFFERENCE BETWEEN TWO MEANS OF INDEPENDENT SAMPLES: USING THE T-TEST

Formula for the t Test—For Testing the Difference Between Two Means—Independent Samples

Variances are assumed to be unequal

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$$

where the degrees of freedom are equal to the smaller of $n_1 - 1$ or $n_2 - 1$.

TESTING THE DIFFERENCE BETWEEN TWO MEANS OF INDEPENDENT SAMPLES: USING THE T-TEST

When the variances are assumed to be equal, this formula is used

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$df = n_1 + n_2 - 2$$

TESTING THE DIFFERENCE BETWEEN TWO MEANS OF INDEPENDENT SAMPLES: USING THE T-TEST

Example

The average size of a farm in Indiana County, Pennsylvania, is 191 acres. The average size of a farm in Greene County, Pennsylvania, is 199 acres. Assume the data were obtained from two samples with standard deviations of 38 and 12 acres, respectively, and sample sizes of 8 and 10, respectively. Can it be concluded at $\alpha=0.05$ that the average size of the farms in the two counties is different? Assume the populations are normally distributed and the population variances are not equal.

TESTING THE DIFFERENCE BETWEEN TWO MEANS OF INDEPENDENT SAMPLES: USING THE T-TEST

Two-tailed test	Right-tailed test	Left-tailed test
H_0: $\mu_1 = \mu_2$	H_0: $\mu_1 \leq \mu_2$	H_0: $\mu_1 \geq \mu_2$
H_1: $\mu_1 \neq \mu_2$	H_1: $\mu_1 > \mu_2$	H_1: $\mu_1 < \mu_2$

two samples with standard deviations of 38 and 12 acres, respectively, and sample sizes of 8 and 10, respectively. Can it be concluded at $\alpha=0.05$ that the average size of the farms in the two counties is different? Assume the populations are normally distributed and the population variances are not equal.

Values needed: $\alpha=0.05$

Group 1 (Indiana) – $\bar{x}_1=191$, $n_1=8$, $df_1=7$, $s_1=38$

Group 2 (Greene) – $\bar{x}_2=199$, $n_2=10$, $df_2=9$, $s_2=12$

TESTING THE DIFFERENCE BETWEEN TWO MEANS OF INDEPENDENT SAMPLES: USING THE T-TEST

Solution

Step 1 State the hypotheses and identify the claim for the means.

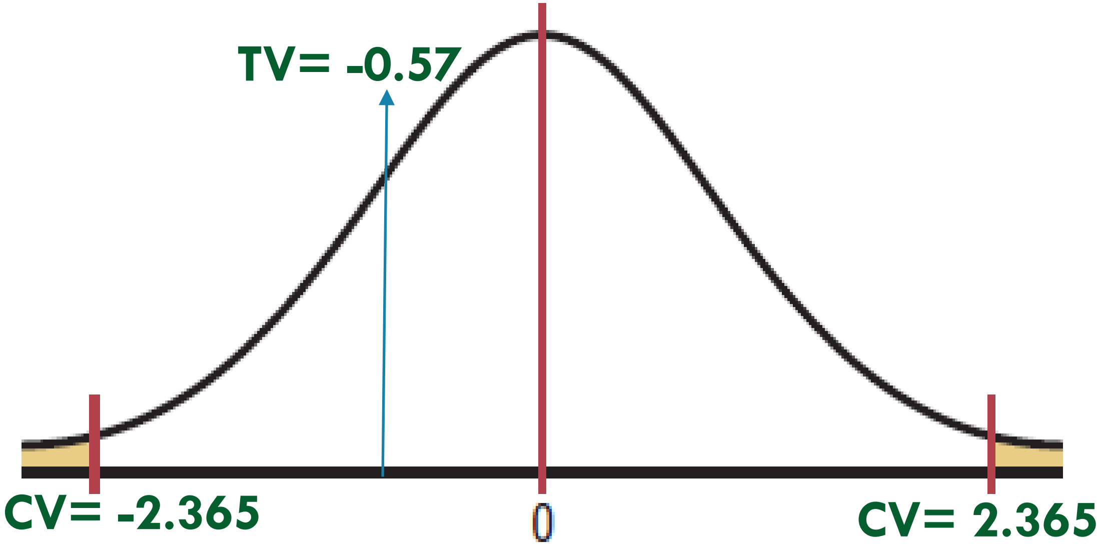
$H_0: \mu_1 = \mu_2$ and $H_1: \mu_1 \neq \mu_2$ (claim) Two-tailed test

Step 2 Find the critical values. Since the test is two-tailed, since $\alpha = 0.05$, and since the variances are unequal, the degrees of freedom are the smaller of $n_1 - 1$ or $n_2 - 1$. In this case, the degrees of freedom are $8 - 1 = 7$. Hence, from Table F, the critical values are $+2.365$ and -2.365.

Step 3 Compute the test value. Since the variances are unequal, use the first formula.

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{(191 - 199) - 0}{\sqrt{\frac{38^2}{8} + \frac{12^2}{10}}} = -0.57$$

TESTING THE DIFFERENCE BETWEEN TWO MEANS OF INDEPENDENT SAMPLES: USING THE T-TEST



Step 4: Make the decision.
H_o: Do not reject and H_1: Reject (claim) since the test value is not in the critical region.

- **Step 5:** Summarize the results. The evidence is not enough to support the claim that the average size of the farms is different.

TESTING THE DIFFERENCE BETWEEN TWO MEANS OF DEPENDENT SAMPLES: USING THE T-TEST

Samples are considered to be **dependent** when the subjects are paired or matched in some way.

Examples of paired data:

- Each person is measured twice where the 2 measurements measure the same thing but under different conditions
- Similar individuals are paired prior to an experiment and each member of a pair receives a different treatment
- Two different variables are measured for each individual and there is interest in the amount of difference between the 2 variables

TESTING THE DIFFERENCE BETWEEN TWO MEANS OF DEPENDENT SAMPLES: USING THE T-TEST

Assumptions for the t Test for Two Means When the Samples Are Dependent

1. The sample or samples are random.
2. The sample data are dependent.
3. When the sample size or sample sizes are less than 30, the population or populations must be normally or approximately normally distributed.

TESTING THE DIFFERENCE BETWEEN TWO MEANS OF DEPENDENT SAMPLES: USING THE T-TEST

Formula for the t-test for Dependent Samples

$$t = \frac{\bar{D} - \mu_D}{\frac{s_D}{\sqrt{n}}}$$

$\bar{D}$ = sample mean of the differences (D) of the values of the pairs of data

μ_D = difference of the population means

s_D = sample standard deviation of the differences (D) of the values of the pairs of data

n = number of data pairs

Anthony S. Morfe_CAS_Math Cluster

TESTING THE DIFFERENCE BETWEEN TWO MEANS OF DEPENDENT SAMPLES: USING THE T-TEST

Example

A sample of nine local banks shows their deposits (in billions of dollars) 3 years ago and their deposits (in billions of dollars) today. At α=0.05, can it be concluded that the average in deposits for the banks is greater today than it was 3 years ago?

Bank	1	2	3	4	5	6	7	8	9
3 years ago (x1)	11.42	8.41	3.98	7.37	2.28	1.1	1	0.9	1.35
Today (x2)	16.69	9.44	6.53	5.58	2.92	1.88	1.78	1.5	1.22

TESTING THE DIFFERENCE BETWEEN TWO MEANS OF DEPENDENT SAMPLES: USING THE T-TEST

Two-tailed test	Right-tailed test	Left-tailed test
H_0: $\mu_1 = \mu_2$	H_0: $\mu_1 \leq \mu_2$	H_0: $\mu_1 \geq \mu_2$
H_1: $\mu_1 \neq \mu_2$	H_1: $\mu_1 > \mu_2$	H_1: $\mu_1 < \mu_2$

At $\alpha = 0.05$, can it be concluded that the average in deposits for the banks is greater today than it was 3 years ago?

Bank	1	2	3	4	5	6	7	8	9
3 years ago (x1)	11.42	8.41	3.98	7.37	2.28	1.1	1	0.9	1.35
Today (x2)	16.69	9.44	6.53	5.58	2.92	1.88	1.78	1.5	1.22

$\mu_{1\ (3\ \text{years ago})} < \mu_{2\ (\text{today})}$

Values Needed: $\alpha = 0.05$, $n = 9$, $df = 8$

Group 1 (3 years ago) and Group 2 (Today)

TESTING THE DIFFERENCE BETWEEN TWO MEANS OF DEPENDENT SAMPLES: USING THE T-TEST

Step 1: State the hypotheses and identify the claim.

H_0: $\mu_1 \geq \mu_2$

H_1: $\mu_1 < \mu_2$ (claim) Left-tailed test

Step 2: Find the critical value. Since $\alpha = 0.05$, df = n-1 = 9-1 = 8 and the test is a left-tailed test, the critical value is –1.860.

TESTING THE DIFFERENCE BETWEEN TWO MEANS OF DEPENDENT SAMPLES: USING THE T-TEST

Step 3: Compute the test value.

Find differences (D) of the values of the pairs of data.

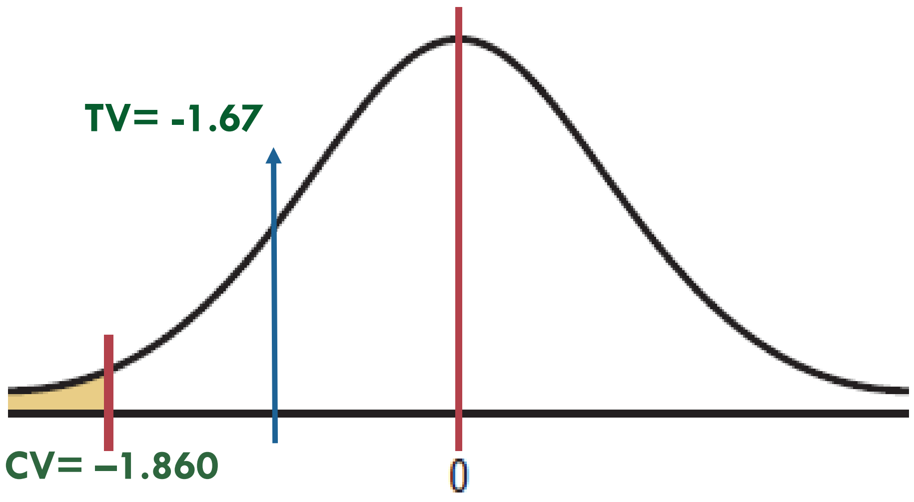
Bank	1	2	3	4	5	6	7	8	9
3 years ago (x1)	11.42	8.41	3.98	7.37	2.28	1.1	1	0.9	1.35
Today (x2)	16.69	9.44	6.53	5.58	2.92	1.88	1.78	1.5	1.22
D = x1 - x2	-5.27	-1.03	-2.55	1.79	-0.64	-0.78	-0.78	-0.6	0.13

Use the shortcut in your calculator to compute for the mean of D ($\overline{D}$) and standard deviation of D (s_D).

$\overline{D} = -1.081111111$ and $s_D=1.937281887$

$$t = [\overline{D} - \mu_D] / [s_D/\sqrt{n}] = [-1.081111111 - 0] / [1.937281887/\sqrt{9}] = -1.67.$$

TESTING THE DIFFERENCE BETWEEN TWO MEANS OF DEPENDENT SAMPLES: USING THE T-TEST



Step 4: Make the decision.

H_o: Do not reject and H_1: Reject (claim) since the test value is not in the critical region.

- Step 5: Summarize the results. The evidence is not enough to support the claim that the average in deposits for the banks is greater today than it was 3 years ago.

TESTING THE DIFFERENCE BETWEEN TWO PROPORTIONS

Anthony S. Morfe_CAS_Math Cluster

TESTING THE DIFFERENCE BETWEEN PROPORTIONS

The ***z* test** with some modifications can be used to test the equality of two proportions.

For example, a researcher might ask, Is the proportion of men who exercise regularly less than the proportion of women who exercise regularly? Is there a difference in the percentage of students who own a personal computer and the percentage of nonstudents who own one?

TESTING THE DIFFERENCE BETWEEN PROPORTIONS

Formula for the z Test for Comparing Two Proportions

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\bar{p}\,\bar{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where

$$\bar{p} = \frac{X_1 + X_2}{n_1 + n_2} \qquad \hat{p}_1 = \frac{X_1}{n_1}$$

$$\bar{q} = 1 - \bar{p} \qquad \hat{p}_2 = \frac{X_2}{n_2}$$

TESTING THE DIFFERENCE BETWEEN PROPORTIONS

Assumptions for the *z* Test for Two Proportions

1. The samples must be random samples.
2. The sample data are independent of one another.
3. For both samples $np \geq 5$ and $nq \geq 5$.

TESTING THE DIFFERENCE BETWEEN PROPORTIONS-EXAMPLE 1

In a nursing home study, the researchers found that 12 out of 34 small nursing homes had a resident vaccination rate of less than 80%, while 17 out of 24 large nursing homes had a vaccination rate of less than 80%. At $\alpha=0.05$, test the claim that there is no difference in the proportions of the small and large nursing homes with a resident vaccination rate of less than 80%.

TESTING THE DIFFERENCE BETWEEN

Two-tailed test	Right-tailed test	Left-tailed test
H_0: $p_1 = p_2$	H_0: $p_1 \leq p_2$	H_0: $p_1 \geq p_2$
H_1: $p_1 \neq p_2$	H_1: $p_1 > p_2$	H_1: $p_1 < p_2$

80%. At $\alpha = 0.05$, test the claim that there is no difference in the proportions of the small and large nursing homes with a resident vaccination rate of less than 80%.

Values Needed:

Group 1(small nursing homes) – $X_1 = 12$, $n_1 = 34$, $\hat{p}_1 = 12/34$

Group 2(large nursing homes) – $X_2 = 17$, $n_2 = 24$, $\hat{p}_2 = 17/24$

$\alpha = 0.05$, $\bar{p} = (12+17)/(34+24) = 1/2$, $\bar{q} = 1 - \bar{p} = 1/2$

TESTING THE DIFFERENCE BETWEEN PROPORTIONS-EXAMPLE 1

Step 1: State the hypotheses and identify the claim.

H_0: $p_1 = p_2$ (claim)

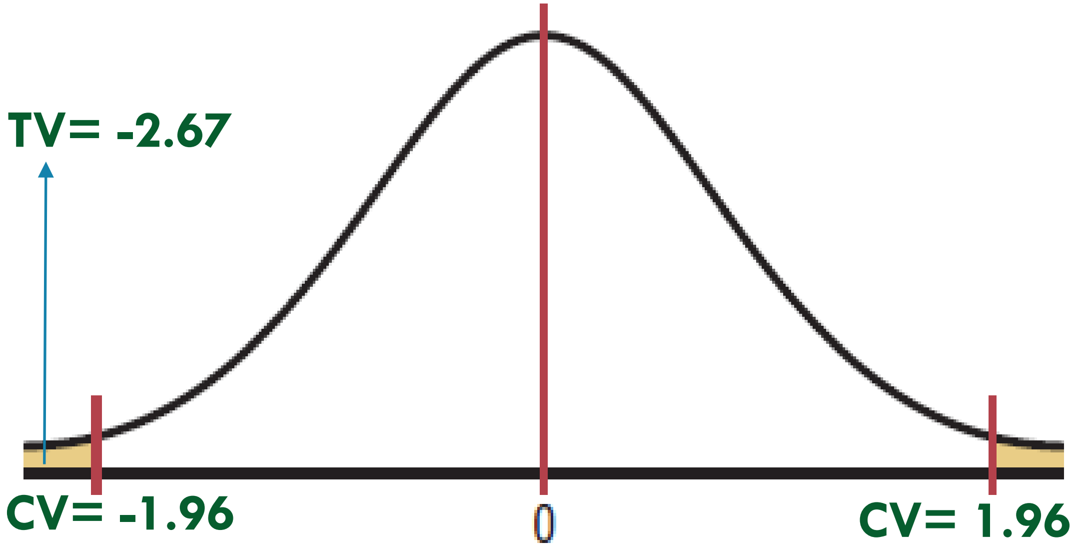
H_1: $p_1 \neq p_2$ Two-tailed test

Step 2: Find the critical values. Since $\alpha = 0.05$ and the test is a two-tailed test, the critical values are ±1.96.

Step 3: Compute the test value.

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\bar{p}\bar{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{\left(\frac{12}{34} - \frac{17}{24}\right) - 0}{\sqrt{\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{34} + \frac{1}{24}\right)}} = -2.67$$

TESTING THE DIFFERENCE BETWEEN PROPORTIONS-EXAMPLE 1



Step 4: Make the decision.
H_o: Reject (claim) and H_1: Do not reject since the test value falls in the critical region.

- **Step 5:** Summarize the results. The evidence is not enough to support the claim that there is no difference in the proportions of small and large nursing homes with a resident vaccination rate of less than 80%.

TESTING THE DIFFERENCE BETWEEN PROPORTIONS-EXAMPLE 2

A survey of 1000 drivers this year showed that 29% of the people send text messages while driving. Last year a survey of 1000 drivers showed that 17% of those send text messages while driving. At $\alpha=0.01$, can it be concluded that there has been an increase in the number of drivers who text while driving?

TESTING THE DIFFERENCE BETWEEN PROPORTIONS EXAMPLE 2

Two-tailed test	Right-tailed test	Left-tailed test
H_0: $p_1 = p_2$	H_0: $p_1 \le p_2$	H_0: $p_1 \ge p_2$
H_1: $p_1 \ne p_2$	H_1: $p_1 > p_2$	H_1: $p_1 < p_2$

rs

$\alpha = 0.01$, can it be concluded that there has been an increase in the number of drivers who text while driving?

p_1 *(survey this year)* $>$ p_2 *(survey last year)*

Values Needed:

Group 1(survey this year) – $n_1 = 1000$, $\hat{p}_1 = 0.29$, $X_1 = 1000(0.29) = 290$

Group 2(survey last year) – $n_2 = 1000$, $\hat{p}_2 = 0.17$, $X_2 = 1000(0.17) = 170$

$\alpha = 0.01$, $\bar{p} = (290+170)/(1000+1000) = 23/100$, $\bar{q} = 1 - \bar{p} = 77/100$

TESTING THE DIFFERENCE BETWEEN PROPORTIONS-EXAMPLE 2

Step 1: State the hypotheses and identify the claim.

H_0: $p_1 \leq p_2$

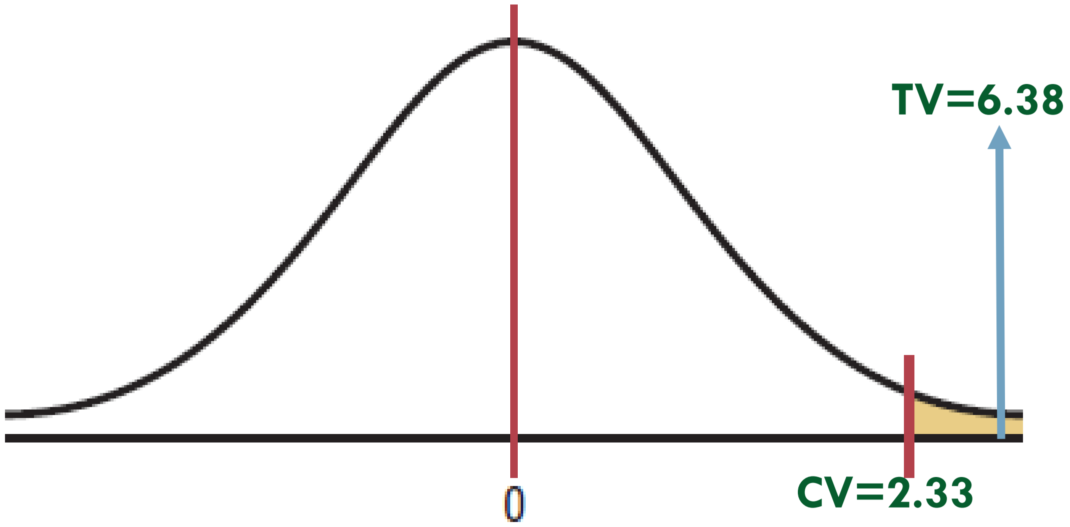
H_1: $p_1 > p_2$ (claim) Right-tailed test

Step 2: Find the critical value. Since $\alpha = 0.01$ and the test is a right-tailed test, the critical value is 2.33.

Step 3: Compute the test value.

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\bar{p}\bar{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.29 - 0.17) - 0}{\sqrt{\left(\frac{23}{100}\right)\left(\frac{77}{100}\right)\left(\frac{1}{1000} + \frac{1}{1000}\right)}} = 6.38$$

TESTING THE DIFFERENCE BETWEEN PROPORTIONS-EXAMPLE 2



Step 4: Make the decision.
H_o: Reject and H_1: Do not reject (claim) since the test value falls in the critical region.

- **Step 5:** Summarize the results. The evidence is enough to support the claim that the proportion of drivers who send text messages is larger today than it was last year.

TESTING THE DIFFERENCE BETWEEN TWO VARIANCES OR TWO STANDARD DEVIATIONS

Anthony S. Morfe_CAS_Math Cluster

TESTING THE DIFFERENCE BETWEEN TWO VARIANCES OR TWO STANDARD DEVIATIONS

If two independent samples are selected from two normally distributed populations in which the variances are equal ($\sigma_1^2 = \sigma_2^2$) and if the variances s_1^2 and s_2^2 are compared as $\frac{s_1^2}{s_2^2}$, the sampling distribution of the variances is called the ***F* distribution.**

Assumptions for Testing the Difference Between Two Variances

1. The samples must be random samples.
2. The populations from which the samples were obtained must be normally distributed. (*Note:* The test should not be used when the distributions depart from normality.)
3. The samples must be independent of one another.

TESTING THE DIFFERENCE BETWEEN TWO VARIANCES OR TWO STANDARD DEVIATIONS

- F distribution is not symmetric, and there are no negative values, so you may not simply take the opposite of the right critical value to find the left critical value.

- Since the left critical values are a pain to calculate, they are often avoided altogether. You can force the F test into a right tail test by placing the sample with the large variance in the numerator and the smaller variance in the denominator. It does not matter which sample has the larger sample size, only which sample has the larger variance.

TESTING THE DIFFERENCE BETWEEN TWO VARIANCES OR TWO STANDARD DEVIATIONS

Characteristics of the *F* Distribution

1. The values of F cannot be negative, because variances are always positive or zero.
2. The distribution is positively skewed.
3. The mean value of F is approximately equal to 1.
4. The F distribution is a family of curves based on the degrees of freedom of the variance of the numerator and the degrees of freedom of the variance of the denominator.

TESTING THE DIFFERENCE BETWEEN TWO VARIANCES OR TWO STANDARD DEVIATIONS

Formula for the *F* Test

$$F = \frac{s_1^2}{s_2^2}$$

where the larger of the two variances is placed in the numerator regardless of the subscripts.

The F test has two terms for the degrees of freedom: that of the numerator, $n_1 - 1$, and that of the denominator, $n_2 - 1$, where n_1 is the sample size from which the larger variance was obtained.

TESTING THE DIFFERENCE BETWEEN TWO VARIANCES OR TWO STANDARD DEVIATIONS

Notes for the Use of the *F* Test

1. The larger variance should always be placed in the numerator of the formula regardless of the subscripts.

$$F = \frac{s_1^2}{s_2^2}$$

2. For a two-tailed test, the α value must be divided by 2 and the critical value placed on the right side of the F curve.
3. If the standard deviations instead of the variances are given in the problem, they must be squared for the formula for the F test.
4. When the degrees of freedom cannot be found in Table H, the closest value on the smaller side should be used.

TESTING THE DIFFERENCE BETWEEN TWO VARIANCES OR TWO STANDARD DEVIATIONS

Example: Find the critical value for a right-tailed *F* test when a=0.05, the degrees of freedom for the numerator (abbreviated d.f.N.) are 15, and the degrees of freedom for the denominator (d.f.D.) are 21.

Table H *(continued)*

d.f.D.: degrees of freedom, denominator	$\alpha = 0.05$												
	d.f.N.: degrees of freedom, numerator												
	1	2	3	4	5	6	7	8	9	10	12	15	20
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	245.9	248.0
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07

Answer: C.V.=2.18

TESTING THE DIFFERENCE BETWEEN TWO VARIANCES OR TWO STANDARD DEVIATIONS

Example: Find the critical value for a two-tailed *F* test with a=0.05 when the sample size from which the variance for the numerator was obtained was 21 and the sample size from which the variance for the denominator was obtained was 12.

Given: a=0.05, a/2=0.025, $n_{numerator}$=21, dfN=20, $n_{denominator}$=12, dfD=11

Table H *(continued)*

d.f.D.: degrees of freedom, denominator	α = 0.025 d.f.N.: degrees of freedom, numerator													
	1	2	3	4	5	6	7	8	9	10	12	15	20	24
1	647.8	799.5	864.2	899.6	921.8	937.1	948.2	956.7	963.3	968.6	976.7	984.9	[illegible]	997.2
2	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40	39.41	39.43	[illegible]	39.46
3	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42	14.34	14.25	[illegible]	14.12
4	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.75	8.66	[illegible]	8.51
9	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.87	3.77	[illegible]	3.61
10	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.62	3.52	[illegible]	3.37
11	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	3.23	3.17
12	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.28	3.18	3.07	3.02

Answer: C.V.=3.23

TESTING THE DIFFERENCE BETWEEN TWO VARIANCES OR TWO STANDARD DEVIATIONS

Example: The standard deviation of the average waiting time to see a doctor for non-life threatening problems in the emergency room at an urban hospital is 28 minutes. At a second hospital, the standard deviation is 32 minutes. If a sample of 18 patients was used in the first case and 16 in the second case, is there enough evidence to conclude at the 0.01 significance level that the standard deviation of the waiting times in the first hospital is less than the standard deviation of the waiting times in the second hospital?

Given: $\alpha=0.01$, $s_1=28$, $s_2=32$, $F=s_2^2/s_1^2$, $n_1=18$, dfD=17, $n_2=16$, dfN=15

TESTING THE DIFFERENCE BETWEEN TWO VARIANCES OR TWO STANDARD DEVIATIONS

Two-tailed test	Right-tailed test	Left-tailed test
H_0: $\sigma_1=\sigma_2$ H_1: $\sigma_1\neq\sigma_2$	H_0: $\sigma_1\leq\sigma_2$ H_1: $\sigma_1>\sigma_2$	H_0: $\sigma_1\geq\sigma_2$ H_1: $\sigma_1<\sigma_2$

deviation is 32 minutes. If a sample of 18 patients was used in the first case and 16 in the second case, is there enough evidence to conclude at the 0.01 significance level that the standard deviation of the waiting times in the first hospital is less than the standard deviation of the waiting times in the second hospital?

Given: $\alpha=0.01$, $s_1=28$, $s_2=32$, $F=s_2^2/s_1^2$, $n_1=18$, $dfD=17$, $n_2=16$, $dfN=15$

TESTING THE DIFFERENCE BETWEEN TWO VARIANCES OR TWO STANDARD DEVIATIONS

Step 1: State the hypotheses and identify the claim.

H_0: $\sigma_1 \geq \sigma_2$ H_1: $\sigma_1 < \sigma_2$ (claim) Left-tailed test

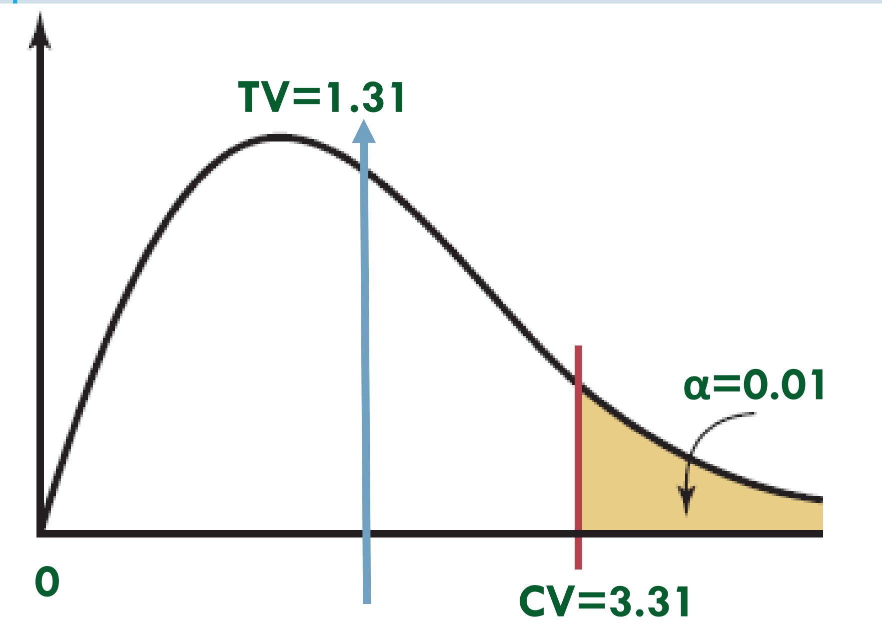
Step 2: Find the critical value given α=0.01, dfN=15, and dfD=17

Table H *(continued)*

d.f.D.: degrees of freedom, denominator	α = 0.01												
	d.f.N.: degrees of freedom, numerator												
	1	2	3	4	5	6	7	8	9	10	12	15	20
1	4052	4999.5	5403	5625	5764	5859	5928	5982	6022	6056	6106	6157	6209
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.42	99.43	99.45
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23	27.05	26.87	26.69
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.37	14.20	14.02
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08

C.V.=3.31

TESTING THE DIFFERENCE BETWEEN TWO VARIANCES OR TWO STANDARD DEVIATIONS



Step 3: Compute the test value.

Given: $s_1=28$, $s_2=32$,

$F=s_2^2/s_1^2=32^2/28^2=1.31$

Step 4: Make the decision.

H_o: Do not Reject and H_1: Reject (claim) since the test value is not in the critical region.

- Step 5: Summarize the results. The evidence is not enough to support the claim that the standard deviation of the waiting times in the first hospital is less than the standard deviation of the waiting times in the second hospital

TESTING THE DIFFERENCE BETWEEN TWO VARIANCES OR TWO STANDARD DEVIATIONS

Example: A medical researcher wishes to see whether the variance of the heart rates (in beats per minute) of smokers is different from the variance of heart rates of people who do not smoke. Two samples are selected, and the data are as shown. Using a = 0.05, is there enough evidence to support the claim?

Smokers	Nonsmokers
$n_1=25$	$n_2=18$
$s_1^2=36$	$s_2^2=10$

Given: $\alpha=0.05$, $s_1^2=36$, $s_2^2=10$, $F=s_1^2/s_2^2$, $n_1=25$, dfN=24, $n_2=18$, dfD=17

Two-tailed test	Right-tailed test	Left-tailed test
H_0: $\sigma_1^2=\sigma_2^2$	H_0: $\sigma_1^2\le\sigma_2^2$	H_0: $\sigma_1^2\ge\sigma_2^2$
H_1: $\sigma_1^2\ne\sigma_2^2$	H_1: $\sigma_1^2>\sigma_2^2$	H_1: $\sigma_1^2<\sigma_2^2$

Example: A medical researcher wishes to see whether the variance of the heart rates (in beats per minute) of smokers is different from the variance of heart rates of people who do not smoke. Two samples are selected, and the data are as shown. Using a = 0.05, is there enough evidence to support the claim?

Smokers	Nonsmokers
$n_1=25$	$n_2=18$
$s_1^2=36$	$s_2^2=10$

Given: $\alpha=0.05$, $\alpha/2=0.025$, $s_1^2=36$, $s_2^2=10$, $F=s_1^2/s_2^2$, $n_1=25$, dfN=24, $n_2=18$, dfD=17

TESTING THE DIFFERENCE BETWEEN TWO VARIANCES OR TWO STANDARD DEVIATIONS

Step 1: State the hypotheses and identify the claim.

H_0: $\sigma_1^{\,2} = \sigma_2^{\,2}$ H_1: $\sigma_1^{\,2} \neq \sigma_2^{\,2}$ (claim) Two-tailed test

Step 2: Find the critical value given $\alpha/2=0.025$, dfN=24, and dfD=17

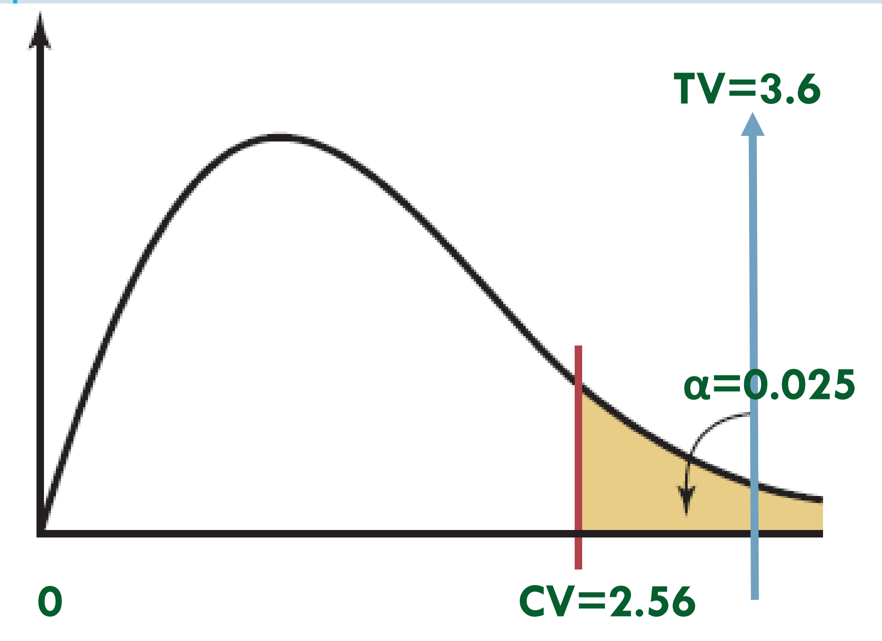
Table H *(continued)*

d.f.D.: degrees of freedom, denominator	$\alpha = 0.025$													
	d.f.N.: degrees of freedom, numerator													
	1	2	3	4	5	6	7	8	9	10	12	15	20	24
1	647.8	799.5	864.2	899.6	921.8	937.1	948.2	956.7	963.3	968.6	976.7	984.9	993.1	997.2
2	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40	39.41	39.43	39.45	39.46
3	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42	14.34	14.25	14.17	14.12
4	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.75	8.66	8.56	8.51
15	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	2.96	2.86	2.76	2.70
16	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05	2.99	2.89	2.79	2.68	2.63
17	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98	2.92	2.82	2.72	2.62	2.56
18	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93	2.87	2.77	2.67	2.56	2.50

C.V.=2.56

Anthony S. Morfe_CAS_Math Cluster

TESTING THE DIFFERENCE BETWEEN TWO VARIANCES OR TWO STANDARD DEVIATIONS



Step 3: Compute the test value.

Given: $s_1^2=36$, $s_2^2=10$

$F=s_1^2/s_2^2=36/10=3.6$

Step 4: Make the decision.

H_o: Reject and H_1: Do not reject (claim) since the test value falls in the critical region.

- **Step 5:** Summarize the results. The evidence is enough to support the claim that the variance of the heart rates (in beats per minute) of smokers is different from the variance of heart rates of people who do not smoke.

REFERENCE:

Bluman, A. (2012). *Elementary Statistics: A Step by Step Approach*,8e. McGraw-Hill Higher Ed.

End of Lecture 10

Anthony S. Morfe_CAS_Math Cluster