Ch 10

- The means of two independent populations

Example (1)

You are a financial analyst for a brokerage firm. Is there a difference in dividend yield between stocks listed on the NYSE & NASDAQ? You collect the following data:

NYSE: $n_1 = 21$, $\bar{X}_1 = 3.27$, $S_1 = 1.30$

NASDAQ: $n_2 = 25$, $\bar{X}_2 = 2.53$, $S_2 = 1.16$

Assuming both populations are approximately normal <u>with equal variances, is there a difference in mean</u>
yield (α = 0.05)?

Solution:

$(\sigma_1 \& \ \sigma_2 \ \ unknown)\ (\sigma_1 = \sigma_2)$ t

Step 1: state the hypothesis:

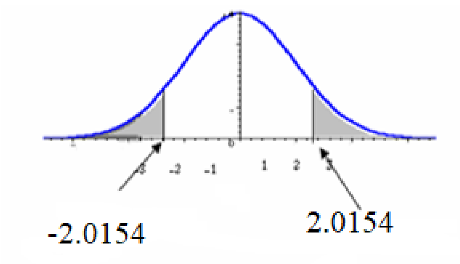
$$H_0: \mu_1 = \mu_2$$
$$H_1: \mu_1 \neq \mu_2$$

Test is two-tailed test (key word is difference between 2 samples)

Step 2- Select the level of significance and critical value.

α = 0.05 as stated in the problem

$$\pm t_{\left(\frac{\alpha}{2}, n_1+n_2-2\right)} = \pm t_{\left(\frac{0.05}{2}, 21+25-2\right)} = \pm t_{(0.025, 44)} = \pm 2.0154$$



Step 3: Select the test statistic.

Use Z-distribution since the assumptions are met

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 - 1) + (n_2 - 1)} = \frac{(21-1) \times 1.3^2 + (25-1) \times 1.16^2}{(21-1) + (25-1)}$$
$$= \frac{33.8 + 32.2944}{44} = \frac{66.0944}{44} = 1.5021$$

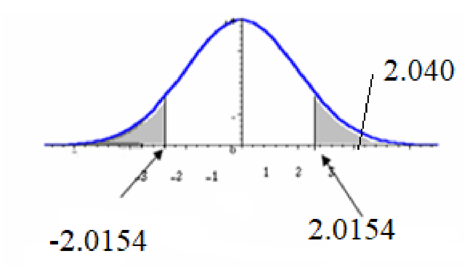
$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$t_{stat} = \frac{3.27 - 2.53}{\sqrt{1.5021(\frac{1}{21} + \frac{1}{25})}} = \frac{0.74}{0.3627} = 2.040$$

Step 4: Formulate the decision rule. (Critical value)

Reject H_0if t_c>2.0154 Or $t_c < -2.0154$

Step 5: Make a decision and interpret the result.



× $H_0: \mu_1 \le \mu_2$

√ $H_1: \mu_1 > \mu_2$

Reject H_0 at α= 0.05, There is evidence of a difference in means.

Reject H_0. There is sufficient evidence that $\mu_{NYSE} > \mu_{NASDAQ}$

95% Confidence Interval for $\mu_{NYSE} - \mu_{NASDAQ}$

$$\left(\overline{X}_1 - \overline{X}_2\right) \pm t_{\alpha/2} \sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = 0.74 \pm 2.0154 \times 0.3628 = (0.009,\ 1.471)$$

Since 0 is less than the entire interval, we rejected H_0 and we can be 95% confident that $\mu_{NYSE} > \mu_{NASDAQ}$

- The means of two related populations

Example (2)

Assume you send your salespeople to a "customer service" training workshop. Has the training made a difference in the number of complaints (at the 0.01 level)? You collect the following data:

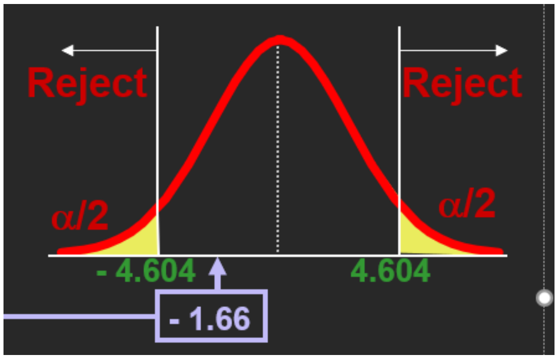
Salesperson	**Number of Complaints** **Before**	**Number of Complaints** **After**
C.B.	6	4
T.F.	20	6
M.H.	3	2
R.K.	0	0
M.O.	4	0

Solution:

Step 1: state the hypothesis:

H_0: $\mu_D = 0$

H_1: $\mu_D \neq 0$



Step2: Select the level of significance and critical value.

$t_{0.005} = \pm 4.604$ d.f. = n - 1 = 4

Step 3: Find the appropriate test statistic.

Salesperson	**Number of Complaints** **Before**	**Number of Complaints** **After**	D **(X2-X1)**	$\left(D_i - \overline{D}\right)$	$\left(D_i - \overline{D}\right)^2$
C.B.	6	4	-2	-2-(-4.2)= 2.2	4.84
T.F.	20	6	-14	-14-(-4.2)=-9.8	96.04
M.H.	3	2	-1	-1-(-4.2)= 3.2	10.24
R.K.	0	0	0	0-(-4.2)=4.2	17.64
M.O.	4	0	-4	4-(-4.2)=0.2	0.04
Total			-21		128.8

$$\bar{D} = \frac{\sum D_i}{n} = \frac{-21}{5} = -4.2$$

$$S_D = \sqrt{\frac{\sum(D_i - \bar{D})^2}{n-1}} = \sqrt{\frac{128.8}{4}} = 5.6745$$

$$t_c = \frac{\bar{D}}{S_D/\sqrt{n}} = \frac{-4.2}{5.6745/2.2361} = \frac{-4.2}{2.5377} = -1.66$$

Step 4: State the decision rule

Reject H_0 if

$$t_c > 4.604$$
$$or$$
$$t_c < -4.604$$

Step 5: Decision Reject H_0

$$\surd \qquad H_0: \mu_D = 0$$
$$\times \qquad H_1: \mu_D \neq 0$$

Do not reject H_0 (t_{stat} is not in the rejection region)

Do not reject H_0. There is insufficient of a change in the number of complaints.

$$\mu_{comlaints\ betore\ training\ workshop} = \mu_{complaints\ \ after\ training\ workshop}$$

-The Paired Difference Confidence Interval μ_D is:

$$\hat{\mu}_D = \overline{D} \pm t_{\alpha/2} \frac{S_D}{\sqrt{n}}$$
$$= -4.2 \pm 4.604 \frac{5.6745}{\sqrt{5}}$$
$$= -4.2 \pm 11.6836$$
$$-15.87 < \hat{\mu}_D < 7.48$$

Since this interval contains 0 you are 99% confident that $\mu_D = 0$

Do not reject H_0

$$\mu_{comlaints\ betore\ training\ workshop} = \mu_{complaints\ \ after\ training\ workshop}$$

- The proportions of two independent populations

Example (3)

Is there a significant difference between the proportion of men and the proportion of women who will vote Yes on Proposition A?

In a random sample, 36 of 72 men and 35 of 50 women indicated they would vote Yes

Test at the .05 level of significance

Solution:

Step 1: State the null and alternate hypotheses.

H0: π1 – π2 = 0 (the two proportions are equal)

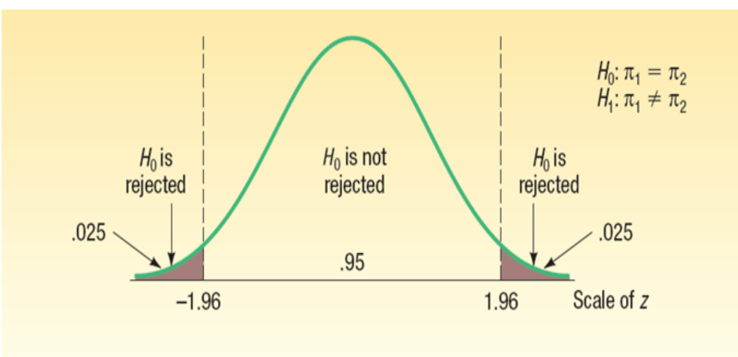
H1: π1 – π2 ≠ 0 (there is a significant difference between proportions)

Step 2: State the level of significance and critical value.

The .05 significance level is stated in the problem.

$$\pm Z_{\frac{\alpha}{2}} = \pm Z_{\frac{0.05}{2}} = \pm Z_{0.025}$$

$$\pm Z_{0.025} = \pm 1.96$$



Step 3: Find the appropriate test statistic.

The sample proportions are:

Men: $p_1 = 36/72 = 0.50$

Women: $p_2 = 35/50 = 0.70$

The pooled estimate for the overall proportion is:

$$\overline{p} = \frac{X_1 + X_2}{n_1 + n_2} = \frac{36 + 35}{72 + 50} = \frac{71}{122} = .582$$

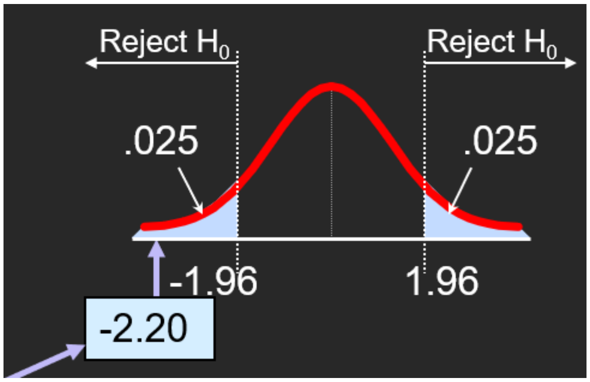
$$z_{\text{STAT}} = \frac{(p_1 - p_2) - (\pi_1 - \pi_2)}{\sqrt{\overline{p}(1-\overline{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$= \frac{(.50 - .70) - (0)}{\sqrt{.582(1 - .582)\left(\frac{1}{72} + \frac{1}{50}\right)}} = -2.20$$

Step 4: State the decision rule

Reject H_0 if $Z_c > Z_{\frac{\alpha}{2}}$ Or $Z_c < -Z_{\frac{\alpha}{2}}$

$Z_c > 1.96$ Or $Z_c < -1.96$



Step 5:Decision Reject H_0

There is evidence of a significant difference in the proportion of men and women who will vote yes.

× $H_0: \pi_1 \leq \pi_2$
√ $H_1: \pi_1 > \pi_2$

Reject H_0. There is sufficient evidence that $\pi_{weman} > \pi_{men}$

The confidence interval for

$\pi_1 - \pi_2$ is:

$$(p_1 - p_2) \pm Z_{\alpha/2} \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

The 95% confidence interval for $\pi_1 - \pi_2$ is:

$$(0.50 - 0.70) \pm 1.96 \sqrt{\frac{0.50(0.50)}{72} + \frac{0.70(0.30)}{50}}$$

$= (-0.37, -0.03)$

Since this interval does not contain 0 can be 95% confident the two proportions are different. *There is sufficient evidence that* $\pi_{weman} > \pi_{men}$

- **Testing for the Ratio Of Two Population Variances**

Example Example (4)

Determine the upper-tail critical values of F in each of the following two-tail tests.
What is the value of FSTAT?

$\alpha = 0.01 \quad n_1 = 13 \quad , \qquad S_1^2 = 26$

$n_2 = 16 \quad , \qquad S_2^2 = 11$

Solution:

- In the F table,

- numerator degrees of freedom determine the column ($v_1 = df_1 = n_1 - 1$)
- denominator degrees of freedom determine the row ($v_2 = df_2 = n_2 - 1$)

$$F_{\frac{\alpha}{2}, n_1-1, n_2-1} = F_{\frac{0.01}{2}, 13-1, 16-1} = F_{0.005, 12, 15} = 4.25$$

Cumulative Probabilities = 0.995											
Upper-Tail Areas = 0.005											
	Numerator, df_1										
Denominator, df_2	1	2	3	4	5	6	7	8	9	10	12	15
1	16,211.00	20,000.00	21,615.00	22,500.00	23,056.00	23,437.00	23,715.00	23,925.00	24,091.00	24,224.00	24,426.00	24,630.00
2	198.50	199.00	199.20	199.20	199.30	199.30	199.40	199.40	199.40	199.40	199.40	199.40
3	55.55	49.80	47.47	46.19	45.39	44.84	44.43	44.13	43.88	43.69	43.39	43.08
4	31.33	26.28	24.26	23.15	22.46	21.97	21.62	21.35	21.14	20.97	20.70	20.44
5	22.78	18.31	16.53	15.56	14.94	14.51	14.20	13.96	13.77	13.62	13.38	13.15
6	18.63	14.54	12.92	12.03	11.46	11.07	10.79	10.57	10.39	10.25	10.03	9.81
7	16.24	12.40	10.88	10.05	9.52	9.16	8.89	8.68	8.51	8.38	8.18	7.97
8	14.69	11.04	9.60	8.81	8.30	7.95	7.69	7.50	7.34	7.21	7.01	6.81
9	13.61	10.11	8.72	7.96	7.47	7.13	6.88	6.69	6.54	6.42	6.23	6.03
10	12.83	9.43	8.08	7.34	6.87	6.54	6.30	6.12	5.97	5.85	5.66	5.47
11	12.23	8.91	7.60	6.88	6.42	6.10	5.86	5.68	5.54	5.42	5.24	5.05
12	11.75	8.51	7.23	6.52	6.07	5.76	5.52	5.35	5.20	5.09	4.91	4.72
13	11.37	8.19	6.93	6.23	5.79	5.48	5.25	5.08	4.94	4.82	4.64	4.46
14	11.06	7.92	6.68	6.00	5.56	5.26	5.03	4.86	4.72	4.60	4.43	4.25
15	10.80	7.70	6.48	5.80	5.37	5.07	4.85	4.67	4.54	4.42	4.25	4.07
16	10.58	7.51	6.30	5.64	5.21	4.91	4.69	4.52	4.38	4.27	4.10	3.92
17	10.38	7.35	6.16	5.50	5.07	4.78	4.56	4.39	4.25	4.14	3.97	3.79

$$F_{stat} = \frac{26}{11} = 2.36$$

Example (5)

Determine the upper-tail critical value of F in each of the following one-tail tests for a claim that the variance of sample 1 is greater than the variance of sample 2.

What is the value of FSTAT?

$$\alpha = 0.025 \quad n_1 = 16 \quad , \qquad S_1^2 = 25$$
$$n_2 = 13 \quad , \qquad S_2^2 = 50$$

Solution:

$$F_{\alpha, n_1 - 1, n_2 - 1} = F_{0.025, 13-1, 16-1} = F_{0.025, 12, 15} = 2.96$$

Denominator, df_2	Numerator, df_1: 1	2	3	4	5	6	7	8	9	10	12	15
	Cumulative Probabilities = 0.975; Upper-Tail Areas = 0.025											
1	647.80	799.50	864.20	899.60	921.80	937.10	948.20	956.70	963.30	968.60	976.70	984.90
2	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.39	39.39	39.40	39.41	39.43
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
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
5	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.52	6.43
6	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.37	5.27
7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.67	4.57
8	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.20	4.10
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
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
11	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.43	3.33
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
13	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31	3.25	3.15	3.05
14	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21	3.15	3.05	2.95
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
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
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

$$F_{stat} = \frac{50}{25} = 2$$

Example (6)

Determine the upper-tail critical values of F in each of the following two-tail tests.

$\alpha = 0.10 \quad n_1 = 14 \quad , \qquad n_2 = 16$

Solution:

$$F_{\frac{\alpha}{2}, n_1-1, n_2-1} = F_{\frac{0.10}{2}, 14-1, 16-1} = F_{0.05, \ 13, 15} = \frac{2.48+2.40}{2} = 2.44$$

Upper-Tail Areas = 0.05													
	Numerator, df_1												
Denominator, df_2	1	2	3	4	5	6	7	8	9	10	12	15	20
1	161.40	199.50	215.70	224.60	230.20	234.00	236.80	238.90	240.50	241.90	243.90	245.90	248.00
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
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19
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

Example (7)

You are a financial analyst for a brokerage firm. Is there a difference in dividend yield between stocks listed on the NYSE & NASDAQ? You collect the following data:

NYSE: $n_1 = 21$, $S_1 = 1.30$

NASDAQ: $n_2 = 25$, $S_2 = 1.16$

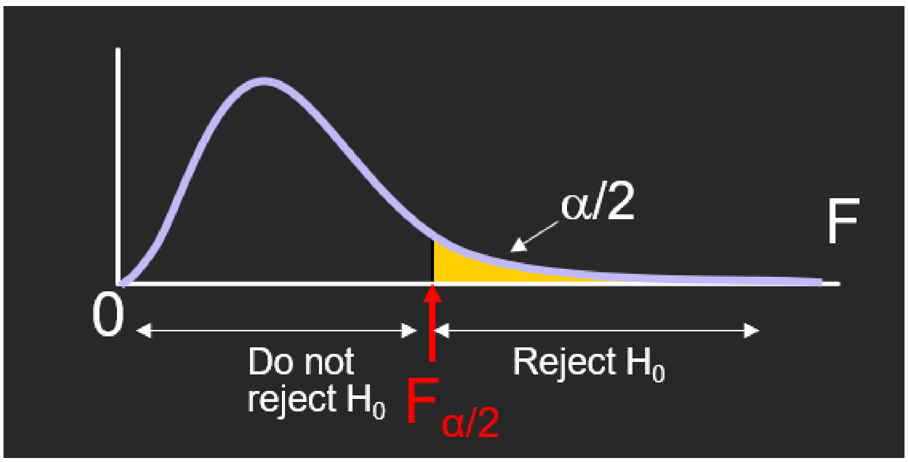
Is there a difference in the variances between the NYSE & NASDAQ at the

$\alpha = 0.05$ level?

Step 1: state the hypothesis:

H_0: $\sigma^2_1 = \sigma^2_2$ (there is no difference between variances)

H_1: $\sigma^2_1 \neq \sigma^2_2$ (there is a difference between variances)



Step2: Select the level of significance and critical value.

Find the F critical value for $\alpha = 0.05$:

Numerator d.f. = $n_1 - 1 = 21 - 1 = 20$

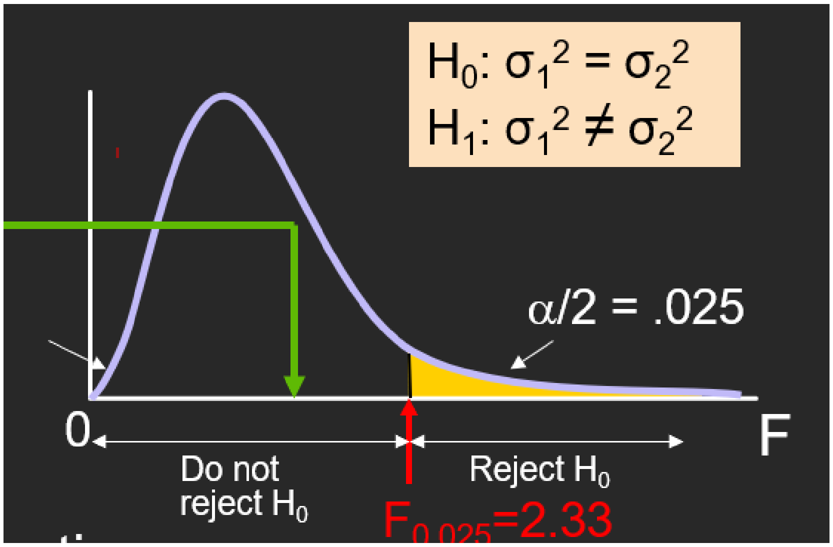
Denominator d.f. = $n_2 - 1 = 25 - 1 = 24$

$F_{\alpha/2} = F_{.025,\,20,\,24} = 2.33$

Step 3: Find the appropriate test statistic.

The test statistic is:

$$F_{STAT} = \frac{S_1^2}{S_2^2} = \frac{1.30^2}{1.16^2} = 1.256$$



Step 4: State the decision rule

Reject H_0 if $F_{STAT} > F_{\alpha/2} = 2.33$

$F_{STAT} = 1.256$ is not in the rejection region, so we do not reject H_0

Step 5: Decision Reject H_0

because $F_{stat} = 1.256 < 2.33$,you do not reject H_0.using a 0.05 level of significance, you conclude that there is evidence that there is no variability in the NYSE & NASDAQ.

Do not reject the null hypothesis, there was no significant difference between two variances ($\sigma_1^2 = \sigma_2^2$). The test statistic will be use is Pooled variance t-test

√ $H_0: \sigma^2_1 = \sigma^2_2$ (there is no difference between variances)

× $H_1: \sigma^2_1 \neq \sigma^2_2$ (there is a difference between variances)

	Cumulative Probabilities = 0.975												
	Upper-Tail Areas = 0.025												
	Numerator, df_1												
Denominator, df_2	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**12**	**15**	**20**
1	647.80	799.50	864.20	899.60	921.80	937.10	948.20	956.70	963.30	968.60	976.70	984.90	993.10
2	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.39	39.39	39.40	39.41	39.43	39.45
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
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
5	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.52	6.43	6.33
6	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.37	5.27	5.17
7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.67	4.57	4.47
8	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.20	4.10	4.00
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	3.67
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	3.42
11	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.43	3.33	3.23
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
13	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31	3.25	3.15	3.05	2.95
14	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21	3.15	3.05	2.95	2.84
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
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
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
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
19	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88	2.82	2.72	2.62	2.51
20	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.68	2.57	2.46
21	5.83	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.80	2.73	2.64	2.53	2.42
22	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76	2.70	2.60	2.50	2.39
23	5.75	4.35	3.75	3.41	3.18	3.02	2.90	2.81	2.73	2.67	2.57	2.47	2.36
24	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70	2.64	2.54	2.44	2.33

Example (8) Text book pages (372 & 373)

Waiting time is a critical issue at fast-food chains, which not only want to minimize the mean service time but also want to minimize the variation in the service time from customer to customer .One fast-food chain carried out a study to measure the variability in the waiting time (defined as the time in minutes from when an order was completed to when it was delivered to the customer) at lunch and breakfast at one of the chain's stores. The results were as follows:

$Lunch: n_1 = 25$, $S_1^2 = 4.4$

Breakfast: $n_2 = 21$, S_2^2=1.9

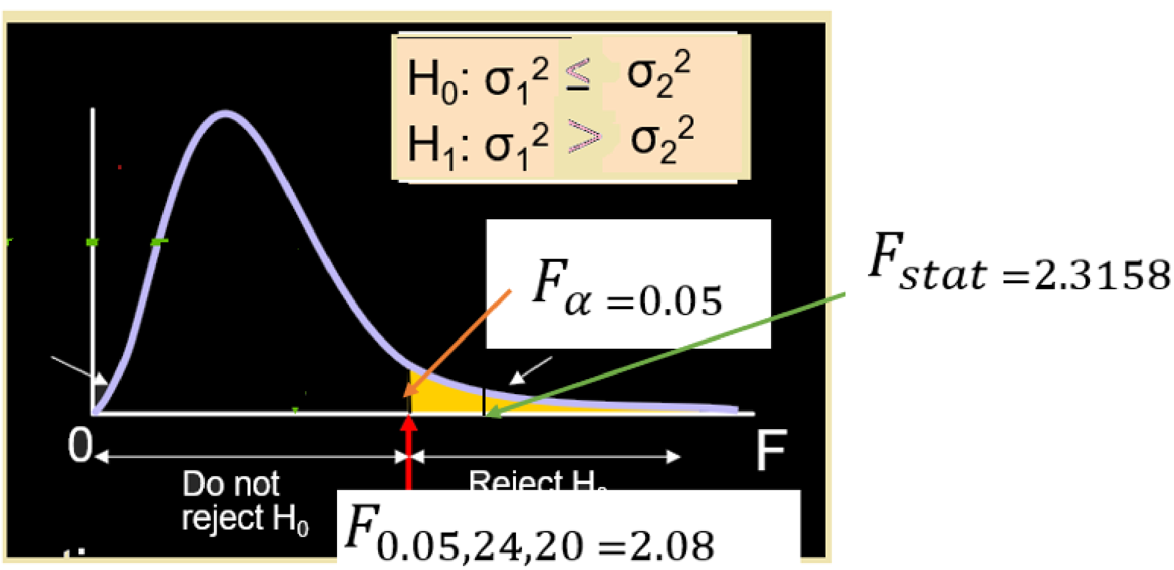
At the 0.05 level of significance, is there evidence that there is more variability in the service time at lunch than at breakfast? Assume that the population service times are normally distributed

Solution

1)

$H_0: \sigma_1^2 \leq \sigma_2^2$

$H_1: \sigma_1^2 > \sigma_2^2$



2)

$F_{\alpha,n_{1-1},n_2-1} = F_{0.05,25-1,21-1} = F_{0.05,24,20} = 2.08$

3)

$$F_{stat} = \frac{S_1^2}{S_2^2} = \frac{4.4}{1.9} = 2.3158$$

4)

Reject H_0 if $F_{stat} > 2.08$

5) because $F_{stat} = 2.3158 > 2.08$,you reject H_0.using a 0.05 level of significance, you conclude that there is evidence that there is more variability in the service time at lunch than at breakfast.

Reject the null hypothesis, there was a significant difference between two variances ($\sigma_1^2 \neq \sigma_2^2$). The test statistic will be use is Separate variance t-test

× $H_0: \sigma_1^2 \leq \sigma_2^2$

√ $H_1: \sigma_1^2 > \sigma_2^2$

	Upper-Tail Areas = 0.05													
	Numerator, df_1													
Denominator, df_2	1	2	3	4	5	6	7	8	9	10	12	15	20	24
1	161.40	199.50	215.70	224.60	230.20	234.00	236.80	238.90	240.50	241.90	243.90	245.90	248.00	249.10
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	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	8.64
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	5.77
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15
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	2.11
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	2.08
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	2.05

- The means of more than two populations

Example (9) An experiment has a Sigle factor with four groups and nine values in each group.

If SSA=752,   SST=1250

Answer the following questions:

1) How many degrees of freedom are there in determining the among-group variation?

2) How many degrees of freedom are there in determining the within-group variation?

3) How many degrees of freedom are there in determining the total variation?

4) What is SSW?

5) What is MSA?

6) What is MSW?

7) What is the value of F_{stat} ?

Solution:

1) How many degrees of freedom are there in determining The among-group variation?

c-1=4-1=3

2) How many degrees of freedom are there in determining The within-group variation?

n = 4 x 9 = 36

n-c=36-4=32

3) How many degrees of freedom are there in determining The total variation?

n-1=36-1=35

4)What is SSW?

SST-SSA=1250-752=498

5)What is MSA?

$$MSA = \frac{752}{3} = 250.67$$

6)What is MSW?

$$MSW = \frac{498}{32} = 15.5625$$

7)What is the value of F_{stat} ?

$$F_{stat} = \frac{MSA}{MSW} = \frac{250.67}{15.5625} = 16.11$$

Example (10) (Slide 65-68)

You want to see if three different golf clubs yield different distances. You randomly select five measurements from trials on an automated driving machine for each club. At the 0.05 significance level, is there a difference in mean distance?

	Club 1	Club 2	Club 3
	254	234	200
	263	218	222
	241	235	197
	237	227	206
	251	216	204
Total	1246	1130	1029
Mean	249.2	226	205.8
$\bar{\bar{X}} = 227$			

C=3 n=15

SSA = 4716.4 , SSW = 1119.6

Solution:

Source of variation (S.V)	Degrees of freedom	Sum of Squares (S.S)	Mean Squares (MS)	F- ratio
Among groups	c-1=3-1=2	SSA=4716.4	MSA =SSA/c-1 $\frac{4716.4}{2} = 2358.2$	F_{STAT} = MSA / MSW $= \frac{2358.2}{93.3} = 25.275$
Within groups	n-c=15-3=12	SSW=1119.6	MSW = SSW/n-c $\frac{1119.6}{12} = 93.3$	
Total	n-1=15-1=14	SST=5836		

Step (1) : State the null and alternate hypotheses :

$H_0 : \mu_1 = \mu_2 = \mu_3$

H_1 : Not all μ_j are equal(Not all the means are equal.).

Step (2): Select the level of significance (α =0.05)

Step (3): The test statistic :

$$F_{STAT} = \frac{MSA}{MSW} = \frac{SSA/c-1}{SSW/n-c} = \frac{4716.4/3-1}{1119.6/15-3} = 25.275$$

Step (4): The critical value:

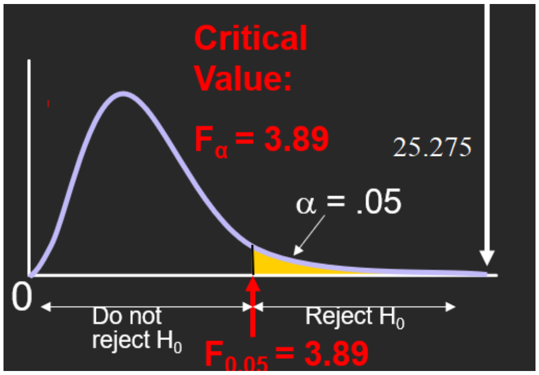
The degrees of freedom for the numerator (c-1) = 3-1 =2

The degrees of freedom for the denominator (n-c) = 15-3 =12

$$F_{(0.05\ ,\ 2,12)} = 3.89$$

Step (5) : Formulate the decision Rule and make a decision

$$F_{STAT}\,(25.275) > F_{(0.05,2,12)}(3.89)$$



Reject Ho at α =0.05

There is evidence that at least one μ_j differs from the rest

Since $F_{stat} > 3.89$, *we reject H_0. And there is sufficient evidence that* at least one μ_j differs .

Reject H_0.There is evidence to conclude that there is a difference in the population means of the groups

								Cumulative Probabilities = 0.95			
								Upper-Tail Areas = 0.05			
								Numerator, df_1			
Denominator, df_2	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**12**
1	161.40	199.50	215.70	224.60	230.20	234.00	236.80	238.90	240.50	241.90	243.90
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69

Example (11) Text book page 386

Consider an experiment with four groups, with eight values in each. For the ANOVA summary table below, fill in all the missing results:

ANOVA

Sources of variation	Degree of freedom	Sum of squares	Mean squares MS.	F-ratio
Among groups	c-1=?	SSA= ?	MSA=80	F_{STAT} =?
Within groups	n-c=?	SSW=560	MSW=?	
Total	n-1=?	SST= ?		

a) At the 0.05 level of significance, state the decision rule for testing the null hypothesis that all four groups have equal population means.

b) What is your statistical decision?

Solution:

Step (1) c-1=4-1=3 , n=4x8=32 , n-c=32-4=28 , n-1=32-1=31

Step (2) $SSA = MSA\ x\ (c-1) = 80x3 = 240$

Step (3) $MSW = \frac{560}{28} = 20$

Step (4) $SST = SSA + SSW = 240 + 560 = 800$

Step (5) $F_{stat} = \frac{MSA}{MSW} = \frac{80}{20} = 4$

Step (1) : State the null and alternate hypotheses :

$H_0 : \mu_1 = \mu_2 = \mu_3$

H_1 : Not all μ_j are equal(Not all the means are equal.).

Step (2): Select the level of significance (α =0.05)

Step (3): The test statistic :

$$F_{STAT} = \frac{MSA}{MSW} = \frac{SSA/c-1}{SSW/n-c} = \frac{240/4-1}{560/32-4} = \frac{\frac{240}{3}}{\frac{560}{32}} = \frac{80}{20} = 4$$

Step (4): The critical value:

The degrees of freedom for the numerator (c-1) = 4-1 =3

The degrees of freedom for the denominator (n-c) = 32-4 =28

$$F_{(0.05\ ,\ 3\ ,28\ )} = 2.95$$

Step (5) : Formulate the decision Rule and make a decision

$$F_{STAT}\ (4) > F_{(0.05\ ,3,28)=}(2.95)$$

$$Reject\ H_0$$

Sources of variation	Degree of freedom	Sum of squares	Mean squares MS.	F-ratio
Among groups	3	SSA= 240	MSA=80	F_{STAT} =4
Within groups	28	SSW=560	MSW=20	
Total	31	SST= 800		

a) At the 0.05 level of significance, state the decision rule for testing the null hypothesis that all four groups have equal population means.

$$Reject\ H_0\ \ if\ F_{stat} > 2.95: otherwise\ do\ not\ reject\ H_0$$

b) What is your statistical decision?

$$Because\ F_{stat} = 4 > 2.95\ \ ,reject\ H_0$$

There is evidence that at least one μ_j differs from the rest

Since $F_{stat} > 2.95$ *,we reject H₀. And there is sufficient evidence that* at least one μ_j differs .

Reject H_0.There is evidence to conclude that there is a difference in the population means of the groups

								Cumulative Probabilities = 0.95			
								Upper-Tail Areas = 0.05			
								Numerator, df_1			
Denominator, df_2	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**12**
1	161.40	199.50	215.70	224.60	230.20	234.00	236.80	238.90	240.50	241.90	243.90
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10