CHAPTER 13

2. From the F table,
   a) For a 5% tail, $F_{4, 18} = 2.928$.
   b) For a 5% tail, $F_{1, 12} = 4.747$.
   c) For a 1% tail, $F_{5, 28} = 3.754$.

4. $\bar{x}_1 = 270$ $\quad\quad$ $\bar{x}_2 = 260$
   $s_1 = 16$ $\quad\quad$ $s_2 = 20$
   $n_1 = 10$ $\quad\quad$ $n_2 = 10$

Step 1: Compute the Within-Groups Variation.

$$SSW = (n_1-1)s_1^2 + (n_2-1)s_2^2$$
$$= (10-1)16^2 + (10-1)20^2$$
$$= (9)256 + (9)400 = 5904$$

Step 2: Compute the Grand Mean for the Samples.

$$\bar{\bar{x}} = \frac{n_1(\bar{x}_1) + n_2(\bar{x}_2)}{n_1 + n_2} = \frac{9(270) + 9(260)}{9+9} = 265$$

Note: Because sample sizes are equal here, we could also compute the grand mean as the simple average of the 2 sample averages:

$$\bar{\bar{x}} = \frac{(270) + (260)}{2} = 265$$

Step 3: Compute the Between-Groups Variation

$$SSB = n_1(\bar{x}_1 - \bar{\bar{x}})^2 + n_2(\bar{x}_2 - \bar{\bar{x}})^2$$
$$= 10(270-265)^2 + 10\,(260-265)^2$$
$$= 250 + 250 = 500$$

Step 4: Compute the variance ratio.

$$F_{stat} = \frac{SSB/(k-1)}{SSW/(n_1+n_2-2)} = \frac{500/(2-1)}{5904/(10+10-2)} = \frac{500}{328} = 1.524$$

Step 5: Compare the variance ratio, F_{stat}, to the critical F value, F_c, from the table.

From the F table for a 5% tail, $F_c = F_{1,\,18} = 4.414$.

Since $F_{stat} < F_c$ (that is, $1.524 < 4.414$) we can't reject the "no difference" null hypothesis. There's not enough sample evidence to convince us that the two population means would be different.

6. Step 1: Compute the Within-Groups Variation.

$$\begin{aligned} SSW &= (n_1-1)s_1^2 + (n_2-1)s_2^2 + (n_3-1)s_3^2 \\ &= (21-1)23.3^2 + (21-1)25.2^2 + (21-1)\,24.7^2 \\ &= (20)542.889 + (20)635.04 + (20)610.90 = 35760.38 \end{aligned}$$

Step 2: Compute the Grand Mean for the Samples.

$$\bar{\bar{x}} = \frac{n_1(\bar{x}_1) + n_2(\bar{x}_2) + n_3(\bar{x}_3)}{n_1 + n_2 + n_3} = \frac{21(122.1) + 21(131.0) + 21(118.3)}{21 + 21 + 21} = 123.8$$

Note: Because sample sizes are equal here, we could also compute the grand mean as the simple average of the 3 sample averages:

$$\bar{\bar{x}} = \frac{(122.1) + (131.0) + (118.3)}{3} = 123.8$$

Step 3: Compute the Between-Groups Variation

$$\begin{aligned} SSB &= n_1(\bar{x}_1 - \bar{\bar{x}})^2 + n_2(\bar{x}_2 - \bar{\bar{x}})^2 + n_3(\bar{x}_3 - \bar{\bar{x}})^2 \\ &= 21(122.1-123.8)^2 + 21(131.0-123.8)^2 + 21(118.3-123.8)^2 \\ &= 21(2.89) + 21(51.84) + 21(30.25) = 1784.58 \end{aligned}$$

Step 4: Compute the variance ratio.

$$F_{stat} = \frac{SSB/(k-1)}{SSW/(n_1 + n_2 + n_3 - 3)} = \frac{1784.58/(3-1)}{35760.38/(21+21+21-3)}$$

$$= \frac{892.29}{596.006} = 1.497$$

Step 5: Compare the variance ratio, F_{stat}, to the critical F value, F_c, from the table.

From the F table for a 5% tail, $F_c = F_{2,\,60} = 3.150$.

Since $F_{stat} < F_c$ (that is, 1.497 < 3.150) we can't reject the "no difference" null hypothesis. There's not enough sample evidence to make the case that at least one of the population means would be different.

8. First compute the mean and standard deviation for each of the three samples:

Procedure 1	Procedure 2	Procedure 3
5	14	10
7	10	16
10	6	12
6	10	6
mean = 7	mean = 10	mean = 11
std dev = 2.160	std dev = 3.266	std dev = 4.163

Next,

Step 1: Compute the Within-Groups Variation.

$$SSW = (n_1-1)s_1^2 + (n_2-1)s_2^2 + \ldots + (n_k-1)s_k^2$$
$$= (4-1)2.160^2 + (4-1)3.266^2 + (4-1)\,4.163^2 = 98.0$$

Step 2: Compute the Grand Mean for the Samples.

$$\bar{\bar{x}} = \frac{n_1(\bar{x}_1) + n_2(\bar{x}_2) + n_3(\bar{x}_3)}{n_1 + n_2 + n_3} = \frac{4(7) + 4(10) + 4(11)}{4+4+4} = 9.33$$

Step 3: Compute the Between-Groups Variation

$$SSB = n_1(\bar{x}_1 - \bar{\bar{x}})^2 + n_2(\bar{x}_2 - \bar{\bar{x}})^2 + \ldots + n_k(\bar{x}_k - \bar{\bar{x}})^2$$
$$= 4(7-9.33)^2 + 4(10-9.33)^2 + 4(11-9.33)^2 = 34.67$$

Step 4: Compute the variance ratio.

$$F_{stat} = \frac{SSB/(k-1)}{SSW/(n_1+n_2+n_3-k)} = \frac{34.67/(3-1)}{98.0/(4+4+4-3)} = \frac{17.34}{10.89} = 1.592$$

Step 5: Compare the variance ratio, F_{stat}, to the critical F value, F_c, from the table.

From the F table for a 5% tail, $F_c = F_{2,9} = 4.256$.

Since $F_{stat} < F_c$ (that is, 1.592 < 4.256) we can't reject the "no difference" null hypothesis. There's not enough sample evidence to convince us that at least one of the population means would be different.

10. Using calculations from Exercise 4:

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit(.05)*
Between Groups	500	1	500	1.524	XXX	4.414
Within Groups	5904	18	328			
Total	6404	19				

12. Using calculations from Exercise 7:

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit(.05)*
Between Groups	184667.68	2	92333.84	.157	XXX	3.35
Within Groups	15874200	27	587933.33			
Total	16058867.68	29				

14. From the chi-square table, a) 18.307 b) 23.209 c) 15.987.

16. From the chi-square table, a) 15.507 b) 15.507 c) 20.090.

18. Step 1: Pool the sample proportions.

$$\bar{p}_{pooled} = \frac{500(.32)+1000(.26)}{500+1000} = .28$$

Step 2: Compute each of the sample z scores.

$$z_1 = \frac{.32-.28}{\sqrt{\frac{.28(1-.28)}{500}}} = 2.0$$

$$z_2 = \frac{.26 - .28}{\sqrt{\frac{.28(1-.28)}{1000}}} = -1.41$$

Step 3: Calculate the sample chi-square value.

$$\chi^2_{stat} = 2.0^2 + (-1.41)^2 = 4.0 + 1.96 = 5.96$$

Step 4: Compare the sample chi-square value to the critical chi-square value from the table.

For 5% significance and *df* = 2 – 1 = 1, the table value is 3.841 (i.e., $\chi^2_c = 3.841$).

Since $\chi^2_{stat} > \chi^2_c$ (that is, 5.96 > 3.841), we can reject the "no difference in population proportions" null hypothesis.

Translation: There is enough sample evidence to make the case the proportion of teachers who hold this opinion is *not* the same for the two teacher populations.

20. Step 1: Pool the sample proportions.

$$\bar{p}_{pooled} = \frac{2130(.72) + 2960(.76) + 1560(.73)}{2130 + 2960 + 1560} = .74$$

Step 2: Compute each of the sample z scores.

$$z_1 = \frac{.72 - .74}{\sqrt{\frac{.74(1-.74)}{2130}}} = \frac{-.02}{.0095} = -2.105$$

$$z_2 = \frac{.76 - .74}{\sqrt{\frac{.74(1-.74)}{2960}}} = \frac{.02}{.008} = 2.5$$

$$z_3 = \frac{.73 - .74}{\sqrt{\frac{.74(1-.74)}{1560}}} = \frac{-.01}{.011} = -.909$$

Step 3: Calculate the sample chi-square value.

$$\chi^2_{stat} = -2.105^2 + 2.5^2 + -.909^2 = 4.431 + 6.25 + .826 = 11.5$$

Step 4: Compare the sample chi-square value to the critical chi-square value from the table.

For 10% significance and $df = 3 - 1 = 2$, the table value is 4.605.

Since 11.5 > 4.605, we can reject the "no difference in population proportions" null hypothesis.

Translation: There is enough sample evidence to convince us that there is a difference in the proportion of seatbelt users in at least one of the three populations represented.

22. OBSERVED FREQUENCIES

SUPPLIER	Too Light	Not Too Light	TOTAL
Allen	12	88	100
Baron	16	84	100
TOTAL	28	172	200

"pooled" sample proportion of "too light" baseballs = $\frac{28}{200} = .14$

"pooled" sample proportion of "not too light" baseballs = $\frac{172}{200} = .86$

EXPECTED FREQUENCIES

SUPPLIER	Too Light	Not Too Light	TOTAL
Allen	.14(100) = 14	.86(100) = 86	100
Baron	.14(100) = 14	.86(100) = 86	100
TOTAL	28	172	200

OBSERVED and EXPECTED FREQUENCIES

SUPPLIER	Too Light	Not Too Light	TOTAL
Allen	12 / 14	88 / 86	100
Baron	16 / 14	84 / 86	100
TOTAL	28	172	200

$$\chi 2_{stat} = \frac{(12-14)^2}{14} + \frac{(16-14)^2}{14} + \frac{(88-86)^2}{86} + \frac{(84-86)^2}{86}$$

$$= .286 + .286 + .046 + .046 = .664$$

From the table, for a right tail area of 5% and $df = (2\text{-}1) \times (2\text{-}1) = 1$, $\chi_c^2 = 3.841$.
Since .664 < 3.841, we can't reject the "both proportions are equal" null hypothesis.

24. OBSERVED FREQUENCIES

COUNTRY	Bankrupt	Non Bankrupt	TOTAL
France	.17(100) = 17	83	100
Germany	.26(100) = 26	74	100
Italy	.28(100) = 28	72	100
TOTAL	71	229	300

"pooled" sample proportion of bankrupt companies = $\frac{71}{300} = .237$

"pooled" sample proportion of non-bankrupt companies = $\frac{229}{300} = .763$

EXPECTED FREQUENCIES

COUNTRY	Bankrupt	Non Bankrupt	TOTAL
France	.237(100) = 23.7	.763(100) = 76.3	100
Germany	.237(100) = 23.7	.763(100) = 76.3	100
Italy	.237(100) = 23.7	.763(100) = 76.3	100
TOTAL	71	229	300

OBSERVED and EXPECTED FREQUENCIES

COUNTRY	Bankrupt	Non Bankrupt	TOTAL
France	17 / 23.7	83 / 76.3	100
Germany	26 / 23.7	74 / 76.3	100
Italy	28 / 23.7	72 / 76.3	100
TOTAL	71	229	300

$$\chi 2_{stat} = \frac{(17-23.7)^2}{23.7} + \frac{(26-23.7)^2}{23.7} + \frac{(28-23.7)^2}{23.7} + \frac{(83-76.3)^2}{76.3} + \frac{(74-76.3)^2}{76.3}$$

$$+ \frac{(72-76.3)^2}{76.3} =$$

$$= 1.88 + .23 + .79 + .58 + .07 + .25 = 3.8$$

From the table, for a right tail area of 5% and $df = (3\text{-}1) \times (2\text{-}1) = 2$, $\chi_c^2 = 5.991$. Since $3.8 < 5.991$, we can't reject the "all proportions are equal" null hypothesis.

26. Step 1: Calculate expected frequencies.

a) $ef(1,1) = \frac{60}{100} \times 64 = 38.4$ b) $ef(1,2) = \frac{40}{100} \times 64 = 25.6$

Expected Frequencies

	Male	Female	Totals
Flavor A	38.4	25.6	64
Flavor B	21.6	14.4	36
Totals	60	40	100

c) $ef(2,1) = \frac{60}{100} \times 36 = 38.4$ d) $ef(2,2) = \frac{40}{100} \times 36 = 14.4$

Step 2: Compute the $\chi 2_{stat}$ summary measure.

Using a combined observed/expected frequency table,

	Male	Female	Totals
Flavor A	36 / 38.4	28 / 25.6	64
Flavor B	24 /21.6	12 / 14.4	36
Totals	60	40	100

$$\chi 2_{stat} = \sum_i \sum_j \frac{(of - ef)^2}{ef}$$

$$= \frac{(36-38.4)^2}{38.4} + \frac{(28-25.6)^2}{25.6} + \frac{(24-21.6)^2}{21.6} + \frac{(12-14.4)^2}{14.4}$$

$$= .15 + .225 + .267 + .4 = 1.042$$

Step 3: Compare the sample chi-square value to the critical chi-square value from the table.

From the table, for a tail area of 5% and $df = (2\text{-}1) \times (2\text{-}1) = 1$, $\chi_c^2 = 3.841$.

Since $1.042 < 3.841$ we can't reject the "independence" null hypothesis at the 5% significance level

Translation: There isn't enough sample evidence to make the case that gender and flavor preference are related.

28. Step 1: Calculate expected frequencies.

a) $ef(1,1) = \frac{140}{200} \times 80 = 56$ b) $ef(1,2) = \frac{60}{200} \times 80 = 24$

Expected *F*requencies

	Normal Grass	High Energy Feed	Totals
Low	56	24	80
Moderate	49	21	70
High	35	15	50
Totals	140	60	200

c) $ef(2,1) = \frac{140}{200} \times 70 = 49$ d) $ef(2,2) = \frac{60}{200} \times 70 = 21$

Step 2: Compute the $\chi 2_{stat}$ summary measure.

Using a combined observed/expected frequency table,

	Normal Grass	High Energy Feed	Totals
Low	70 / 56	10 /24	80
Moderate	50 / 49	20 / 21	70
High	20 / 35	30 /15	50
Totals	140	60	200

$$\chi 2_{stat} = \sum_i \sum_j \frac{(of - ef)^2}{ef}$$

$$= \frac{(70-56)^2}{56} + \frac{(10-24)^2}{24} + \frac{(50-49)^2}{49} + \frac{(20-21)^2}{21} + \frac{(20-35)^2}{35} + \frac{(30-15)^2}{15}$$

$$= 3.5 + 8.167 + .020 + .048 + 6.429 + 15 = 33.163$$

Step 3: Compare the sample chi-square value to the critical chi-square value from the table.

From the table, for a tail area of 1% and $df = (3\text{-}1) \times (2\text{-}1) = 2$, $\chi_c^2 = 9.210$.

Since 33.163 > 9.210, we can reject the "independence" null hypothesis at the 1% significance level

Translation: There is enough sample evidence to make the case that feed type and level of mastitis are related.

30. Step 1: Calculate expected frequencies.

$$ef(\text{i,j}) = \frac{\text{Column } j \text{ Total}}{\text{Grand Total}} \times \text{Row } i \text{ Total}$$

$$ef(1,1) = \frac{230}{500} \times 130 = 59.8 \qquad ef(4,2) = \frac{270}{500} \times 145 = 78.3$$

CLASS	YES	NO	TOTAL
Freshman	59.8	70.2	130
Soph	52.9	62.1	115
Junior	50.6	59.4	110
Senior	66.7	78.3	145
TOTAL	230	270	500

Step 2: Compute the $\chi 2_{stat}$ summary measure.

Using a combined observed/expected frequency table,

CLASS	YES	NO	TOTAL
Freshman	68 / 59.8	62 / 70.2	130
Soph	51 / 52.9	64 / 62.1	115
Junior	52 / 50.6	58 / 59.4	110
Senior	59 / 66.7	86 / 78.3	145
TOTAL	230	270	500

$$\chi 2_{stat} = \sum_i \sum_j \frac{(of - ef)^2}{ef} = \frac{(68-59.8)^2}{59.8} + \frac{(62-70.2)^2}{70.2} + \ldots + \frac{(86-78.3)^2}{78.3}$$

$$= 1.124 + .957 + .068 + .058 + .039 + .033 + .889 + .757 = 3.926$$

Step 3: Compare the sample chi-square value to the critical chi-square value from the table.

From the table, for a 5% tail and $df = (4\text{-}1) \times (2\text{-}1) = 3$, $\chi_c^2 = 7.815$.

Since $3.926 < 7.815$ we can't reject the "independence" null hypothesis at the 1% significance level

Translation: We don't have enough sample evidence to make the case that class and response are related.

32. Step 1: Compute the Within-Groups Variation.

$$SSW = (n_1\text{-}1)s_1^2 + (n_2\text{-}1)s_2^2 + \ldots + (n_k\text{-}1)s_k^2$$
$$= (654\text{-}1)2.34^2 + (72\text{-}1)2.16^2 + (81\text{-}1)1.84^2 + (61\text{-}1)2.37^2$$
$$= (653)\,5.48 + (71)4.67 + (80)3.39 + (60)\,5.62 = 4514.7$$

Step 2: Compute the Grand Mean for the Samples.

$$\bar{\bar{x}} = \frac{n_1(\bar{x}_1) + n_2(\bar{x}_2) + \ldots + n_k(\bar{x}_k)}{n_1 + n_2 + \ldots + n_k}$$
$$= \frac{654(6.198) + 72(8.326) + 81(3.827) + 61(5.426)}{654 + 72 + 81 + 61} = 6.1$$

Step 3: Compute the Between-Groups Variation

$$SSB = n_1(\bar{x}_1 - \bar{\bar{x}})^2 + n_2(\bar{x}_2 - \bar{\bar{x}})^2 + \ldots + n_k(\bar{x}_k - \bar{\bar{x}})^2$$
$$= 654(6.198\text{-}6.1)^2 + 72(8.326\text{-}6.1)^2 + 81(3.827\text{-}6.1)^2 + 61(5.426\text{-}6.1)^2$$
$$= 654(.01) + 72(4.96) + 81(5.16) + 61(.45) = 809.2$$

Step 4: Compute the variance ratio.

$$F_{stat} = \frac{SSB / (k-1)}{SSW / (n_1 + n_2 + \ldots + n_k - k)}$$

$$= \frac{809.2 / (4-1)}{4514.7 / (654 + 72 + 81 + 61 - 4)}$$

$$= \frac{269.7}{5.2} = 51.8$$

Step 5: Compare the variance ratio, F_{stat}, to the critical F value, F_c, from the table.

From the F table for a 5% tail, $F_c = F_{3,\,864} = 2.6$. (Use $F_{3,\,\infty}$.)

Since $F_{stat} > F_c$ (that is, 51.8 < 2.6) we can reject the "no difference" null hypothesis. There's enough sample evidence to convince us that at least one of the population means would be different.

34. First compute the mean and standard deviation for each of the four samples:

G	PG	PG-13	R
80	84	80	94
79	81	89	86
86	89	86	85
82	76	79	82
78	85	91	93
mean = 81	83	85	88
std dev = 3.162	4.848	5.339	5.244

Step 1: Compute the Within-Groups Variation.

$$SSW = (n_1\text{-}1)s_1^2 + (n_2\text{-}1)s_2^2 + \ldots + (n_k\text{-}1)s_k^2$$
$$= (5\text{-}1)3.162^2 + (5\text{-}1)4.848^2 + (5\text{-}1)\,5.339^2 + (5\text{-}1)\,5.244^2$$
$$= (4)10 + (4)23.5 + (4)28.5 + (4)27.5 = 358$$

Step 2: Compute the Grand Mean for the Samples.

$$\bar{\bar{x}} = \frac{n_1(\bar{x}_1) + n_2(\bar{x}_2) + \ldots + n_k(\bar{x}_k)}{n_1 + n_2 + \ldots + n_k}$$
$$= \frac{5(81) + 5(83) + 5(85) + 5(88)}{5+5+5+5} = 84.25$$

Step 3: Compute the Between-Groups Variation

$$SSB = n_1(\bar{x}_1 - \bar{\bar{x}})^2 + n_2(\bar{x}_2 - \bar{\bar{x}})^2 + \ldots + n_k(\bar{x}_k - \bar{\bar{x}})^2$$
$$= 5(81\text{-}84.25)^2 + 5(83\text{-}84.25)^2 + 5(85\text{-}84.25)^2 + 5(88\text{-}84.25)^2$$
$$= 133.75$$

Step 4: Compute the variance ratio.

$$F_{stat} = \frac{SSB / (k-1)}{SSW / (n_1 + n_2 + \ldots + n_k - k)}$$

$$= \frac{133.75/(4-1)}{358/(5+5+5+5-4)}$$

$$= \frac{44.58}{22.37} = 1.99$$

Step 5: Compare the variance ratio, F_{stat}, to the critical F value, F_c, from the table.

From the F table for a 1% tail, $F_c = F_{3,16} = 5.29$.

Since $F_{stat} < F_c$ (that is, $1.99 < 5.29$) we can't reject the "no difference" null hypothesis. There's not enough sample evidence to convince us that at least one of the population means would be different.

36. Since F_{stat} (1.253, as shown in the F column of the ANOVA table) is less than F_c (2.866, as shown in the *F crit* column of the ANOVA table), we can't reject (at the 5% significance level) the null hypothesis that there's "no difference in average number of days until the patient attains 75% lateral rotation" for the three populations represented.

38.

Step 1: Compute the Within-Groups Variation.

$$SSW = (n_1-1)s_1^2 + (n_2-1)s_2^2 + \ldots + (n_k-1)s_k^2$$
$$= (800-1)6.8^2 + (1050-1)6.6^2 + (650-1)\,6.1^2$$
$$= (799)46.24 + (1049)43.56 + (649)37.21 = 106789.5$$

Step 2: Compute the Grand Mean for the Samples.

$$\bar{\bar{x}} = \frac{n_1(\bar{x}_1) + n_2(\bar{x}_2) + \ldots + n_k(\bar{x}_k)}{n_1 + n_2 + \ldots + n_k}$$

$$= \frac{800(67) + 1050(69) + 650(70)}{800 + 1050 + 650} = 68.62$$

Step 3: Compute the Between-Groups Variation

$$SSB = n_1(\bar{x}_1 - \bar{\bar{x}})^2 + n_2(\bar{x}_2 - \bar{\bar{x}})^2 + \ldots + n_k(\bar{x}_k - \bar{\bar{x}})^2$$
$$= 800(67-68.62)^2 + 1050(69-68.62)^2 + 650(70-68.62)^2$$
$$= 800(2.62) + 1050(.14) + 650(1.90)$$
$$= 3489$$

Step 4: Compute the variance ratio.

$$F_{stat} = \frac{SSB/_{k-1}}{SSW/_{(n_1 + n_2 + ... + n_k - k\ )}}$$

$$= \frac{3489/_{3-1}}{106789.5/_{(800 + 1050 + 650 - 3\ )}}$$

$$= \frac{1744.5}{42.77} = 40.79$$

Step 5: Compare the variance ratio, F_{stat}, to the critical F value, F_c, from the table.

From the F table for a 1% tail, $F_c = F_{2,\,2497} = 4.61$. (Use $F_{2,\,\infty}$.)

Since $F_{stat} > F_c$ (that is, 40.79 > 4.61) we can reject the "no difference" null hypothesis. There's enough sample evidence to convince us that at least one of the population means would be different.

40. Defining μ_1 as the average score for all patients who would be given the drug and μ_2 as the average score for all patients who would be given the sugar pill, the hypotheses are:

H_0: $\mu_1 = \mu_2$ or $\mu_1 - \mu_2 = 0$
H_a: $\mu_1 \neq \mu_2$ or $\mu_1 - \mu_2 \neq 0$

a)

Compute the mean and the standard deviation for each of the two samples:

Test Group Sample:

$$\bar{x}_1 = \frac{80 + 75 + 84 + 66 + 70}{5} = 75$$

$$s_1 = \sqrt{\frac{(80-75)^2 + (75-75)2 + (84-75)^2 + (66+75)^2 + (70-75)^2}{5-1}} = 7.28$$

Control Group Sample:

$$\bar{x}_2 = \frac{79+60+66+57+63}{5} = 65$$

$$s_2 = \sqrt{\frac{(84-70)^2+(65-70)^2+(71-70)^2+(62+70)^2+(68-70)^2}{5-1}} = 8.52$$

Now "pool" sample standard deviations:

$$s_{pooled} = \sqrt{\frac{(5-1)(7.28)^2+(5-1)(8.52)^2}{5+5-2}} = 7.92$$

Compute the t-score for the sample mean difference assuming that this sample mean difference comes from a sampling distribution of sample mean differences that is centered on μ_1-μ_2 = 0:

$$t_{stat} = \frac{(\bar{x}_1-\bar{x}_2)-0}{\sqrt{\frac{s_{pooled}^2}{n_1}+\frac{s_{pooled}^2}{n_2}}} = \frac{(75-65)-0}{\sqrt{\frac{62.75}{5}+\frac{62.75}{5}}} = \frac{10}{5.01} = 2.0$$

Compare t_{stat} to t_c, the critical t value for 5% significance and 5 + 5 – 2 = 8 degrees of freedom:

From the t table, the t score for a tail area of .025 (remember this is a two tailed test with α = .05) and 8 degrees of freedom is 2.306.

Conclusion: Since $t_{stat} < t_c$ (2.0 < 2.306), we can't reject the "no difference in population means" null hypothesis.

b) Using the analysis of variance approach:

Step 1: Compute the Within-Groups Variation.

$$\begin{aligned} SSW &= (n_1-1)s_1^2 + (n_2-1)s_2^2 \\ &= (5-1)7.28^2 + (5-1)8.52^2 \\ &= (4)53 + (4)72.5 = 502 \end{aligned}$$

Step 2: Compute the Grand Mean for the Samples.

$$\bar{\bar{x}} = \frac{n_1(\bar{x}_1) + n_2(\bar{x}_2)}{n_1+n_2} = \frac{5(75)+5(65)}{5+5} = 70$$

Step 3: Compute the Between-Groups Variation

$$\text{SSB} = n_1(\bar{x}_1 - \bar{\bar{x}})^2 + n_2(\bar{x}_2 - \bar{\bar{x}})^2 = 5(75\text{-}70)^2 + 5(65\text{-}70)^2 = 250$$

Step 4: Compute the variance ratio.

$$F_{stat} = \frac{\text{SSB}/(k-1)}{\text{SSW}/(n_1+n_2-k)} = \frac{250/(2-1)}{502/(5+5-2)} = \frac{250}{62.75} = 4.0$$

Step 5: Compare the variance ratio, F_{stat}, to the critical F value, F_c, from the table.

From the F table for a 5% tail, $F_c = F_{1,8} = 5.32$.

Conclusion: Since $F_{stat} < F_c$ (that is, 4.0 < 5.32) we can't reject the "no difference" null hypothesis. There's not enough sample evidence to convince us that the population means would be different.

c) $t_{stat} = 2$; and $F_{stat} = 4$. In any two-sample case, F_{stat} will always equal t_{stat}^2 for any given significance level.

d) $t_c = 2.306$; $F_c = 5.32$. In any two-sample case, F_c will always equal t_c^2 for any given significance level.

42. Step 1: Pool the sample proportions.

$$\bar{p}_{pooled} = \frac{50(.76) + 100(.68) + 200(.71)}{50 + 100 + 200} = \frac{248}{350} = .708$$

Step 2: Compute each of the sample z scores.

$$z_1 = \frac{.76 - .708}{\sqrt{\frac{.708(1-.708)}{50}}} = .8$$

$$z_2 = \frac{.68 - .708}{\sqrt{\frac{.708(1-.708)}{100}}} = -.63$$

$$z_3 = \frac{.71 - .708}{\sqrt{\frac{.708(1-.708)}{200}}} = .04$$

Step 3: Calculate the sample chi-square value.

$$\chi^2_{stat} = .8^2 + -.63^2 + .04^2 = 1.04$$

Step 4: Compare the sample chi-square value to the critical chi-square value from the table.

For 5% significance and *df* = 3 – 1 = 2, the table value is 5.991 (i.e., $\chi^2_c = 5.991$).

Since $\chi^2_{stat} < \chi^2_c$ (that is, 1.04 < 5.991), we cannot reject the "no difference in population proportions" null hypothesis.

Translation: There is not enough sample evidence to convince us that among the three populations represented there would be a difference in the proportion of company representatives who would report good long term prospects.

44. Step 1: Pool the sample proportions.

$$\bar{p}_{pooled} = \frac{150(.74) + 150(.68) + 150(.70) + 150(.64)}{150 + 150 + 150 + 150} = .69$$

Note: Since sample sizes are equal, we could have produced the same $\bar{p}_{pooled}$ result simply by adding the four sample proportions and dividing by 4.

Step 2: Compute each of the sample z scores.

$$z_1 = \frac{.74 - .69}{\sqrt{\frac{.69(1-.69)}{150}}} = 1.323 \qquad z_2 = \frac{.68 - .69}{\sqrt{\frac{.69(1-.69)}{150}}} = -.265$$

$$z_3 = \frac{.70 - .69}{\sqrt{\frac{.69(1-.69)}{150}}} = .265 \qquad z_4 = \frac{.64 - .69}{\sqrt{\frac{.69(1-.69)}{150}}} = -1.323$$

Step 3: Calculate the sample chi-square value.

$$\chi^2_{stat} = 1.323^2 + -.265^2 + .265^2 + -1.323^2 = 3.64$$

Step 4: Compare the sample chi-square value to the critical chi-square value from the table.

For 5% significance and $df = 4 - 1 = 3$, the table value is 7.815 (i.e., $\chi^2_c = 7.815$).

Since $\chi^2_{stat} < \chi^2_c$ (that is, $3.64 < 7.815$), we cannot reject the "no difference in population proportions" null hypothesis.

Translation: There is not enough sample evidence to convince us that there would be a difference in the proportion of highly favorable responses among the four populations represented. (That is, we can't show that at least one of the population proportions would be different.)

46. Step 1: Pool the sample proportions

$$\bar{p}_{pooled} = \frac{11706(.84)+8280(.85)+3171(.87)+771(.88)+3539(.88)+1642(.90)}{11706+8280+3171+771+3539+1642}$$

$$= \frac{24900}{29109} = .855$$

Step 2: Compute each of the sample z scores.

$$z_1 = \frac{.84-.855}{\sqrt{\frac{.855(1-.855)}{11706}}} = -4.6 \qquad z_2 = \frac{.85-.855}{\sqrt{\frac{.855(1-.855)}{8280}}} = -1.3$$

$$z_3 = \frac{.87-.855}{\sqrt{\frac{.855(1-.855)}{3171}}} = 2.4 \qquad z_4 = \frac{.88-.855}{\sqrt{\frac{.855(1-.855)}{771}}} = 2.0$$

$$z_5 = \frac{.88-.855}{\sqrt{\frac{.855(1-.855)}{3539}}} = 4.2 \qquad z_6 = \frac{.90-.855}{\sqrt{\frac{.855(1-.855)}{1642}}} = 5.2$$

Step 3: Calculate the sample chi-square value.

$$\chi^2_{stat} = -4.6^2 + -1.3^2 + 2.4^2 + 2.0^2 + 4.2^2 + 5.2^2 = 77.2$$

Step 4: Compare the sample chi-square value to the critical chi-square value from the table.

For 1% significance and df = 6 – 1 = 5, the table value is 11.071 (i.e., χ^2_c = 11.071).

Since $\chi^2_{stat} > \chi^2_c$ (that is, 77.2 > 11.071), we can reject the "no difference in population proportions" null hypothesis.

Translation: There is enough sample evidence to convince us that there would be a difference in proportion of "good or fairly good" responses among the six populations represented. (At least one of the population proportions is different.)

48. OBSERVED FREQUENCIES

SPORT	Yes	No	TOTAL
Football	78	222	300
Basketball	48	152	200
Baseball	36	64	100
TOTAL	162	438	600

"pooled" sample proportion of "yes" responses = $\frac{162}{600} = .27$

"pooled" sample proportion of "no" responses = $\frac{438}{600} = .73$

EXPECTED FREQUENCIES

SPORT	Yes	No	TOTAL
Football	.27(300) = 81	.73(300) = 219	300
Basketball	.27(200) = 54	.73(200) = 146	200
Baseball	.27(100) = 27	.73(100) = 73	100
TOTAL	162	438	600

OBSERVED and EXPECTED FREQUENCIES

SPORT	Yes	No	TOTAL
Football	78 / 81	222 / 219	300
Basketball	48 / 54	152 / 146	200
Baseball	36 / 27	64 / 73	100
TOTAL	162	438	600

$$\chi 2_{stat} = \frac{(78-81)^2}{81} + \frac{(48-54)^2}{54} + \frac{(36-27)^2}{27} + \frac{(222-219)^2}{219} + \frac{(152-146)^2}{146}$$

$$+\frac{(64-73)^2}{73} \quad = .111 + .667 + 3.0 + .041 + .246 + 1.11 = 5.175$$

From the table, for a right tail area of 5% and df = (3-1) x (2-1) = 2, χ_c^2 = 5.991.

Since 5.175 < 5.991, we can't reject the "all proportions are equal" null hypothesis.

50. OBSERVED FREQUENCIES

Operator	Defective	Not Defective	TOTAL
Smith	12	188	200
Roberts	16	184	200
Buckley	18	182	200
Phillips	10	190	200
TOTAL	56	744	800

"pooled" sample proportion defectives = $\frac{56}{800} = .07$

"pooled" sample proportion defective = $\frac{744}{800} = .93$

EXPECTED FREQUENCIES

Operator	Defective	Not Defective	TOTAL
Smith	.07(200) = 14	.93(200) = 186	200
Roberts	.07(200) = 14	.93(200) = 186	200
Buckley	.07(200) = 14	.93(200) = 186	200
Phillips	.07(200) = 14	.93(200) = 186	200
TOTAL	56	744	800

OBSERVED and EXPECTED FREQUENCIES

Operator	Defective	Not Defective	TOTAL
Smith	12 / 14	188 / 186	200
Roberts	16 / 14	184 / 186	200
Buckley	18 / 14	182 / 186	200
Phillips	10 / 14	190 / 186	200
TOTAL	56	744	800

$$\chi 2_{stat} = \frac{(12-14)^2}{14} + \frac{(16-14)^2}{14} + \frac{(18-14)^2}{14} + \frac{(10-14)^2}{14} + \frac{(188-186)^2}{186}$$
$$+ \frac{(184-186)^2}{186} + \frac{(182-186)^2}{186} + \frac{(190-186)^2}{186}$$
$$= .286 + .286 + 1.143 + 1.143 + .021 + .021 + .086 + .086 = 3.072$$

From the table, for a right tail area of 1% and $df = (4\text{-}1) \times (2\text{-}1) = 3$, $\chi_c^2 = 11.345$.
Since 3.072 < 11.345, we can't reject the "all proportions are equal" null hypothesis.

52. Step 1: Calculate expected frequencies.

$$ef(i,j) = \frac{\text{Column } j \text{ Total}}{\text{Grand Total}} \times \text{Row } i \text{ Total}$$

For example,

$$ef(1,1) = \frac{1220}{1550} \times 722 = 568.3 \qquad ef(3,3) = \frac{97}{1550} \times 359 = 22.5$$

Guest Category	Positive	Negative	No Opinion	TOTALS
Business	568.3	108.5	45.2	722
Tourist	369.1	70.5	29.3	449
Weekend Get-a-Way	282.6	54.0	22.5	359
TOTALS	1220	233	97	1550

Step 2: Compute the $\chi2_{stat}$ summary measure.

Combined observed/expected frequency table:

Guest Category	Positive	Negative	No Opinion	TOTALS
Business	561 / 568.3	114 / 108.5	47 / 45.2	722
Tourist	376 / 369.1	68 / 70.5	25 / 29.3	449
Weekend Get-a-Way	283 / 282.6	51 / 54	25 / 22.5	359
TOTALS	1220	233	97	1550

$$\chi2_{stat} = \sum_i \sum_j \frac{(of - ef)^2}{ef} = \frac{(561-568.3)^2}{568.3} + \frac{(114-108.5)^2}{108.5} + \ldots + \frac{(25-22.5)^2}{22.5}$$

$$= .09 + .27 + .07 + .13 + .09 + .64 + 00 + .16 + .29 = 1.75$$

Step 3: Compare the sample chi-square value to the critical chi-square value from the table.

From the table, for a tail area of 1% and $df = (3\text{-}1) \times (3\text{-}1) = 4$, $\chi_c^2 = 13.277$.

Since $1.75 < 13.277$ we cannot reject the "independence" null hypothesis at the 1% significance level

Translation: There is not enough sample evidence to make the case that guest category and reaction to the new program are related.

54. Preliminary step: Since the expected value for the lower right-hand cell (65 and older, definitely no) is less than 5, we'll combine columns 4 and 5 into a single column headed "50 and older" and proceed:

. Step 1: Calculate expected frequencies.

$$ef(i,j) = \frac{\text{Column } j \text{ Total}}{\text{Grand Total}} \times \text{Row } i \text{ Total}$$

For example,

$$ef(1,1) = \frac{164}{834} \times 524 = 103.0 \qquad ef(4,4) = \frac{286}{834} \times 51 = 17.5$$

Response	18-29	30-39	40-49	50 and older	TOTALS
Definitely yes	103.0	118.7	122.5	179.7	524
Probably yes	35.8	41.2	42.6	62.4	182
Probably no	15.1	17.4	18.0	26.4	77
Definitely no	10.0	11.6	11.9	17.5	51
TOTALS	164	189	195	286	834

Step 2: Compute the $\chi2_{stat}$ summary measure.

Using a combined observed/expected frequency table,

Response	18-29	30-39	40-49	50 and older	TOTS
Definite yes	90 / 103	113 / 118.7	125 / 122.5	196/179.7	524
Prob yes	51 / 35.8	47 / 41.2	36 / 42.6	48 / 62.4	182
Prob no	17 / 15.1	11 / 17.4	21 / 18	28 / 26.4	77
Definite no	6 / 10	18 / 11.6	13 / 11.9	14 / 17.5	51
TOTALS	164	189	195	286	834

$$\chi2_{stat} = \sum_i \sum_j \frac{(of - ef)^2}{ef} = \frac{(90-103)^2}{103} + \frac{(113-118.7)^2}{118.7} + \ldots + \frac{(14-17.5)^2}{17.5}$$

$$= 1.7 + .28 + .05 + 1.48 + 6.46 + .8 + 1.01 + 3.33 + .23 + 2.38 + .5 + .1 + 1.62 + 3.59 + .1 + .7 = 24.3$$

Step 3: Compare the sample chi-square value to the critical chi-square value from the table.

From the table, for a tail area of 1% and *df* = (3-1) x (3-1) = 4, $\chi_c^2 = 21.666$.

Since 24.3 > 21.666 we can reject the “independence” null hypothesis at the 1% significance level

Translation: There is enough sample evidence to make the case that age and survey response are related.

56. First compute the observed frequencies as follows:

OBSERVED FREQUENCIES
Response

Program	Highly Favorable	Not Highly Favorable	TOTALS
A	150(.74) = 111	39	150
B	150(.68) = 102	48	150
C	150(.70) = 105	45	150
D	150(.64) = 96	54	150
TOTALS	414	186	600

Then,

Step 1: Calculate expected frequencies.

$$ef(\text{i,j}) = \frac{\text{Column } j \text{ Total}}{\text{Grand Total}} \times \text{Row } i \text{ Total}$$

For example,

$$ef(1,1) = \frac{414}{600} \times 150 = 103.5 \qquad ef(4,2) = \frac{186}{600} \times 150 = 46.5$$

Program	Highly Favorable	Not Highly Favorable	TOTALS
A	103.5	46.5	150
B	103.5	46.5	150
C	103.5	46.5	150
D	103.5	46.5	150
TOTALS	414	186	600

Step 2: Compute the $\chi2_{stat}$ summary measure.

Using a combined observed/expected frequency table,

Program	Highly Favorable	Not Highly Favorable	TOTALS
A	111 / 103.5	39 / 46.5	150
B	102 / 103.5	48 / 46.5	150
C	105 / 103.5	45 / 46.5	150
D	96 / 103.5	54 / 46.5	150
TOTALS	414	186	600

$$\chi 2_{stat} = \sum_i \sum_j \frac{(of-ef)^2}{ef} = \frac{(111-105.5)^2}{105.5} + \frac{(39-46.5)^2}{46.5} + \ldots + \frac{(54-46.5)^2}{46.5}$$

$$= .543 + 1.21 + .022 + .048 + .022 + .048 + .543 + 1.21 = 3.64$$

Step 3: Compare the sample chi-square value to the critical chi-square value from the table.

For 5% significance and $df = (4-1)\text{x}(2-1) = 3$, the table value is 7.815 (i.e., $\chi^2_c = 7.815$).

Since $\chi^2_{stat} < \chi^2_c$ (that is, $3.64 < 7.815$), we cannot reject the "independence" null hypothesis at the 5% significance level

Translation: There is not enough sample evidence to make the case that response is related to program—which is to say, we can't use this sample to make the case that the response would be different for at least one of the programs.

The results are *identical* to the test shown in Exercise 27, where we couldn't reject the "all proportions are equal" null hypothesis.

58. a. Let π_1 be the proportion of completely satisfied customers in the population of Swedish car buyers
and π_2 be the proportion of completely satisfied customers in the population of Korean car buyers

The hypotheses are:

H_0: $\pi_1 - \pi_2 = 0$ (There's no difference in the two population proportions.)
H_1: $\pi_1 - \pi_2 \neq 0$ (There is a difference in the two population proportions.)

The sample proportions here are:

$$\bar{p}_1 = \frac{160}{200} = .80 \qquad \bar{p}_1 = \frac{105}{150} = .70$$

Step 1: Pool the sample proportions.

$$\bar{p}_{pooled} = \frac{200(.80)+(150)(.70)}{200+150} = .757$$

Step 2: Calculate the standard error for the null sampling distribution.

$$s_{\bar{p}_1-\bar{p}_2} = \sqrt{\frac{.757(.243)}{200}+\frac{.757(.243)}{150}} = .046$$

Step 3: Compute the test statistic.

$$z_{stat} = \frac{(.80-.70)-0}{.046} = 2.17$$

Step 4: Compare z_{stat} to z_c.

From the normal table, with a significance level of 5% (two-tailed-test),

$$z_c = 1.96$$

Conclusion: Since 2.17 > 1.96, we can reject the "no difference" null hypothesis. Sample evidence is strong enough to convince us that the population proportions would be different.

b. Step 1: Pool the sample proportions.

$$\bar{p}_{pooled} = \frac{200(.80)+(150)(.70)}{200+150} = .757$$

Step 2: Compute each of the sample z scores.

$$z_1 = \frac{.80-.757}{\sqrt{\frac{.757(1-.757)}{200}}} = 1.43 \qquad z_2 = \frac{.70-.757}{\sqrt{\frac{.757(1-.757)}{150}}} = -1.63$$

Step 3: Calculate the sample chi-square value.

$$\chi^2_{stat} = 1.43^2 + -1.63^2 = 4.7$$

Step 4: Compare the sample chi-square value to the critical chi-square value from the table.

For 5% significance and $df = 2 - 1 = 1$, the table value is 3.841 (i.e., $\chi^2_c = 3.841$).

Conclusion: Since 4.7 > 3.841, we can reject the "no difference" null hypothesis. Sample evidence is strong enough to convince us that the population proportions would be different. This conclusion is perfectly consistent with the conclusion we reached in part a). In fact, this sort of consistency will always be the case since it will always be true that

$$\chi^2_{stat} = z_{stat}^2 \quad \text{and} \quad \chi^2_c = z_c^2 \qquad \text{when } df = 1.$$

In this example, $z_{stat}^2 = 2.17^2 = 4.7 = \chi^2_{stat}$ (approx) and

$$z_c^2 = 1.96^2 = 3.84 = \chi^2_c$$

c. Show the observed frequencies as follows:

OBSERVED FREQUENCIES
Promptness of Payment

Make of Car	Completely Satisfied	Not Completely Satisfied	TOTALS
Swedish	160	40	200
Korean	105	45	150
TOTALS	265	85	350

Step 1: Calculate expected frequencies.

$$ef(i,j) = \frac{\text{Column } j \text{ Total}}{\text{Grand Total}} \times \text{Row } i \text{ Total}$$

For example,

$$ef(1,1) = \frac{265}{350} \times 200 = 151.4 \qquad ef(2,2) = \frac{85}{350} \times 150 = 36.4$$

Make of Car	Completely Satisfied	Not Completely Satisfied	TOTALS
Swedish	151.4	48.6	200
Korean	113.6	36.4	150
TOTALS	265	85	350

Step 2: Compute the $\chi 2_{stat}$ summary measure.

Using a combined observed/expected frequency table,

Make of Car	Completely Satisfied	Not Completely Satisfied	TOTALS
Swedish	160 / 151.4	40 / 48.6	200
Korean	105 / 113.6	45 / 36.4	150
TOTALS	265	85	350

$$\chi 2_{stat} = \sum_i \sum_j \frac{(of - ef)^2}{ef}$$

$$= \frac{(160-151.4)^2}{151.4} + \frac{(40-48.6)^2}{48.6} + \frac{(105-113.6)^2}{113.6} + \frac{(45-36.4)^2}{36.4}$$

$$= .49 + 1.51 + .65 + 2.02 = 4.7$$

Step 3: Compare the sample chi-square value to the critical chi-square value from the table.

For 5% significance and $df = (2 - 1) \times (2 - 1) = 1$, the table value is 3.841 (i.e., $\chi^2_c = 3.841$).

Conclusion: Since 4.7 > 3.841, we can reject the "independence" null hypothesis. Sample evidence is strong enough to convince us that satisfaction is related to make of car purchased. This conclusion is perfectly consistent with the conclusion we reached in parts a) and b).