CHAPTER 13

Section 13.1

1. We reject H_0 if the calculated χ^2 value is greater than or equal to the tabled value of $\chi^2_{\alpha,k-1}$ from Table A.6.

a. Since $12.25 \geq \chi^2_{.05,4} = 9.488$, we would reject H_0.

b. Since $8.54 < \chi^2_{.01,3} = 11.344$, we would fail to reject H_0.

c. Since $4.36 < \chi^2_{.10,2} = 4.605$, we would fail to reject H_0.

d. Since $10.20 < \chi^2_{.01,5} = 15.085$, we would fail to reject H_0.

3. Using the number 1 for business, 2 for engineering, 3 for social science, and 4 for agriculture, let p_i = the true proportion of all clients from discipline i. If the Statistics Department's expectations are correct, then the relevant null hypothesis is $H_0 : p_1 = .40, p_2 = .30, p_3 = .20, p_4 = .10$, versus H_a: The Statistics Department's expectations are not correct. With df = k – 1 = 4 – 1 = 3, we reject H_0 if $\chi^2 \geq \chi^2_{.05,3} = 7.815$. Using the proportions in H_0, the expected number of clients are:

Client's Discipline	Expected Number
Business	(120)(.40) = 48
Engineering	(120)(.30) = 36
Social Science	(120)(.20) = 24
Agriculture	(120)(.10) = 12

Since all the expected counts are at least 5, the chi-squared test can be used. The value of the test statistic is $\chi^2 = \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i} = \sum_{\text{all cells}} \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$

$= \left[\frac{(52-48)^2}{48} + \frac{(38-36)^2}{36} + \frac{(21-24)^2}{24} + \frac{(9-12)^2}{12} \right] = 1.57$, which is < 7.815, so we fail to reject H_0.

Alternatively, P-value = $P(\chi^2 \geq 1.57) > .10$, and since the P-value is not < .05, we do not reject H_0. Thus we have no significant evidence to suggest that the statistics department's expectations are incorrect.

5. We will reject H_0 if the P-value < .10. The observed values, expected values, and corresponding χ^2 terms are:

Obs	4	15	23	25	38	21	32	14	10	8
Exp	6.67	13.33	20	26.67	33.33	33.33	26.67	20	13.33	6.67
χ^2 contr.	1.069	.209	.450	.105	.654	.163	1.065	1.800	.832	.265

$\chi^2 = 1.069 + \cdots + .265 = 6.612$. With df = 10 – 1 = 9, our χ^2 value of 6.612 is less than $\chi^2_{.10,9} = 14.684$, so the P-value > .10 and we cannot reject H_0. There is no significant evidence that the data is not consistent with the previously determined proportions.

7. Let p_i = the true proportion of sales on the ith day of the week (1 = Monday, … , 5 = Friday). The hypotheses are H_0: $p_1 = p_2 = p_3 = p_4 = p_5 = .2$ (all the same) vs H_a: not all p_i's equal .2. The expected count for each day of the week is 92(.2) = 18.4, so the test statistic value is $\chi^2 = \frac{(22-18.4)^2}{18.4} + \cdots + \frac{(24-18.4)^2}{18.4} = 4.41$. Since 4.41 < 7.779 = $\chi^2_{.10,5-1}$, H_0 cannot be rejected at the .10 level (equivalently, the P-value > .10). The data do not provide convincing evidence that sales are not evenly distributed throughout the week.

9.

a. Denoting the 5 intervals by $[0, c_1)$, $[c_1, c_2)$, …, $[c_4, \infty)$, we wish c_1 for which $.2 = P(0 \le X \le c_1) = \int_0^{c_1} e^{-x}dx = 1 - e^{-c_1}$, so $c_1 = -\ln(.8) = .2231$. Then $.2 = P(c_1 \le X \le c_2) \Rightarrow .4 = P(0 \le X_1 \le c_2) = 1 - e^{-c_2}$, so $c_2 = -\ln(.6) = .5108$. Similarly, $c_3 = -\ln(.4) =$ .0163 and $c_4 = -\ln(.2) = 1.6094$. the resulting intervals are [0, .2231), [.2231, .5108), [.5108, .9163), [.9163, 1.6094), and [1.6094, ∞).

b. Each expected cell count is 40(.2) = 8, and the observed cell counts are 6, 8, 10, 7, and 9, so

$$\chi^2 = \left[\frac{(6-8)^2}{8} + \ldots + \frac{(9-8)^2}{8}\right] = 1.25$$

Because 1.25 is not ≥ $\chi^2_{.10,4} = 7.779$, even at level .10 H_0 cannot be rejected; the data is quite consistent with the specified exponential distribution.

11.

a. The six intervals must be symmetric about 0, so denote the 4th, 5th and 6th intervals by $[0, a)$, $[a, b)$, $[b, \infty)$. The value a must be such that $\Phi(a) = .6667\left(\frac{1}{2} + \frac{1}{6}\right)$, which from Table A.3 gives $a \approx .43$. Similarly, $\Phi(b) = .8333$ implies $b \approx .97$, so the six intervals are (–∞, –.97), [–.97, –.43), [–.43, 0), [0, .43), [.43, .97), and [.97, ∞).

b. The six intervals are symmetric about the mean of .5. From **a**, the fourth interval should extend from the mean to .43 standard deviations above the mean, i.e., from .5 to .5 + .43(.002), which gives [.5, .50086). Thus the third interval is [.5 – .00086, .5) = [.49914, .5). Similarly, the upper endpoint of the fifth interval is .5 + .97(.002) = .50194, and the lower endpoint of the second interval is .5 – .00194 = .49806. The resulting intervals are (–∞, .49806), [.49806, .49914), [.49914, .5), [.5, .50086), [.50086, .50194), and [.50194, ∞).

c. Each expected count is 45(1/6) = 7.5, and the observed counts are 13, 6, 6, 8, 7, and 5, so $\chi^2 = 5.53$. With 5 df, the P-value > .10, so we would fail to reject H_0 at any of the usual levels of significance. There is no significant evidence to suggest that the bolt diameters are *not* normally distributed with $\mu = .5$ and $\sigma = .002$.

13.

a. Let θ denote the probability of a male (as opposed to female) birth under the binomial model. The four cell probabilities (corresponding to x = 0, 1, 2, 3) are $\pi_1(\theta) = (1-\theta)^3$, $\pi_2(\theta) = 3\theta(1-\theta)^2$, $\pi_3(\theta) = 3\theta^2(1-\theta)$, and $\pi_4(\theta) = \theta^3$. The likelihood is $3^{n_2+n_3} \cdot (1-\theta)^{3n_1+2n_2+n_3} \cdot \theta^{n_2+2n_3+3n_4}$. Forming the log likelihood, taking the derivative with respect to θ, equating to 0, and solving yields $\hat{\theta} = \dfrac{n_2+2n_3+3n_4}{3n} = \dfrac{66+128+48}{480} = .504$. The estimated expected counts are $160(1-.504)^3 = 19.52$, $480(.504)(.496)^2 = 59.52$, 60.48, and 20.48, so

$$\chi^2 = \left[\frac{(14-19.52)^2}{19.52} + \ldots + \frac{(16-20.48)^2}{20.48}\right] = 1.56 + .71 + .20 + .98 = 3.45$$

. The number of degrees of freedom for the test is 4 – 1 – 1 = 2. H_0 of a binomial distribution will be rejected using significance level .05 if $\chi^2 \ge \chi^2_{.05,2} = 5.992$. Because 3.45 < 5.992, H_0 is *not* rejected, and the binomial model is judged to be plausible.

b. Now $\hat{\theta} = \dfrac{53}{150} = .353$ and the estimated expected counts are 13.54, 22.17, 12.09, and 2.20. The last estimated expected count is much less than 5, so the chi-squared test based on 2 df should not be used.

15. The mle of μ is $\hat{\mu} = \bar{x} = \dfrac{24(0)+16(1)+\cdots+12(1)}{24+16+\cdots+1} = \dfrac{380}{120} = 3.167$, so under the null hypothesis of a Poisson-distributed population, the estimated probabilities are $\hat{p}(x) = \dfrac{e^{-3.167}3.167^x}{x!}$. We apply this formula for x = 0, 1, 2, 3, 4, 5, 6, and then use 1 – [sum of these probs.] to estimate $P(X \ge 7)$. The expected count for each of the categories (other than the last one) is $n\hat{p}(x) = 120\hat{p}(x)$; the last expected count is 120 – [sum of other expected counts].

x	0	1	2	3	4	5	6	≥ 7
$\hat{p}(x)$	.0421	.1334	.2113	.2230	.1766	.1119	.0590	.0427
Exp.	5.05	16.00	25.36	26.76	21.19	13.43	7.08	5.12
Obs.	24	16	16	18	15	9	6	16

The resulting test statistic value is χ^2 = 103.98, and when compared to either $\chi^2_{.01,8-1} = 18.474$ or $\chi^2_{.01,8-1-1} = 16.812$, it is obvious that H_0 is strongly rejected and the Poisson model fits very poorly.

17. Using the $k = 9$ categories suggested, we will ultimately reject H_0: the data follow a NB distribution if $\chi^2 \geq \chi^2_{.10,9-1-2} = \chi^2_{.10,6} = 10.645$. For $x = 0, \ldots, 6$, the estimated probability is based on $nb(x;\hat{r},\hat{p})$; for the next-to-last category, the estimated probability is $nb(7;\hat{r},\hat{p}) + nb(8;\hat{r},\hat{p})$; and the estimated probability for the last category is 1 – [sum of other probabilities].

x	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7-8**	**≥ 9**
Obs.	17	15	17	19	9	5	4	8	9
Prob.	.1785	.1685	.1420	.1149	.0911	.0712	.0553	.0753	.1032
Exp.	18.39	17.36	14.63	11.83	9.38	7.33	5.70	7.76	10.63

The observed test statistic value is $\chi^2 = 6.668 < 10.645$, so H_0 is not rejected at the .10 level. The data do not provide convincing evidence that a generalized NB distribution is inappropriate for this data.

Section 13.2

19. The tax holiday survey data from the earlier exercises is displayed below. Expected counts, calculated as (row total)(column total)/(grand total) are in parentheses.

	Tax Holiday Is Important?	
	Yes	**No**
Men	195 (203.24)	55 (46.76)
Women	370 (361.76)	75 (83.24)

The test statistic value is $\chi^2 = \frac{(195-203.24)^2}{203.24} + \cdots + \frac{(75-83.24)^2}{83.24} = 2.788$. At df = (2 – 1)(2 – 1) = 1, the P-value is $P(\chi^2_1 \geq 2.788) = .095$ from software. Hence, we'd reject H_0 at the .10 level but not the .05 level. In Chapter 10 Exercise 60(b), the test statistic value was $z = 1.67$, and $1.67^2 \approx 2.788$ (other than a little rounding error). Also, the *one*-tailed P-value in that exercise was .047, so the two-sided P-value is 2(.047) = .094, essentially identical to the chi-squared P-value here (again, the only reason they're not literally identical is rounding error).

21. For $i = 1, 2, 3, 4, 5$ and $j = 1, 2$, let p_{ij} = the proportion of all Italian 12-year-olds that would fall into cell (i, j). The null hypothesis is H_0: $p_{ij} = p_{i\cdot}p_{\cdot j}$ for all i and j, meaning the two underlying variables are independent.

a. The accompanying table shows the observed and estimated expected counts. The resulting chi-squared value is $\chi^2 = 4.504$; at df = (5 – 1)(2 – 1) = 4, this is < $\chi^2_{.05,4} = 9.488$, so H_0 is not rejected. The data do *not* provide convincing evidence that presence/absence of cavities is associated with number of times children brush their teeth in the population of Italian 12-year-olds!

	Cavities	No cavities	All
Never	11	7	18
	9.52	8.48	
Once a day	24	21	45
	23.80	21.20	
2 times a day	99	77	176
	93.10	82.90	
3 times a day	107	117	224
	118.49	105.51	
4 times a day	42	30	72
	38.09	33.91	
All	283	252	535

b. Given the fairly large sample size, it's unlikely (though technically possible) that we've committed a type II error. Perhaps the most likely explanation is that children are not reliable reporters of how often they brush their teeth — in that case, how often children *say* they brush their teeth daily may be unrelated to how often they *actually* do. Perhaps the latter is indeed related to cavities.

23. Let p_{1j} = the true proportion of Canadian women whose typical number of vacation days falls into the jth category in the table ($j = 1, \ldots , 7$), and define p_{2j} similarly for men. The hypotheses are H_0: $p_{1j} = p_{2j}$ for all j versus H_a: H_0 is not true.
The accompanying table shows the observed and estimated expected counts, from which $\chi^2 = 9.858$. At df = $(2 - 1)(7 - 1) = 6$, $9.858 < \chi^2_{.05,6} = 12.592$, so H_0 is not rejected at the .05 level. The data do *not* provide convincing statistical evidence that the distribution of the number of vacation days taken is different for Canadian women and men.

	None	1-5	6-10	11-15	16-20	20-25	>25	All
Female	42	25	79	94	70	58	79	447
	42.95	21.24	67.42	94.66	65.11	64.65	90.97	
Male	51	21	67	111	71	82	118	521
	50.05	24.76	78.58	110.34	75.89	75.35	106.03	
All	93	46	146	205	141	140	197	968

25. We need counts for each class/response category. For instance, .204(3168) = 646.272, so presumably 646 of the 3168 lower-class students said they talk frequently with faculty outside class. It follows that 3168 – 646 = 2522 such students said they didn't. Similar calculations result in the following contingency table.

	Talk frequently w/ faculty outside of class?	
	Yes	**No**
Lower-class	646	2522
Middle-class	3120	13654
Upper-class	1654	6534

a. Let p_{ij} = the proportion of *all* college students who would fall into the (i, j)th cell of the table above (for i = 1, 2, 3 and j = 1, 2). The hypotheses are H_0: $p_{ij} = p_{i.}p_{.j}$ vs H_a: H_0 is not true. Note that this is actually a test of independence, because students were cross-classified; the researchers did not contact independent random samples of lower-, middle-, and upper-class students. With the aid of software, the test statistic and P-value are χ^2 = 11.954 at 2 df and P-value = .003. Hence, at most reasonable significance levels, we conclude that there *is* an association between economic class and whether a student speaks frequently to faculty outside the classroom.

b. With a total sample size of 28,130, even the slightest deviation from perfect equality will be declared statistically significant. Although the conditional proportions of 20.4%, 18.6%, 20.2% might not sound *practically* different, the null hypothesis was still rejected.

27. Under the null hypothesis, we compute estimated expected cell counts by

$\hat{e}_{ijk} = n\hat{p}_{ijk} = n\hat{p}_{i..}\hat{p}_{.j.}\hat{p}_{..k} = n\dfrac{n_{i..}}{n}\dfrac{n_{.j.}}{n}\dfrac{n_{..k}}{n} = \dfrac{n_{i..}n_{.j.}n_{..k}}{n^2}$. This is a 3 x 3 x 4 situation, so there are 36 cells. Only the total sample size, n, is fixed in advance of the experiment, so there are 35 freely determined cell counts. We must estimate all of $p_{1..}, p_{2..}, p_{3..}, p_{.1.}, p_{.2.}, p_{.3.}, p_{..1}, p_{..2}, p_{..3}, p_{..4}$, but $\Sigma p_{i..} = \Sigma p_{.j.} = \Sigma p_{..k} = 1$, so only 2 + 2 + 3 = 7 freely-determined parameters are estimated. The general df rule now gives df = 36 – 1 – 7 = 28.

In general, the degrees of freedom for independence in an I x J x K array equals
$(IJK - 1) - [(I - 1) + (J - 1) + (K - 1)] = IJK - (I+J+K) + 2$.

29.

a.

Observed

13	19	28	60
7	11	22	40
20	30	50	100

Estimated Expected

12	18	30
8	12	20

$\chi^2 = \dfrac{(13-12)^2}{12} + \ldots + \dfrac{(22-20)^2}{20} = .6806$. Because $.6806 < \chi^2_{.10,2} = 4.605$, H_0 is not rejected.

b. Each observation count here is 10 times what it was in **a**, and the same is true of the estimated expected counts, so now $\chi^2 = 6.806 \geq 4.605$, and H_0 is rejected. With the much larger sample size, the departure from what is expected under H_0, the independence hypothesis, is statistically significant – it cannot be explained just by random variation.

c. The observed counts are .13n, .19n, .28n, .07n, .11n, .22n, whereas the estimated expected are .12n, .18n, .30n, .08n, .12n, .20n, yielding $\chi^2 = .006806n$. H_0 will be rejected at level .10 iff $.006806n \geq 4.605$, i.e., iff $n \geq 676.6$, so the minimum n = 677.

31.

a. Minitab output for the test of gender homogeneity among ranks appears below. We see the test statistic is χ^2 = 6.454, with a P-value of 0.04 at 2 df. Thus, at the α = .05 level, we reject the null hypothesis and conclude that the proportion of Professors, Associate Professors, and Assistant Professors who are female are different.

```
          M       F  Total
 1       25       9     34
     21.42   12.58
     0.598   1.019

 2       20       8     28
     17.64   10.36
     0.316   0.538

 3       18      20     38
     23.94   14.06
     1.474   2.510

Total      63      37    100

Chi-Sq = 6.454, DF = 2, P-Value = 0.040
```

b. Minitab output for the appropriate logistic regression (with Success = Female, 1 = Professor, 2 = Assoc. Prof., 3 = Asst. Prof.) is shown below. From the output, the test of H_0: $\beta_1 = 0$ yields z = 2.29 with a P-value 0.022. Thus, again we reject the hypothesis that rank and gender are unrelated. In particular, the output shows that the odds a faculty member is female increases significantly as we continue <u>down</u> the table (i.e., down the ranks from Professor to Assistant).

```
Logistic Regression Table

                                                   Odds     95% CI
Predictor       Coef   SE Coef       Z       P  Ratio  Lower  Upper
Constant   -1.76732  0.593645  -2.98  0.003
C8         0.589227  0.257333   2.29  0.022   1.80   1.09   2.98
```

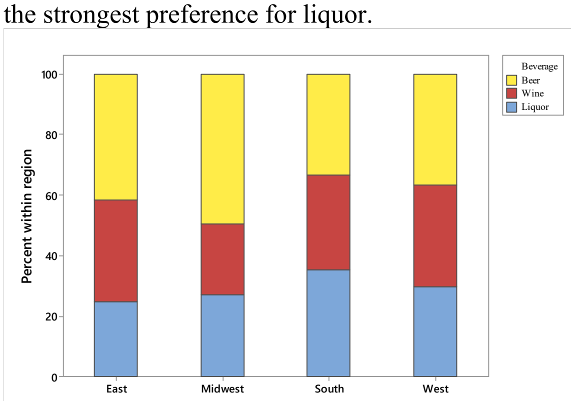
c. Yes: with the extra assumption of order among the factor, we anticipate the P-value from logistic regression will be lower than the P-value from the chi-squared test.

d. The gender imbalance among faculty is less pronounced in lower ranks and more pronounced in higher ranks. This suggests more female faculty are being hired than before. If all these faculty remain at the university, in 10-15 years' time we will see noticeably less gender imbalance among the full professors.

Supplementary Exercises

33. Among the population of all U.S. college students with credit cards, let p_{ij} = the true proportion that would fall into the (i, j)th cell of the table; e.g., (1, 1) = (pay off full balance each month, male). The hypotheses are $H_0 : p_{ij} = p_{i.}p_{.j}$ for all i, j versus H_a: H_0 is not true. With the aid of software, the test statistic value is $\chi^2 = 5.934$. At df = (3 – 1)(2 – 1) = 2, the P-value is .059, just barely above .05. Equivalently, $\chi^2_{.05,2}$ = 5.992, and 5.934 < 5.992 (again, barely). So, at the .05 level, we (barely) fail to reject H_0. At the .05 level, the data do not provide convincing statistical evidence that credit card payoff habits depend upon students' sex.

35. Let p_{ij} = the true proportion of adults in region i (1 = East, 2 = Midwest, etc.) whose favorite alcoholic beverage is j (1 = liquor, 2 = wine, 3 = beer). We test $H_0 : p_{1j} = p_{2j} = p_{3j} = p_{4j}$ for j = 1, 2, 3. With the aid of software, the test statistic value is $\chi^2 = 29.775 \geq \chi^2_{.05,(4-1)(3-1)} = 12.592$. So, H_0 is rejected at the .05 level. The data provide convincing statistical evidence that adults beverage preferences indeed vary by region. As seen in the graph below, Midwesterners have the strongest preference for beer, while Southerners have the strongest preference for liquor.

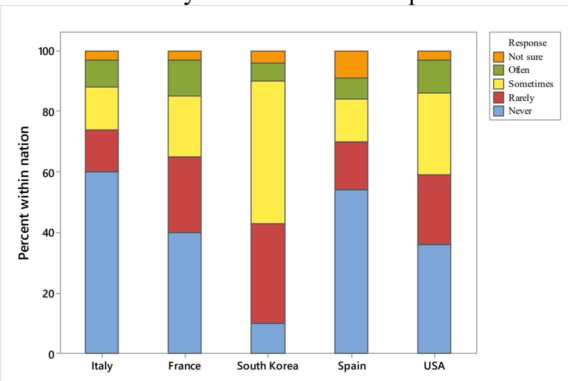


37.

a. H_0: The proportion of Late Game Leader Wins is the same for all four sports; H_a: The proportion of Late Game Leader Wins is not the same for all four sports. With 3 df, the computed $\chi^2 = 10.518$, and the P-value < .015 < .05, so reject H_0. There appears to be a relationship between the late-game leader winning and the sport played.

b. Quite possibly: Baseball had many fewer late-game leader losses than expected.

39. Let p_{ij} = the true proportion of people in the ith country (1 = Italy, 2 = France, etc.) whose view on torturing terror suspects is category j (1 = Never, 2 = Rarely, etc.). The null hypothesis for this homogeneity test is $H_0 : p_{1j} = p_{2j} = p_{3j} = p_{4j} = p_{5j}$ for j = 1, 2, 3, 4, 5. With the aid of software, χ^2 = 881.36; even at df = (5 – 1)(5 – 1) = 16, the P-value is effectively zero. In particular, P-value < .01, and we resoundingly reject H_0 in favor of the alternative hypothesis that attitudes on torture differed in 2005 in these five nations. The accompanying segmented bar graph shows that attitudes in the European nations were quite similar and strongly disposed against torturing terror suspects. USA respondents were more amenable to torture than the Europeans, while South Korean respondents were vastly more likely than anyone else to say it's "sometimes" okay to torture terror suspects.



41.

a.

obs.	22	10	5	11
exp.	13.189	10	7.406	17.405

H_0: probabilities are as specified.
H_a: probabilities are not as specified.

Test Statistic: $\chi^2 = \dfrac{(22-13.189)^2}{13.189} + \dfrac{(10-10)^2}{10} + \dfrac{(5-7.406)^2}{7.406} + \dfrac{(11-17.405)^2}{17.405}$

$= 5.886 + 0 + 0.782 + 2.357 = 9.025$. Rejection Region: $\chi^2 > \chi^2_{.05,3} = 7.815$.

Since 9.025 > 7.815, we reject H_0. The model postulated in the exercise is not a good fit.

b.

p_i	0.45883	0.18813	0.11032	0.24272
exp	22.024	9.03	5.295	11.651

$$\chi^2 = \frac{(22-22.024)^2}{22.024} + \frac{(10-9.03)^2}{9.03} + \frac{(5-5.295)^2}{5.295} + \frac{(11-11.651)^2}{11.651}$$

$= .0000262 + .1041971 + .0164353 + .0363746 = .1570332$

With the same rejection region as in a, we do not reject the null hypothesis. This model does provide a good fit.

43.

a. H_0: $p_0 = p_1 = \cdots = p_9 = .10$ vs H_a: at least one $p_i \neq .10$, with df = 9.

b. H_0: $p_{ij} = .01$ for i and $j = 0,1,2,\ldots,9$ vs H_a: at least one $p_{ij} \neq .01$, with df = 99.

c. For this test, the number of p's in the Hypothesis would be $10^5 = 100{,}000$ (the number of possible combinations of 5 digits). Using only the first 100,000 digits in the expansion, the number of non-overlapping groups of 5 is only 20,000. We need a much larger sample size.

d. Based on these P-values, we could conclude that the digits of π behave as though they were randomly generated!