14.11a. $X^2 = \sum_{j=1}^{c} \sum_{i=1}^{r} \frac{[n_{ij} - E(\hat{n}_{ij})]^2}{E(\hat{n}_{ij})} = \sum_{j=1}^{c} \sum_{i=1}^{r} \left[\frac{\left(n_{ij} - \frac{r_i c_j}{n}\right)^2}{\frac{r_i c_j}{n}} \right]$

$= n \sum_{j=1}^{c} \sum_{i=1}^{r} \left[\frac{n_{ij}^2 - \frac{2n_{ij} r_i c_j}{n} + \frac{r_i^2 c_j^2}{n^2}}{r_i c_j} \right]$

$= n \left[\sum_{j=1}^{c} \sum_{i=1}^{r} \frac{n_{ij}^2}{r_i c_j} - 2 \sum_{j=1}^{c} \sum_{i=1}^{r} \frac{n_{ij}}{n} + \sum_{j=1}^{c} \sum_{i=1}^{r} \frac{r_i c_j}{n^2} \right]$

$= n \left[\sum_{j=1}^{c} \sum_{i=1}^{r} \frac{n_{ij}^2}{r_i c_j} - 2 + \frac{\left(\sum_j c_j\right)\left(\sum_i r_i\right)}{n^2} \right]$

$= n \left[\sum_{j=1}^{c} \sum_{i=1}^{r} \frac{n_{ij}^2}{r_i c_j} - \frac{2n}{n} + \frac{n^2}{n^2} \right] = n \left[\sum_{j=1}^{c} \sum_{i=1}^{r} \frac{n_{ij}^2}{r_i c_j} - 1 \right]$

b. When every entry in the contingency is multiplied by the same $k > 0$,

$$X^2 = kn \left[\sum_{j=1}^{c} \sum_{i=1}^{r} \frac{(kn_{ij})^2}{(kr_i)(kc_j)} - 1 \right] = kn \left[\sum_{i=1}^{c} \sum_{i=1}^{r} \frac{n_{ij}^2}{r_i c_j} - 1 \right]$$

Thus, if the pattern of responses is the same, then the X^2 will be increased k times.

14.12 The data are analyzed as a 2×6 contingency table with estimated expected cell counts shown in parentheses.

Sex	Walking	Cycling	Type of Activity Aerobics	Running	Calisthenics	Swimming	Total
Male	60 (81.9)	85 (81.9)	28 (81.9)	113 (81.9)	79 (81.9)	179 (81.9)	544
Female	106 (84.1)	81 (84.1)	138 (84.1)	55 (84.1)	89 (84.1)	90 (84.1)	559
Total	166	16	166	168	168	269	1103

The test statistic is

$$X^2 = \frac{(60-81.9)^2}{81.9} + \frac{(85-81.9)^2}{81.9} + \cdots + \frac{(90-84.1)^2}{84.1} = 135.62$$

using calculator accuracy on the expected cell counts. The rejection region with $\alpha = .05$ and 5 d.f. is $X^2 > 11.0705$, and H_0 is rejected. There is a difference in the proportion according to sex. The value $X^2 = 135.62$ is greater than $16.75 = x_{5,.005}^2$, so that the p-value $< .005$.

14.14 The table of observed and estimated expected cell counts is given below.

	Age 1	2	3
Low	8(13.16)	12(13.67)	21(14.17)
High	18(12.84)	15(13.33)	7(13.83)

Then

$$X^2 = \frac{(8-13.16)^2}{13.16} + \cdots + \frac{(7-13.83)^2}{13.83} = 11.18$$

and the rejection region is $X^2 > x_2^2 = 5.99$ with $\alpha = .05$. The null hypothesis is rejected. The two classifications are dependent.

15.4 **a.** Let $p = P$(high elevation exceeds low elevation) and M = number of nights during which high elevation exceeds low elevation. Then
$$H_0:\ p = \tfrac{1}{2} \qquad \text{vs.} \qquad H_a:\ p > \tfrac{1}{2}$$
and $n = 10$. Large values of M will tend to favor H_a, and an upper-tailed rejection region is used. The observed value of M is $m = 9$. The p-value is
$$P(M \geq 9) = 1 - P(M \leq 8) = 1 - .989 = .011.$$
The data supports H_a.

b. Extremal variables, such as the minimum temperatures in this example, often have skewed distributions, making the assumptions of the t test invalid.

15.5 **a.** The hypothesis to be tested is
$$H_0:\ p = \tfrac{1}{2} \qquad \text{vs.} \qquad H_a:\ p \neq \tfrac{1}{2}$$
where $p = P$(response for stimulus 1 exceeds that for stimulus 2). Then the test statistic is M, the number of times the response for stimulus 1 exceeds that for stimulus 2. Again, denote a positive difference by a plus sign and a negative difference by a minus sign. Then M will be equivalent to the number of plus signs observed. These signs are

Subject	1	2	3	4	5	6	7	8	9
Sign	−	−	+	−	−	−	−	+	−

and $m = 2$. The next step is to select a rejection region such that $\alpha = P\left(\text{reject } H_0 | p = \tfrac{1}{2}\right)$ is close to .05. Using the rejection region $M = 0, 1, 8, 9$, we find that
$$\alpha = P\left(M = 0, 1, 8, 9 | p = \tfrac{1}{2}\right) = \binom{9}{0}\left(\tfrac{1}{2}\right)^0\left(\tfrac{1}{2}\right)^9$$
$$+ \binom{9}{1}\left(\tfrac{1}{2}\right)^1\left(\tfrac{1}{2}\right)^8 + \binom{9}{8}\left(\tfrac{1}{2}\right)^8\left(\tfrac{1}{2}\right)^1 + \binom{9}{9}\left(\tfrac{1}{2}\right)^9\left(\tfrac{1}{2}\right)^0$$
$$= .04$$
Using the large rejection region $M = 0, 1, 2, 7, 8, 9$ we find that
$$\alpha = P\left(M = 0, 1, 2, 7, 8, 9 | p = \tfrac{1}{2}\right) = .180.$$
Therefore we see that $M = 0, 1, 8, 9$ is the correct rejection region to have .05 level test. Examining the test statistic $(m = 2)$, we fail to reject the null hypothesis.

b. The two samples are <u>not</u> random and independent. Rather, the experiment has been conducted in a paired manner, and a paired-difference analysis is used. The differences and the associated t test are given below.

(1) $H_0:\ \mu_1 - \mu_2 = 0 \qquad$ vs. $\qquad H_a:\ \mu_1 - \mu_2 \neq 0$

(2) $\bar{d} = \dfrac{\sum d_i}{n} = \dfrac{-9.2}{9} = -1.022$

$s^2 = \dfrac{\sum_i d_i^2 - \left(\sum_i d_i/n\right)^2}{n-1} = \dfrac{37.14 - 9.404}{8}$

$= \dfrac{27.736}{8} = 3.467$

(3) Test statistic: $t = \dfrac{\bar{d} - 0}{\frac{s}{\sqrt{n}}} = \dfrac{-1.022}{1.86373} = -1.65$

d_i	d_i^2
−.9	.81
−1.1	1.21
1.5	2.25
−2.6	6.76
−1.8	3.24
−2.9	8.41
−2.5	6.25
2.5	6.25
−1.4	1.96

(4) Rejection region: The critical value of t for a two-tailed test, based on 8 degrees of freedom, will be $t_{.025,8} = 2.306$, and the rejection region is $|t| > 2.306$. The test statistic does not fall in the rejection region. Hence the null hypothesis cannot be rejected.

15.10 The differences, with their ranks (according to absolute magnitude) are given in the table at the right. Then $T^- = 1$ and $T^+ = 54$. Hence $T = 1$. Using Table 9 we see that $T^{-1} = 1 < 3$, so that the p-value $< .005$ for this one-tailed test.

d_i	Rank $|d_i|$	d_i	Rank $|d_i|$
1.1	6	.7	4
1.3	8.5	1.1	6
2.8	10	.6	3
−.1	1	1.3	8.5
.5	2	1.1	6

15.11 The experimenter's design was paired using people as blocks. The Wilcoxon rank-sum test is the appropriate test for this experiment, with the null hypothesis that the populations follow the same distributions. The data are shown below.

Subject	Normal	Stress	Difference $(N - S)$	Rank
1	126	130	−4	5
2	117	118	−1	1
3	115	125	−10	7
4	118	120	−2	2
5	118	121	−3	3.5
6	128	125	+3	3.5
7	125	130	−5	6
8	120	120	0	−

The rank sum for positive values is $T^+ = 3.5$, and for negative values $= 24.5$. For $n = 7$ (one tie) and for a one-tailed test, we use the rank sum $T^* = 3.5$. Since $2 < T^* < 4$, $.025 < p$-value $< .05$.

15.16 Use the results of Exercise 15.15 with $\xi_0 = 15,000$. The differences, $d_i = Y_i - 15,000$, are given in the table at the right.

d_i	Rank $|d_i|$	d_i	Rank $|d_i|$
−200	2	3500	7
1900	5	5000	10
3000	6	4200	9
4100	8	100	1
−1800	4	1500	3

a. Using the sign test with $p = P(Y < 15,000) = \frac{1}{2}$ under the null hypothesis, a lower-tailed rejection region is used. The observed value of M is $m = 2$. The p-value is $P(M \le 2) = .055$ (use Table 1).

b. $T^- = 6, T^+ = 49$, and $T = 6$. Using a one-tailed test, the p-value is between .01 and .025. Thus, H_0 is rejected.