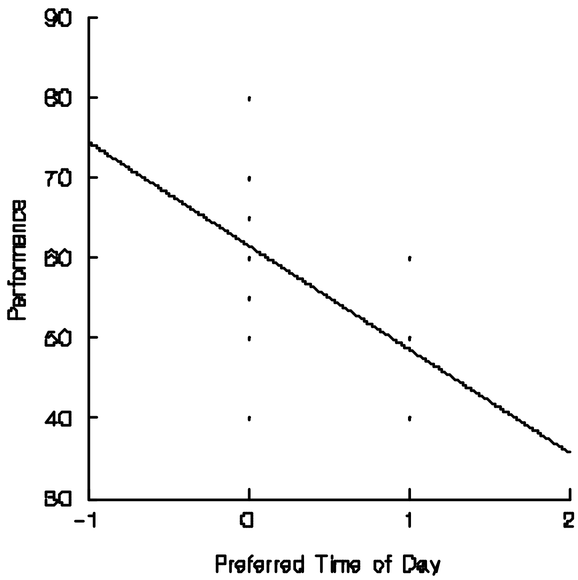
Chapter 10 - Alternative Correlational Techniques

10.1 Performance ratings in the morning related to perceived peak time to day:

a. Plot of data with regression line:



b.

$$s_X = 0.489$$

$$s_Y = 11.743$$

$$\text{cov}_{XY} = -3.105$$

$$r_{pb} = \frac{\text{cov}_{XY}}{s_X s_Y} = \frac{-3.105}{(0.489)(11.743)} = -.540$$

$$t = \frac{r\sqrt{(N-2)}}{\sqrt{1-r^2}} = \frac{(-.540)\sqrt{18}}{\sqrt{.708}} = \frac{-2.291}{.842} = -2.723 \quad [p < .01]$$

c. Performance in the morning is significantly related to people's perceptions of their peak periods.

10.3 It looks as if morning people vary their performance across time, but that evening people are uniformly poor.

10.5 Running a t test on the data in Exercise 10.1:

$\bar{X}_1 = 61.538$ $\quad s_1^2 = 114.103$ $\quad n_1 = 13$

$\bar{X}_2 = 48.571$ $\quad s_2^2 = 80.952$ $\quad n_2 = 7$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{(13-1)114.103 + (7-1)80.952}{13+7-2} = 103.053$$

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{61.538 - 48.571}{\sqrt{103.053\left(\frac{1}{13} + \frac{1}{7}\right)}} = 2.725$$

[$t_{.025(18)} = \pm 2.101$] Reject H_0

The t calculated here (2.725) is equal to the t calculated to test the significance of the r calculated in Exercise 10.1.

10.7 Regression equation for relationship between college GPA and completion of Ph.D. program:

$$b = \frac{\text{cov}_{XY}}{s_X^2} = \frac{0.051}{.503^2} = .202$$

$$a = \frac{\Sigma Y - b\Sigma X}{N} = \frac{17 - .202(72.58)}{25} = .093$$

$$\hat{Y} = bX + a = .202X + .093$$

When $X = \overline{X} = 2.9032$, $\hat{Y} = .202(2.9032) + .093 = .680 = \overline{Y}$.

10.9 Establishment of a GPA cutoff of 3.00:

a. Ph.D. (Y):	0	0	0	0	0	0	0	0	1	1	1

b.

$$s_X = 0.507$$
$$s_Y = 0.476$$
$$cov_{XY} = 0.062$$
$$\phi = \frac{0.062}{(0.507)(0.476)} = .256$$

c.

$$t = \frac{r\sqrt{(N-2)}}{\sqrt{1-r^2}} = \frac{(.256)\sqrt{23}}{\sqrt{.934}} = \frac{1.228}{.967} = 1.27 \quad \text{[not significant]}$$

10.11 Alcoholism and childhood history of ADD:

a.

$$s_X = 0.471$$
$$s_Y = 0.457$$
$$cov_{XY} = 0.135$$
$$\phi = \frac{0.135}{(0.471)(0.457)} = .628$$

b. $\chi^2 = N\phi^2 = 32\left(.628^2\right) = 12.62 \ [p < .05]$

10.13 Development ordering of language skills using Kendall's τ

a.

$$\tau = 1 - \frac{2(\text{\# inversions})}{\text{\# pairs}} = 1 - \frac{2(6)}{15(14)/2} = 1 - \frac{23}{105} = .886$$

b.

$$z = \frac{\tau}{\sqrt{\frac{2(2N+5)}{9N(N-1)}}} = \frac{.886}{\sqrt{\frac{2(30+5)}{9(15)(14)}}} = \frac{.886}{\sqrt{.037}} = 4.60 \quad [p < .05]$$

10.15 Ranking of videotapes of children's behaviors by clinical graduate students and experienced clinicians using Kendall's τ:

Experienced	New	Inversions
1	2	1
2	1	0
3	4	1
4	3	0
5	5	0

Experienced	New	Inversions
6	8	2
7	6	0
8	10	2
9	7	0
10	9	0

$$\tau = 1 - \frac{2(\#\text{ inversions})}{\#\text{ pairs}} = 1 = \frac{2(6)}{10(9)/2} = 1 - \frac{12}{45} = .733$$

10.17 Verification of Rosenthal and Rubin's statement

	Improvement	No Improvement	Total
Therapy	66 (50)	34 (50)	100
No Therapy	34 (50)	66 (50)	100
Total	100	100	200

a.

$$\chi^2 = \Sigma \frac{(O-E)^2}{E} = \frac{(66-50)^2}{50} + \frac{(34-50)^2}{50} + \frac{(34-50)^2}{50} + \frac{(66-50)^2}{50}$$
$$= 20.48$$

b. An $r^2 = .0512$ would correspond to $\chi^2 = 10.24$. The closest you can come to this result is if the subjects were split 61/39 in the first condition and 39/61 in the second (rounding to integers.)

10.19 ClinCase against Group in Mireault's data

	ClinCase 0	ClinCase 1
Loss	69	66
Married	108	73
Divorced	36	23

a. $\chi^2 = 2.815$ [$p = .245$]
$\phi_C = .087$

c. This approach would be preferred over the approach used in Chapter 7 if you had reason to believe that differences in depression scores below the clinical cutoff were of no importance and should be ignored.

10.21 Small Effects:

a. If a statistic is not significant, that means that we have no reason to believe that it is reliably different from 0 (or whatever the parameter under H_0). In the case of a correlation, if it is not significant, that means that we have no reason to believe that there is a relationship between the two variables. Therefore it cannot be important.

b. With the exceptions of issues of power, sample size will not make an effect more important than it is. Increasing N will increase our level of significance, but the magnitude of the effect will be unaffected.