

17.4 Exercises

Basic Concepts

1. What is the correlation coefficient? How is this different from the Spearman rank correlation coefficient?
2. What is the formula for calculating Spearman's rho?
3. Can you calculate Spearman's rho if there are ties in the rank data?
4. Identify the difference in notation between Spearman's rho for population and sample data.
5. Explain the similarities in the behavior of the parametric correlation coefficient and Spearman's rho.
6. Identify one main advantage of the Spearman's rank correlation coefficient versus the parametric correlation coefficient.
7. Explain the procedure for ranking data when calculating Spearman's rho.
8. What are the null and alternative hypotheses for the rank correlation test?
9. Consider the value $r_s = 0.12$. Interpret this value in terms of the x and y variables used to calculate Spearman's rho.

Exercises

10. Chris is a new cashier assigned to a scanner in a supermarket. Each day a sample of purchases at that scanner is examined and a percent of pricing errors is recorded along with the total number of customers who used that scanner. Does the following data indicate an association between Chris' performance and how busy his scanner was? Use $\alpha = 0.05$.

% Pricing Errors and Total Customers			
Number of Customers	Errors (%)	Number of Customers	Errors (%)
57	4.2	67	2.5
44	5.5	71	2.9
32	5.7	69	2.6
60	3.9	56	1.0
55	3.2	51	2.0
59	4.1	70	1.7
63	3.3		

11. Twelve new runners were randomly assigned to different training programs, where they were required to run a certain number of miles every week for a year prior to a major race. After the training, the participants ran the race and their finishing times were recorded.

Miles of Training and Race Times			
Miles Logged	Race Time (Minutes)	Miles Logged	Race Time (Minutes)
35	198	30	189

Miles of Training and Race Times			
Miles Logged	Race Time (Minutes)	Miles Logged	Race Time (Minutes)
25	165	29	240
45	155	42	224
60	148	24	201
70	135	19	246
21	243	55	166

- With 95% confidence, is there evidence that the number of miles logged in a week during training affects the runner's race time?
 - Can the linear correlation coefficient, r , be calculated in order to fit a least squares regression line to the data in the table in an effort to predict the finish time of runners based on the number of miles logged during training? Why or why not?
12. The following data consists of college rankings of five universities by two different magazines. Is there a correlation between the rankings of the magazines? Use $\alpha = 0.10$.

College Rankings by Magazines					
College	A	B	C	D	E
Magazine 1	1	4	2	3	5
Magazine 2	4	3	1	5	2

13. An anthropologist records the heights (in inches) of ten fathers and their sons. Does the following data support (at the 5% level) that taller fathers tend to have taller sons?

Heights of Fathers and Sons (Inches)			
Son's Height	Father's Height	Son's Height	Father's Height
72	70	65	71
68	73	70	78
74	72	69	67
66	68	67	65
71	69	80	66

14. After a mother-daughter golf tournament, mothers and daughters were ranked among themselves. Does the following data show (at the 5% level) a correlation between the daughters' and mothers' golf skills?

Golf Rankings			
Daughter's Ranking	Mother's Ranking	Daughter's Ranking	Mother's Ranking
1	5	5	3
9	4	3	6
10	8	7	7
2	2	6	10
4	1	8	9