

Practice problems – Spearman's r and regression

1. For the following set of data, find the linear regression equation for predicting Y from X:

<u>X</u>	<u>Y</u>
0	9
2	9
4	7
6	3

answer: $SS_x = 20$; $SP = -20$; $y' = 10 - x$

2. For the set of $n = 6$ pairs of X and Y values, there is a correlation of $r = .60$ and $SS_y = 100$. If you are using this data as the basis for a regression equation,

- a) On average, how much error would you expect if the regression equation were used to predict the Y score for a specific individual? That is, find the standard error of estimate.

answer: $S_{y.x} = 4$

- b) How much error would you expect if the sample size was $n = 102$.

answer: if $n = 102$, $S_{y.x} = .80$

- c) How much error would you expect if the sample correlation were $r = .80$ instead of $.60$?

answer: if $r = .80$, $S_{y.x} = 3$ (where $n = 6$).

3. A photographer has noticed that a freshly mixed batch of chemicals will develop photographs faster than an old batch of chemicals. The photographer keeps records of the time needed to develop a print and the chemicals:

<u>age of chemicals</u>	<u>time to develop</u>
1	35
4	38
6	40
9	44
12	49
14	52

Use this data to find the regression equation for predicting time as a function of age of the chemicals.

answer: $y' = 32.8 + 1.33x$

4. Using the regression equation from problem #3:

- a) Predict how long it should take to develop a print if the chemical are 10 days old.

answer: if $x = 10$, then $y' = 46.10$

- b) Calculate the standard error of estimate to determine the degree of accuracy for your prediction.

answer: $S_{y.x} = .73$

5. An investigator wants to arrange the 15 items on her scale of language impairment on the basis of the order in which language skills appear in development. Not being entirely confident that she has selected the correct ordering of skills, she asks another professional to rank the items from 1 to 15 in terms of the order in which he thinks they should appear. The data are given below:

Investigator:

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

Consultant:

1 3 2 4 7 5 6 8 10 9 11 12 15 13 14

Use Pearson's formula (r) to calculate Spearman's r_s .

answer: $SP = 272$; $SS_x = 280$; $SS_y = 280$; $r = .97$

6. Determine whether the following Beatles album popularity in terms of sales were related to when they were released. That is, we have ranked the albums in terms of what was released first (rank = 1) and what was released last (rank = 8). Determine whether release period is related to popularity. Based on the correlation coefficient, did the Beatles get more or less popular with time?

Album	year	release rank	sales rank
Beatles for Sale	65	1	7
Rubber Soul	65	2	6
Revolver	66	3	8
Sgt. Pepper	67	4	1
Magical Mystery Tour	67	5	4
The Beatles (white album)	68	6	2
Abbey Road	69	7	3
Let it Be	70	8	5

answer: $SP = -22$; $SS_x = 42$; $SS_y = 42$ $r = -.52$

7. The data here represents approximate time in hours spent training for a marathon (x) and approximate time to complete the marathon in hours (y).

(a) Compute a Pearson's Product Moment correlation on the following data; (b) determine if the relationship is significant; (c) sketch a scatterplot of the data; (d) what is the standard error of estimate?

x	y
8	2
4	2
2	4
1	5
5	2

answer: (a) $SP = -13$; $SS_x = 30$; $SS_y = 8$; $r = -.84$ (b) $r_{crit} = .878$; $r(3) = -.84$, $p > .05$
 (d) $S_{y.x} = .88$

8. A previous statistics class kept records of the number of hours they studied and their score on the final 20 point quiz. Here is the data collected.

hours studied	quiz score
\underline{X}	\underline{Y}
8	18
6	14
5	13
7	15
4	10

- Graph the observed values. Compute the regression equation and report and graph the predicted values. **answer:** $y' = 3.2 + 1.8x$
- Calculate the proportionate reduction in error. **answer:** $PRE = r^2 = .95$
- What is the correlation between X and Y? **answer:** $r = .98$ (rounded from .976)
- What is the coefficient of alienation? **answer:** $1 - r^2 = 1 - .95 = .05$

9. In a study of diagnostic processes, entering clinical graduate students are shown a 20-minute videotape of children's behavior and asked to rank-order 10 behavioral events on the tape in the order of the importance each has for a behavioral assessment. (1 = most important.) The data are then averaged to produce an average rank ordering for the entire class. The same thing was then done using experienced clinicians. The data follow:

<u>Events</u>	1	2	3	4	5	6	7	8	9	10
Experienced Clinicians:	1	3	2	7	5	4	8	6	9	10
New Students	2	4	1	6	5	3	10	8	7	9

Use Spearman's Γ_s to measure the agreement between experienced and novice clinicians.
answer $r = .89$