Radiology

Kelly H. Zou, PhD
Kemal Tuncali, MD
Stuart G. Silverman, MD

Index terms:
Data analysis
Statistical analysis

Published online
10.1148/radiol.2273011499
Radiology 2003; 227:617–628

[1] From the Department of Radiology, Brigham and Women's Hospital (K.H.Z., K.T., S.G.S.) and Department of Health Care Policy (K.H.Z.), Harvard Medical School, 180 Longwood Ave, Boston, MA 02115. Received September 10, 2001; revision requested October 31; revision received December 26; accepted January 21, 2002. **Address correspondence to** K.H.Z. (e-mail: *zou@bwh.harvard.edu*).

Correlation and Simple Linear Regression[1]

In this tutorial article, the concepts of correlation and regression are reviewed and demonstrated. The authors review and compare two correlation coefficients, the Pearson correlation coefficient and the Spearman ρ, for measuring linear and nonlinear relationships between two continuous variables. In the case of measuring the linear relationship between a predictor and an outcome variable, simple linear regression analysis is conducted. These statistical concepts are illustrated by using a data set from published literature to assess a computed tomography–guided interventional technique. These statistical methods are important for exploring the relationships between variables and can be applied to many radiologic studies.

Results of clinical studies frequently yield data that are dependent of each other (eg, total procedure time versus the dose in computed tomographic [CT] fluoroscopy, signal-to-noise ratio versus number of signals acquired during magnetic resonance imaging, and cigarette smoking versus lung cancer). The statistical concepts correlation and regression, which are used to evaluate the relationship between two continuous variables, are reviewed and demonstrated in this article.

Analyses between two variables may focus on *(a)* any association between the variables, *(b)* the value of one variable in predicting the other, and *(c)* the amount of agreement. Agreement will be discussed in a future article. Regression analysis focuses on the form of the relationship between variables, while the objective of correlation analysis is to gain insight into the strength of the relationship (1,2). Note that these two techniques are used to investigate relationships between continuous variables, whereas the χ^2 test is an example of a test for association between categorical variables. Continuous variables, such as procedure time, patient age, and number of lesions, have no gaps on the measurement scale. In contrast, categorical variables, such as patient sex and tissue classification based on segmentation, have gaps in their possible values. These two types of variables and the assumptions about their measurement scales were reviewed and distinguished in an article by Applegate and Crewson (3) published earlier in this Statistical Concepts Series in *Radiology*.

Specifically, the topics covered herein include two commonly used correlation coefficients, the Pearson correlation coefficient (4,5) and the Spearman ρ (6–10) for measuring linear and nonlinear relationship, respectively, between two continuous variables. Correlation analysis is often conducted in a retrospective or observational study. In a clinical trial, on the other hand, the investigator may also wish to manipulate the values of one variable and assess the changes in values of another variable. To evaluate the relative impact of the predictor variable on the particular outcome, simple regression analysis is preferred. We illustrate these statistical concepts with existing data to assess patient skin dose based on total procedure time by using a quick-check method in CT fluoroscopy–guided abdominal interventions (11).

These statistical methods are useful tools for assessing the relationships between continuous variables collected from a clinical study. However, it is also important to distinguish these statistical methods. While they are similar mathematically, their purposes are different. Correlation analysis is generally overused. It is often interpreted incorrectly (to establish "causation") and should be reserved for generating hypotheses rather than for testing them. On the other hand, regression modeling is a more useful statistical technique that allows us to assess the strength of the relationships in the data and the uncertainty in the model by using confidence intervals (12,13).

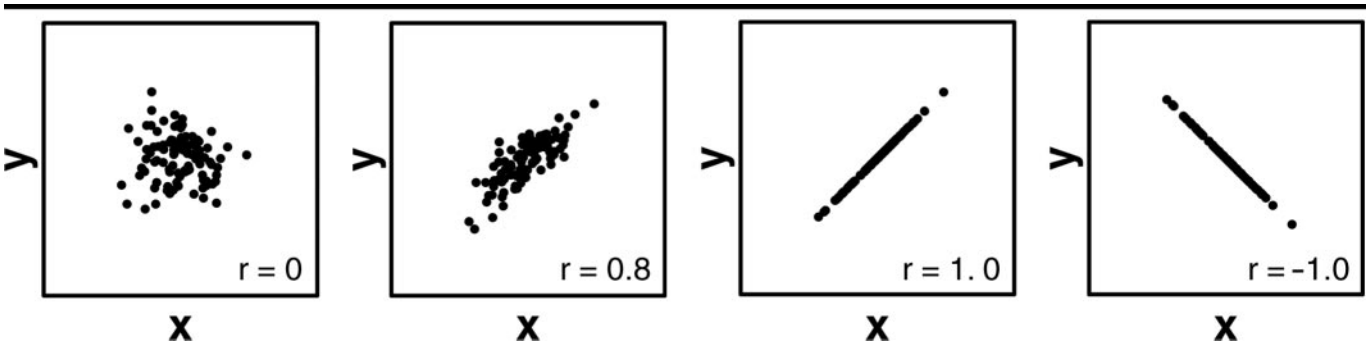

Figure 1. Scatterplots of four sets of data generated by means of the following Pearson correlation coefficients (from left to right): $r = 0$ (uncorrelated data), $r = 0.8$ (strongly positively correlated), $r = 1.0$ (perfectly positively correlated), and $r = -1$ (perfectly negatively correlated).

CORRELATION

The purpose of correlation analysis is to measure and interpret the strength of a linear or nonlinear (eg, exponential, polynomial, and logistic) relationship between two continuous variables. When conducting correlation analysis, we use the term *association* to mean "linear association" (1,2). Herein, we focus on the Pearson and Spearman ρ correlation coefficients. Both correlation coefficients take on values between −1 and +1, ranging from being negatively correlated (−1) to uncorrelated (0) to positively correlated (+1). The sign of the correlation coefficient (ie, positive or negative) defines the direction of the relationship. The absolute value indicates the strength of the correlation (Table 1, Fig 1). We elaborate on two correlation coefficients, linear (eg, Pearson) and rank (eg, Spearman), that are commonly used for measuring linear and general relationships between two variables.

Linear Correlation

The Pearson correlation coefficient is also known as the sample correlation coefficient *(r),* product-moment correlation coefficient, or coefficient of correlation (14). It was introduced by Galton in 1877 (15,16) and developed later by Pearson (17). It measures the linear relationship between two random variables. For example, when the value of the predictor is manipulated (increased or decreased) by a fixed amount, the outcome variable changes proportionally (linearly). A linear correlation coefficient can be computed by means of the data and their sample means (Appendix A). When a scientific study is planned, the required sample size may be computed on the basis of a certain hypothesized value with the desired statistical power at a specified level of significance (Appendix B) (18).

Rank Correlation

The Spearman ρ is the sample correlation coefficient *(r_s)* of the ranks (the relative order) based on continuous data (19,20). It was first introduced by Spearman in 1904 (6). The Spearman ρ is used to measure the monotonic relationship between two variables (ie, whether one variable tends to take either a larger or smaller value, though not necessarily linearly) by increasing the value of the other variable.

Linear versus Rank Correlation Coefficients

The Pearson correlation coefficient necessitates use of interval or continuous measurement scales of the measured outcome in the study population. In contrast, rank correlations also work well with ordinal rating data, and continuous data are reduced to their ranks (Appendix C) (20,21). The rank procedure will also be illustrated briefly with our example data. The smallest value in the sample has rank 1, and the largest has the highest rank. In general, rank correlations are not easily influenced by the presence of skewed data or data that are highly variable.

Statistical Hypothesis Tests for a Correlation Coefficient

The null hypothesis states that the underlying linear correlation has a hypothesized value, ρ_0. The one-sided alternative hypothesis is that the underlying value exceeds (or is less than) ρ_0. When the sample size *(n)* of the paired data is large ($n \geq 30$ for each variable), the standard error *(s)* of the linear correlation *(r)* is approximately $s(r) = (1 - r^2)/\sqrt{n}$. The test statistic value $(r - \rho_0)/s(r)$ may be computed by means of the z test (22). If the P value is below .05, the null hypothesis is rejected. The P value based on the Spearman ρ can be found in the literature (20,21).

TABLE 1
Interpretation of Correlation Coefficient

Correlation Coefficient Value	Direction and Strength of Correlation
−1.0	Perfectly negative
−0.8	Strongly negative
−0.5	Moderately negative
−0.2	Weakly negative
0.0	No association
+0.2	Weakly positive
+0.5	Moderately positive
+0.8	Strongly positive
+1.0	Perfectly positive

Note.—The sign of the correlation coefficient (ie, positive or negative) defines the direction of the relationship. The absolute value indicates the strength of the correlation.

Limitations and Precautions

It is worth noting that even if two variables (eg, cigarette smoking and lung cancer) are highly correlated, it is not sufficient proof of causation. One variable may cause the other or vice versa, or a third factor is involved, or a rare event may have occurred. To conclude causation, the causal variables must precede the variable it causes, and several conditions must be met (eg, reversibility, strength, and exposure response on the basis of the Bradford-Hill criteria or the Rubin causal model) (23–26).

SIMPLE LINEAR REGRESSION

The purpose of simple regression analysis is to evaluate the relative impact of a predictor variable on a particular outcome. This is different from a correlation analysis, where the purpose is to examine the strength and direction of the rela-

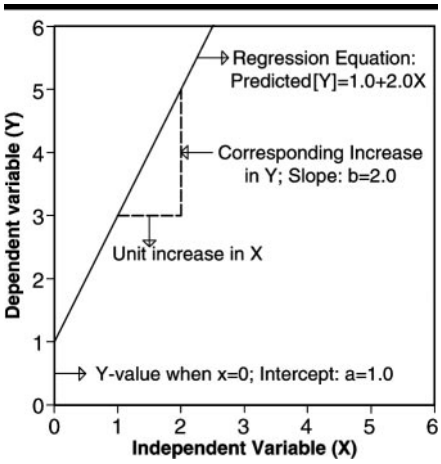

Figure 2. Simple linear regression model shows that the expectation of the dependent variable *Y* is linear in the independent variable *X*, with an intercept $a = 1.0$ and a slope $b = 2.0$.

tionship between two random variables. In this article, we deal with only linear regression of one continuous variable on another continuous variable with no gaps on each measurement scale (3). There are other types of regression (eg, multiple linear, logistic, and ordinal) analyses, which will be provided in a future article in this Statistical Concepts Series in *Radiology*.

A simple regression model contains only one independent (explanatory) variable, X_i, for $i = 1, \ldots, n$ subjects, and is linear with respect to both the regression parameters and the dependent variable. The corresponding dependent (outcome) variable is labeled. The model is expressed as

$$Y_i = a + bX_i + e_i, \qquad (1)$$

where the regression parameter a is the intercept (on the y axis), and the regression parameter b is the slope of the regression line (Fig 2). The random error term e_i is assumed to be uncorrelated, with a mean of 0 and constant variance. For convenience in inference and improved efficiency in estimation (27), analyses often incur an additional assumption that the errors are distributed normally. Transformation of the data to achieve normality may be applied (28,29). Thus, the word *line* (linear, independent, normal, equal variance) summarizes these requirements.

Typical steps for regression model analysis are the following: *(a)* determine if the assumptions underlying a normal relationship are met in the data, *(b)* obtain the equation that best fits the data, *(c)* evaluate the equation to determine the strength of the relationship for prediction and estimation, and *(d)* assess whether the data fit these criteria before the equation is applied for prediction and estimation.

Least Squares Method

The main goal of linear regression is to fit a straight line through the data that predicts *Y* based on *X*. To estimate the intercept and slope regression parameters that determine this line, the least squares method is commonly used. It is not necessary for the errors to have a normal distribution, although the regression analysis is more efficient with this assumption (27). With this regression method, a set of regression parameters are found such that the sum of squared residuals (ie, the differences between the observed values of the outcome variable and the fitted values) are minimized (14). The fitted y value is then computed as a function of the given x value and the estimated intercept and slope regression parameter (Appendix D). For example, in Equation (1), once the estimates of a and b are obtained from the regression analysis, the predicted y value at any given x value is calculated as $a + bx$.

Coefficient of Determination, R^2

It is meaningful to interpret the value of the Pearson correlation coefficient r by squaring it; hence, the term R-square (R^2) or coefficient of determination. This measure (with a range of 0–1) is the fraction of the variability in *Y* that can be explained by the variability in *X* through their linear relationship, or vice versa. That is, $R^2 = SS_{regression}/SS_{total}$, where SS stands for the sum of squares. Note that R^2 is calculated only on the basis of the Pearson correlation coefficient in the linear regression analysis. Thus, it is not appropriate to compute R^2 on the basis of rank correlation coefficients such as the Spearman ρ.

Statistical Hypothesis Tests

There are several hypotheses in the context of regression analysis, for example, to test if the slope of the regression line is $b = 0$ (hypothesis, there is no linear association between *Y* and *X*). One may also test whether intercept a takes on a certain value. The significance of the effects of the intercept and slope may also be computed by means of a Student t statistic introduced earlier in this Statistical Concepts Series in *Radiology* (30).

Limitations and Precautions

The following understandings should be considered when regression analysis is performed. *(a)* To understand whether the assumptions have been met, determine the magnitude of the gap between the data and the assumptions of the model. *(b)* No matter how strong a relationship is demonstrated with regression analysis, it should not be interpreted as causation (as in the correlation analysis). *(c)* The regression should not be used to predict or estimate outside the range of values of the independent variable of the sample (eg, extrapolation of radiation cancer risk from the Hiroshima data to that of diagnostic radiologic tests).

AN EXAMPLE: DOSE VERSUS TOTAL PROCEDURE TIME IN CT FLUOROSCOPY

We applied these statistical methods to help assess the benefit of the use of CT fluoroscopy to guide interventions in the abdomen (11). During CT fluoroscopy–guided interventions, one might postulate that the radiation dose received by a patient is related to (or correlated with) the total procedure time, because the more difficult the procedure is, the more CT fluoroscopic scanning is required, which means a longer procedure time. The rationale was to assess whether radiation dose could be estimated by simply measuring the total CT fluoroscopic procedure time, with the null hypothesis that the slope of the regression line is 0.

Earlier, we discussed two methods to target lesions with CT fluoroscopy. In one method, continuous CT scanning is used during needle placement. In the other method, short CT scanning is used to image the needle after it is placed. The latter method, the so-called quick-check method, has been adopted almost exclusively at our institution. Now, we demonstrate correlation and regression analyses based on a subset of the interventional procedures ($n = 19$). With the quick-check method, we examine the relationship between total procedure time (in minutes) and dose (in rads) on a natural log scale. We also examine the marginal ranks of the x (log of total time) and y (log of dose) components (Table 2). For convenience, the x data are given in ascending order.

In Table 2, each set of rank data is derived by first placing the 19 observations in each sample in ascending order and then assigning ranks 1–19. Ties are broken by means of averaging the respective adjacent ranks. Finally, the ranks are identified for the observations of each of the paired x and y samples.

The natural log (ln) transformation of the total time is used to make the data appear normal, for more efficient analysis (Appendix D), with normality verified statistically (31). However, normality is not necessary in the subsequent regression analysis. We created a scatterplot of the data, with the log of dose (ln[rad]) on the x axis and the log of total time (ln[minutes]) on the y axis (Fig 3).

For illustration purposes, we will conduct both correlation and regression analyses; however, the choice of analysis depends on the aim of research. For example, if the investigators wish to assess whether there is a relationship between time and dose, then correlation analysis is appropriate. In comparison, if the investigators wish to evaluate the impact of the total time on the resulting dose, then regression analysis is preferred.

Correlations

To compute the Spearman ρ with a Pearson correlation coefficient of $r = 0.85$, the marginal ranks of time and dose were derived separately; consequently, $r_s = 0.84$. Both correlation coefficients confirm that the log of total time and the log of dose are correlated strongly and positively.

Regression

We first conducted a simple linear regression analysis of the data on a log scale ($n = 19$); results are shown in Table 3. The value calculated for R^2 was 0.73, which suggests that 73% of the variability of the data could be explained by the linear regression.

The regression line, expressed in the form given in Equation (1), is $Y = -9.28 + 2.83X$, where the predictor variable X represents the log of total time, and the outcome variable Y represents the log of dose. The estimated regression parameters are $a = -9.28$ (intercept) and $b = 2.83$ (slope) (Fig 4). This regression line can be interpreted as follows: At $X = 0$, the value of Y is −9.28. For every one-unit increase in X, the value of Y will increase on average by 2.83. Effects of both the intercept and slope are statistically significant ($P < .005$) (Excel; Microsoft, Redmond, Wash); therefore, the null hypothesis (H_0, the dose remains constant as the total procedure time increases) is rejected. Thus, we confirm the alternative hypothesis (H_1, the dose increases in the total procedure time).

The regression line may be used to give predicted values of Y. For example, if in a future CT fluoroscopy procedure, the log total time is specified at $x = 4$ (translated to $e^4 = 55$ minutes, approximately), then the log dose that is to be applied is approximately $y = -9.28 + 2.83 \times 4 = 2.04$ (translated to $e^{2.04} = 7.69$ rad). On the other hand, if the log total time is specified at $x = 4.5$ (translated to $e^{4.5} = 90$ minutes, approximately), then the log dose that is to be applied is approximately $y = -9.28 + 2.83 \times 4.5 = 3.46$ (translated to $e^{3.46} = 31.82$ rad). Such prediction can be useful for future clinical practice.

TABLE 2
Total Procedure Time and Dose of CT Fluoroscopy–guided Procedures, by Means of the Quick-Check Method

Subject No.	*x* Data: Log Time (ln[min])	Ranks of *x* Data	*y* Data: Log Dose (ln[rad])	Ranks of *y* Data
1	3.61	1	1.48	2
2	3.87	2	1.24	1
3	3.95	3	2.08	5.5
4	4.04	4	1.70	3
5	4.06	5	2.08	5.5
6	4.11	6	2.94	10
7	4.19	7	2.24	7
8	4.20	8	1.85	4
9	4.32	9.5	2.84	9
10	4.32	9.5	3.93	16
11	4.42	11.5	3.03	11
12	4.42	11.5	3.23	13
13	4.45	13	3.87	15
14	4.50	14	3.55	14
15	4.52	15	2.81	8
16	4.57	16	4.07	17
17	4.58	17	4.44	19
18	4.61	18	3.16	12
19	4.74	19	4.19	18

Source.—Reference 11.

Note.—Paired *x* and *y* data are sorted according to the *x* component; therefore, the log of the total procedure time and the log of the corresponding rank have an increasing order. When ties are present in the data, the average of their adjacent ranks is used. Pearson correlation coefficient between log time and log dose, $r = 0.85$; Spearman $\rho = 0.84$.

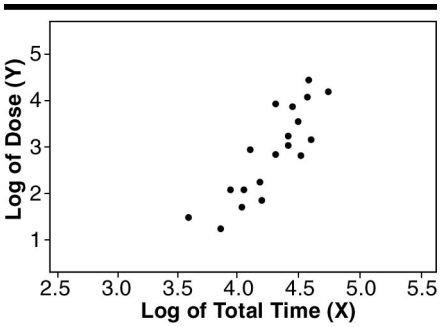

Figure 3. Scatterplot of the log of dose (y axis) versus the log of total time (x axis). Each point in the scatterplot represents the values of two variables for a given observation.

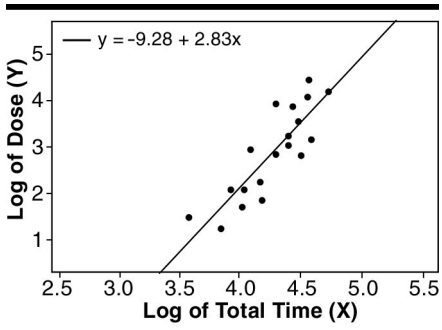

Figure 4. Scatterplot of the log of dose (y axis) versus the log of total time (x axis). The regression line has the intercept $a = -9.28$ and slope $b = 2.83$. We conclude that there is a possible association between the radiation dose and the total time of the procedure.

TABLE 3
Results based on Correlation and Regression Analysis for Example Data

Regression Statistic	Numerical Result
Correlation coefficient *r*	0.85
R-square (R^2)	0.73
Regression parameter	
Intercept	−9.28
Slope	2.83

Source.—Reference 11.

SUMMARY AND REMARKS

Two important statistical concepts, correlation and regression, which are used commonly in radiology research, are reviewed and demonstrated herein. Addi-

tional sources of information and electronic textbooks on statistical analysis methods found on the World Wide Web are listed in Appendix E. A glossary of the statistical terms used in this article is presented in Appendix F.

When correlation analysis is conducted to measure the association between two random variables, either the Pearson linear correlation coefficient or the Spearman rank correlation coefficient ρ may be adopted. The former coefficient is used to measure the linear relationship but is not recommended for use with skewed data or data with extremely large or small values (often called the outliers). In contrast, the latter coefficient is used to measures a general association, and it is recommended for use with data that are skewed or that have outliers.

When simple regression analysis is conducted to assess the linear relationship of a dependent variable as a function of the independent variable, caution must be used when determining which of the two variables is viewed as the independent variable that makes sense clinically. A useful graphical aid is a scatterplot. Once the regression line is obtained, caution should also be used to avoid prediction of a y value for any value of x that is outside the range of the data. Finally, correlation and regression analyses do not infer causality, and more rigorous analyses are required if causal inference is to be made (23–26).

APPENDIX A

Formula for computing the Pearson correlation coefficient, r: The formula for computing r between bivariate data, X_i and Y_i values ($i = 1,\ldots,n$) is

$$r = \frac{\sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n} (X_i - \overline{X})^2 \sum_{i=1}^{n} (Y_i - \overline{Y})^2}},$$

where $\overline{X}$ and $\overline{Y}$ are the sample means of the X_i and Y_i values, respectively.

The Pearson correlation coefficient may be computed by means of a computer-based statistics program (Excel; Microsoft) by using the option "Correlation" under the option "Data Analysis Tools". Alternatively, it may also be computed by means of a built-in software function "Cor" (Insightful; MathSoft, Seattle, Wash [MathSoft S-Plus 4 guide to statistics, 1997; 89–96]. Available at: *www.insightful.com*) or with a free software program (R Software. Available at: *lib.stat.cmu.edu/R)*.

APPENDIX B

Total sample size based on the Pearson correlation coefficient: Specify r = expected correlation coefficient, $C = 0.5 \times \ln[(1 + r)/(1 - r)]$, N = total number of subjects required, α = type I error (ie, significance level, typically fixed at 0.05), β = type II error (ie, 1 minus statistical power, typically fixed at 0.10). Then $N = [(Z_\alpha + Z_\beta)/C]^2 + 3$, where Z_α is the inverse of the cumulative probability of a standard normal distribution with the tail probability of α. Similarly, Z_β is the inverse of the cumulative probability of a standard normal distribution with the tail probability of β. Consequently, compute the smallest integer, n, such that $n \geq N$, as the required sample size.

For example, an investigator wishes to conduct a clinical trial of a paired design based on a one-tailed hypothesis test of the correlation coefficient. The null hypothesis is that the correlation between two variables is $r = 0.60$ (ie, $C = 0.693$) in the population of interest. The alternative hypothesis is that the correlation is $r > 0.60$. Type I error is fixed to be 0.05 (ie, $Z_\alpha = 1.645$), while type II error is fixed to be 0.10 (ie, $Z_\beta = 1.282$). Thus, the required sample size is $N = 21$ subjects. A sample size table may also be found in reference 18.

APPENDIX C

Formula for computing Spearman ρ and Pearson r_s: Replace bivariate data, X_i and Y_i ($i = 1,\ldots,n$), by their respective ranks R_i = rank(X_i) and S_i = rank(Y_i). Rank correlation coefficient, r_s, is defined as the Pearson correlation coefficient between the R_i and S_i values, which can be computed by means of the formula given in Appendix A. An alternative direct formula was given by Hettmansperger (19).

The Spearman ρ may also be computed by first reducing the continuous data to their marginal ranks by using the "rank and percentile" option with Data Analysis Tools (Excel; Microsoft) or the "rank" function (Insightful; MathSoft) or the free software. Both software programs correctly rank the data in ascending order. However, the rank and percentile option in Excel ranks the data in descending order (the largest is 1). Therefore, to compute the correct ranks, one may first multiply all of the data by −1 and then apply the rank function. Excel also gives integer ranks in the presence of ties compared with the methods that yield possible noninteger ranks, as described in the standard statistics literature (19).

Subsequently, the sample correlation coefficient is computed on the basis of the ranks of the two marginal data by using the Correlation option in Data Analysis Tools (Excel; Microsoft) or by using the Cor function (Insightful; MathSoft) or the free software.

APPENDIX D

Simple regression analysis: Regression analysis may be performed by using the "Regression" option with Data Analysis Tools (Excel; Microsoft). This regression analysis tool yields the sample correlation R^2; estimates of the regression parameters, along with their statistical significance on the basis of the Student t test; residuals; and standardized residuals. Scatter, line fit, and residual plots may also be created. Alternatively, the analyses can be performed by using the function "lsfit" (Insightful; MathSoft) or the free software.

With either program, one may choose to transform the data or exclude outliers before conducting a simple regression analysis. A commonly used variance-stabilizing transformation is the natural log function (ln) applied to one or both variables. Other transformation (eg, Box-Cox transformation) and weighting methods in regression analysis may also be used (28,29).

APPENDIX E

Uniform resource locator, or URL, links to electronic statistics textbooks: *www.davidmlane.com/hyperstat/index.html*, *www.statsoft.com/textbook/stathome.html*, *www.ruf.rice.edu/~lane/rvls.html*, *www.bmj.com/collections/statsbk/index.shtml*, *espse.ed.psu.edu/statistics/investigating.htm*.

APPENDIX F

Glossary of statistical terms:

Bivariate data.—Measurements obtained on more than one variable for the same unit or subject.

Correlation coefficient.—A statistic between −1 and 1 that measures the association between two variables.

Intercept.—The constant a in the regression equation, which is the value for y when $x = 0$.

Least squares method.—The regression line that is the best fit to the data for which the sum of the squared residuals is minimized.

Outlier.—An extreme observation far away from the bulk of the data, often caused by faulty measuring equipment or recording error.

Pearson correlation coefficient.—Sample correlation coefficient for measuring the linear relationship between two variables.

R^2.—The square of the Pearson correlation coefficient r, which is the fraction of the variability in Y that can be explained by

the variability in X through their linear relationship or vice versa.

Rank.—The relative ordering of the measurements in a variable, which can be noninteger numbers in the presence of ties.

Residual.—The difference between the observed values of the outcome variable and the fitted values based on a linear regression analysis.

Scatterplot.—A plot of the observed bivariate outcome variable (y axis) against its predictor variable (x axis), with a dot for each pair of bivariate observations.

Simple linear regression analysis.—A linear regression analysis with one predictor and one outcome variable.

Skewed data.—A distribution is skewed if there are more extreme data on one side of the mean. Otherwise, the distribution is symmetric.

Slope.—The constant b in the regression equation, which is the change in y that corresponds to a one-unit increase (or decrease) in x.

Spearman ρ.—A rank correlation coefficient for measuring the monotone relationship between two variables.

Acknowledgments: We thank Kimberly E. Applegate, MD, MS, and Philip E. Crewson, PhD, co-editors of this Statistical Concepts Series in *Radiology* for their constructive comments on earlier versions of this article.

References

1. Krzanowsk WJ. Principles of multivariate analysis: a user's perspective. Oxford, England: Clarendon, 1988; 405–432.
2. Rodriguez RN. Correlation. In: Kotz S, Johnson NL, eds. Encyclopedia of statistical sciences. New York, NY: Wiley, 1982; 193–204.
3. Applegate KE, Crewson PE. An introduction to biostatistics. Radiology 2002; 225: 318–322.
4. Goldman RN, Weinberg JS. Statistics: an introduction. Upper Saddle River, NJ: Prentice Hall, 1985; 72–98.
5. Freund JE. Mathematical statistics. 5th ed. Upper Saddle River, NJ: Prentice Hall, 1992; 494–546.
6. Spearman C. The proof and measurement of association between two things. Am J Psychol 1904; 15:72–101.
7. Fieller EC, Hartley HO, Pearson ES. Tests for rank correlation coefficient. I. Biometrika 1957; 44:470–481.
8. Fieller EC, Pearson ES. Tests for rank correlation coefficients. II. Biometrika 1961; 48:29–40.
9. Kruskal WH. Ordinal measurement of association. J Am Stat Assoc 1958; 53:814–861.
10. David FN, Mallows CL. The variance of Spearman's rho in normal samples. Biometrika 1961; 48:19–28.
11. Silverman SG, Tuncali K, Adams DF, Nawfel RD, Zou KH, Judy PF. CT fluoroscopy-guided abdominal interventions: techniques, results, and radiation exposure. Radiology 1999; 212:673–681.
12. Daniel WW. Biostatistics: a foundation for analysis in the health sciences. 7th ed. New York, NY: Wiley, 1999.
13. Altman DG. Practical statistics for medical research. Boca Raton, Fla: CRC, 1990.
14. Neter J, Wasserman W, Kutner MH. Applied linear models: regression, analysis of variance, and experimental designs. 3rd ed. Homewood, Ill: Irwin, 1990; 38–44, 62–104.
15. Galton F. Typical laws of heredity. Proc R Inst Great Britain 1877; 8:282–301.
16. Galton F. Correlations and their measurements, chiefly from anthropometric data. Proc R Soc London 1888; 45:219–247.
17. Pearson K. Mathematical contributions to the theory of evolution. III. Regression, heredity and panmixia. Phil Trans R Soc Lond Series A 1896; 187:253–318.
18. Hulley SB, Cummings SR. Designing clinical research: an epidemiological approach. Baltimore, Md: Williams & Wilkins, 1988; appendix 13.C.
19. Hettmansperger TP. Statistical inference based on ranks. Malabar, Fla: Krieger, 1991; 200–205.
20. Kendall M, Gibbons JD. Rank correlation methods. 5th ed. New York, NY: Oxford University Press, 1990; 8–10.
21. Zou KH, Hall WJ. On estimating a transformation correlation coefficient. J Appl Stat 2002; 29:745–760.
22. Fisher RA. Frequency distributions of the values of the correlation coefficient in samples from an indefinitely large population. Biometrika 1915; 10:507–521.
23. Duncan OD. Path analysis: sociological examples. In: Blalock HM Jr, ed. Causal models in the social sciences. Chicago, Ill: Alpine-Atherton, 1971; 115–138.
24. Rubin DB. Estimating casual effects of treatments in randomized and nonrandomized studies. J Ed Psych 1974; 66:688–701.
25. Holland P. Statistics and causal inference. J Am Stat Assoc 1986; 81:945–970.
26. Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. J Am Stat Assoc 1996; 91:444–455.
27. Seber GAF. Linear regression analysis. New York, NY: Wiley, 1997; 48–51.
28. Carroll RJ, Ruppert D. Transformation and weighting in regression. New York, NY: Chapman & Hall, 1988; 2–61.
29. Box GEP, Cox DR. An analysis of transformation. J R Stat Soc Series B 1964; 42: 71–78.
30. Tello R, Crewson PE. Hypothesis testing II: means. Radiology 2003; 227:1–4.
31. Mudholkar GS, McDermott M, Scrivastava DK. A test of p-variate normality. Biometrika 1992; 79:850–854.