APPENDIX 21

STATISTICAL EVALUATION METHODS

SLIPPERY ROCK CREEK PA STATE GAME LANDS #95

PROJECT SL-110-7-101.5

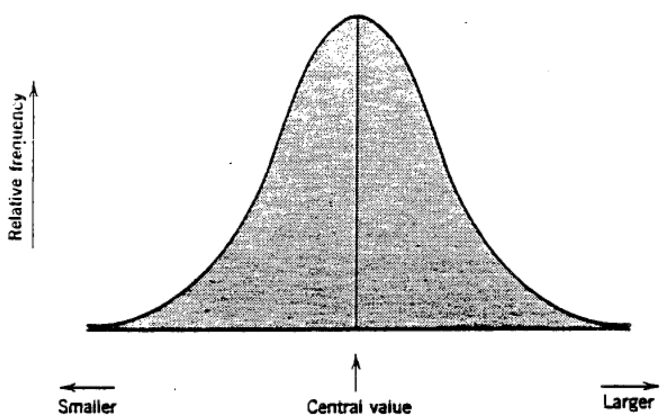
3. Statistical Evaluation Methods

Determination of Means

The mean value of a population is defined as the sum of all observations divided by the number of observations as shown below:

$$M = \frac{N = 1 (Xi)}{n}$$

The mean is another word for the arithmetic average. For populations with a normal frequency distribution, the mean is located at the center with most values clustered around the mean and the frequency of occurrence decreasing away from this central point. This is shown graphically below:



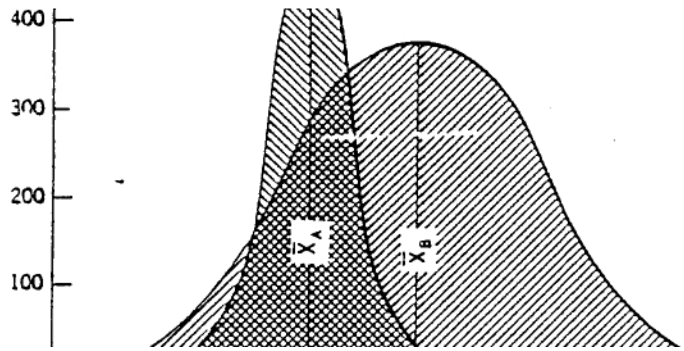
Plot of the normal frequency distribution.

The total area beneath the curve can be defined as being equal to 1.00 (or as 100% of the population distributed around the mean). The dispersion about the population mean can be expressed and standardized with respect to the mean using the normal distribution. This permits calculation of probabilities directly from the curve and allows graphic representations o£ sampling efficiency, proportionality, or range.

Methods of Measuring Dispersion

The dispersion (spread) about the me---,, while being normally distributed, is unique for every copulation. This can be shown graphically below where the dispersion around the mean of A(XA) is considerably less than the dispersion around the mean of B(XB).

\1k



While numerous methods are available for measuring the dispersion, only two methods have achieved widespread use and acceptance. They are:

1. Variance

Variance may be regarded as the average squared deviation of all possible observations from the mean of a population as defined below:

$$\sigma^2 = \frac{\sum_{i=1}^{n} (X_i - \mu)^2}{n}$$

Two algebraically equivalent forms are generally used to avoid the necessity of doing N subtractions, N multiplications, and N summations. The forms of the equation then become:

$$s^2 = \frac{\sum_{i=1}^{n} X_i^2 - n\overline{X}^2}{n-1}$$

$$s^2 = \frac{n \sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2}{n(n-1)}$$

The variance may also be calculated by using an intermediate quantity SS(the corrected sum of squares) as given below:

2. Standard Deviation

Standard deviation is defined as the square root of the variance. The advantages of using standard deviations over variance measures are:

a. The standard deviation is expressed in the units of measurement of the data.

b. The standard deviation is easily calculated from the variance.

c. The areas (expressed as a ,percentage of the total possible observations) under the curve of a normally distributed ,copulation can be precisely calculated for a given range.. That is:

(1) Standard Deviation = 68.3%
(2) (2) Standard Deviations = 95.4%
(3) (3) Standard Deviations = 99.7%
This is shown graphically below:

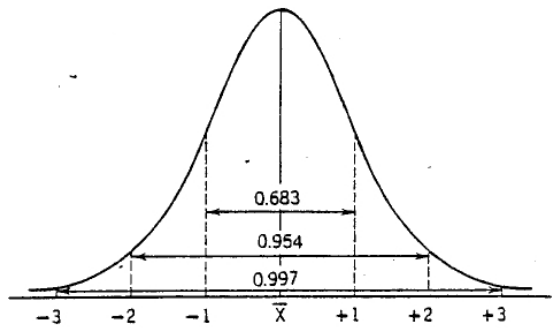


FIGURE 3.7. Areas enclosed by successive standard deviations of the standard normal distribution.

We might wish to restate o. above by relating this to a hypothetical sampling program where 100 samples were collected. First, the mean would be calculated, then by definition, the magnitudes of the first standard deviation would enclose 68 (68%) of our samples (34 on each side of the mean); 95 (95%) of our samples would be enclosed by the second standard deviation; and 5 (5%) of our samples would have-values greater than the second standard deviation and would lie outside that area.

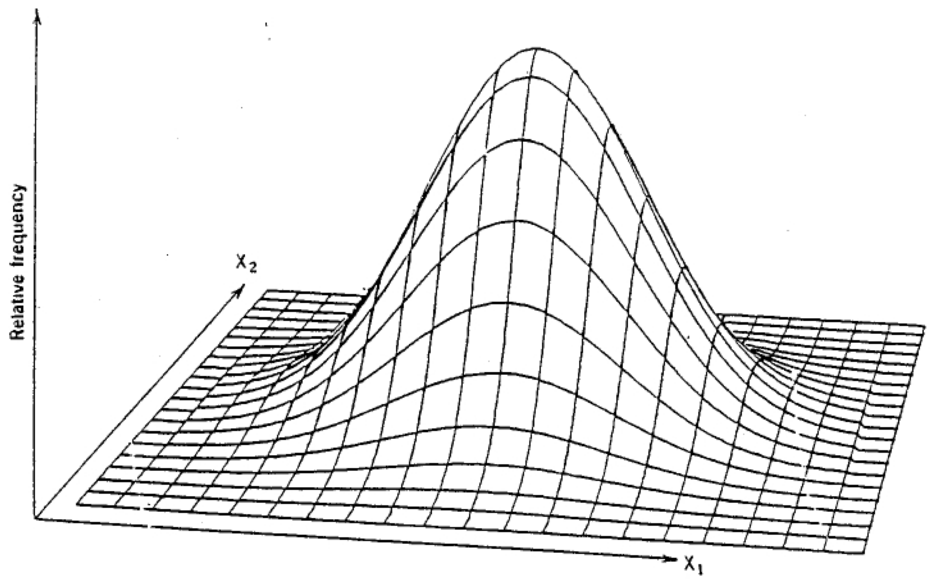
Methods of Comparing Dispersions

The comparing and testing of the properties of dispersion of two well-defined populations has been the aim of numerous statistical methods. The three methods in widespread use are:

1. Covariance

Covariance measures the joint variation of two variables about their common mean. The equation used for calculation purposes is given below:

An example of the joint variation of two variables is shown graphically below:



Joint probability distribution of two independent normal distributions. Both X_1 and X_2 are normally distributed.

Covariance results are interpreted much the same as variance results such that:

a. They may be compared against other covariance results but are not expressed in units similar to the initial measures.

b. They provide a useful intermediate quantity for inclusion in calculation of more efficient comparison methods.

2. <u>Correlation</u>

Correlation is the ratio of the covariance of two variables to the product of their standard deviations. The equations shown below are the definitional equation and the computational equation.

$$r_{jk} = \frac{COV_{jk}}{s_j s_k}$$

$$r_{jk} = \frac{SP_{jk}}{\sqrt{SS_j \cdot SS_k}}$$

$$= \frac{\sum_{i=1}^{n} X_{ij}X_{ik} - \left(\sum_{i=1}^{n} X_{ij} \sum_{i=1}^{n} X_{ik}\right)/n}{\sqrt{\left\{\sum_{i=1}^{n} X_{ij}^2 - \left[\left(\sum_{i=1}^{n} X_{ij}\right)^2/n\right]\right\}\left\{\sum_{i=1}^{n} X_{ik}^2 - \left[\left(\sum_{i=1}^{n} X_{ik}\right)^2/n\right]\right\}}}$$

The advantages of using correlation measures instead of covariance measures are:

a. The calculation of correlation estimates the interrelation between variables in a manner not influenced by measurement units.

b. Because the correlation is a ratio, it is a unitless number. However, it has definite range (between +1.00 and -1.00) which may equal but not exceed the product of the standard deviation of its variables.

c. The magnitude and sign of the correlation (hereinafter called the. <u>correlation coefficient)</u> can be interpreted as follows

(1) <u>positive correlation coefficient</u> - direct relationship between variables

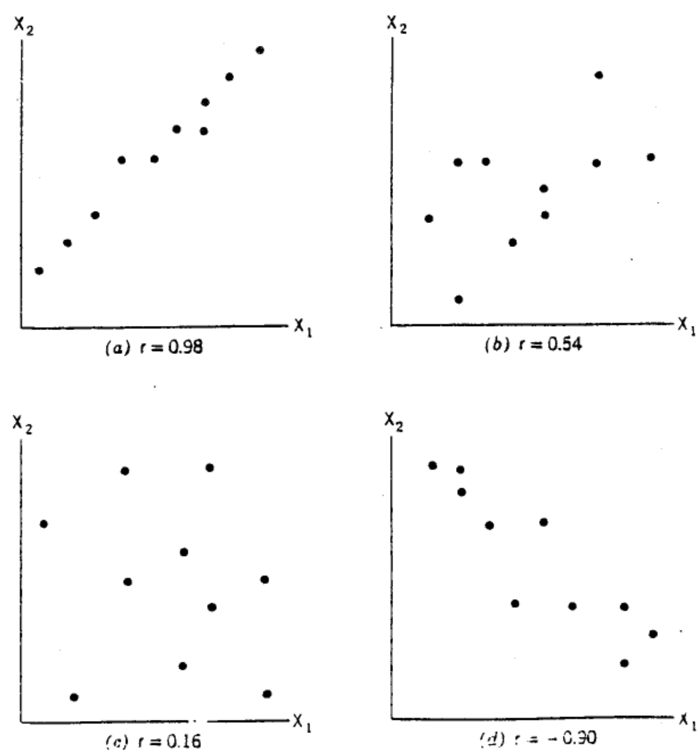
(2) <u>negative correlation coefficient</u> - inverse relationship between variables

(3) <u>zero correlation coefficient</u> - no relationship

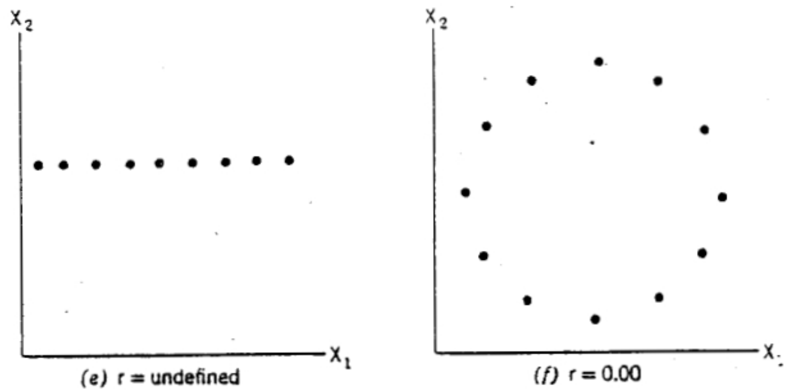
(4) <u>range of .01 to 1.00</u> - <u>direct</u> relationship where the magnitude of the coefficient can be expressed as percent, where +1.00 = perfect direct relationship

(5) <u>range of -.01</u> <u>to -1.00</u> inverse relationship where the magnitude of the coefficient can be expressed as percent, where - 1.00 = perfect inverse relationship

(6) Some examples of correlation coefficients and the plot of the mutual variables are shown graphically below:



(a) $r = 0.98$ (b) $r = 0.54$ (c) $r = 0.16$ (d) $r = -0.90$



d. The correlation coefficient is an expression of the LINEAR relationship between two variables, therefore, the line of dependence between the variables must be linear.

e. The non-linear relationships, an algebraically equivalent and analagous correlation is calculated. This is the multiple correlation coefficient and this can be interpreted in the same fashion as the correlation coefficient; however, this is a corrected coefficient which is not subject to linear relationships only.

3. F - Test Ratios

The F-distribution is based on a probability distribution and tests the equality of variances by comparing:

a. The theoretical distribution of values that would be expected by randomly sampling from a normal population, and expressing the result as a cut-off (critical) value for a predetermined level of significance.

b. The ratio(s) of sample variances for all possible pairs, as shown below:

$$F = \frac{s_1^2}{s_2^2}$$

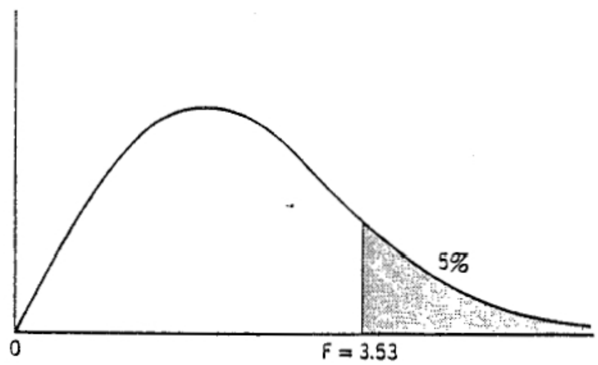
where s_1^2 is the larger variance and s_2^2 is the smaller. We now are testing the hypothesis

$$H_0: \sigma_1^2 = \sigma_2^2$$

against

$$H_1: \sigma_1^2 \neq \sigma_2^2$$

Stated simply, the F-test Treasures the ratio(s) of the sample variance and compares them to a standard Which would be expected if the samples came from the same population. Consider the graph below where a hypothetical F-distribution is given:



A typical F distribution with ν_1 = 10 and ν_2 = 25 degrees of freedom, with critical region (shown by shading) which contains 5% of the area under the curve. Critical value of F = 3.53.

a. If the sample ratios <u>exceeds</u> 3.53, the samples were not drawn from the sane normal population.

b. If the sample ratios <u>do not exceed</u> 3.53, then the samples are drawn from the same normal population.

c. Selection of Critical Values for F-tests

The shape of the F-distribution will change with changes in sample size. Since the F-distribution comperes a ratio of too sample variances to a probability distribution, an unbiased quantity called <u>degrees of freedom</u> is used to determine the critical value of F. The use of the degrees of - freedom <u>compensates</u> for the double use of observations by imposing a reduction of the total number.

<u>Degrees of Freedom</u> can be defined as the number of observations in a sample minus the number of parameters estimated from the sample. Restated, the degrees of freedom represents the number of independent comparisons that can be made between the observations in the estimating sample.

d. Selection of Level of Significance

The level of significance used in a F-distribution is expressed in %. This is a measure of the error associated with the comparison between the probability distribution and the F-test ratio. That is, if you were using a F-test to verify the appropriateness of a model and you wanted 95% accuracy, then the level of significance (acceptable error in the model) would be 0.05%. The level of significance in modeling often closely coincides to the standard deviation being utilized when describing the model variance, that is

F-test (.05 significance) - Standard Deviation - 2 (95%) F-test (.01 significance) - Standard Deviation - 3 (99%)

Methods of Comparing and Testing L' Dispersions for Multiple Populations

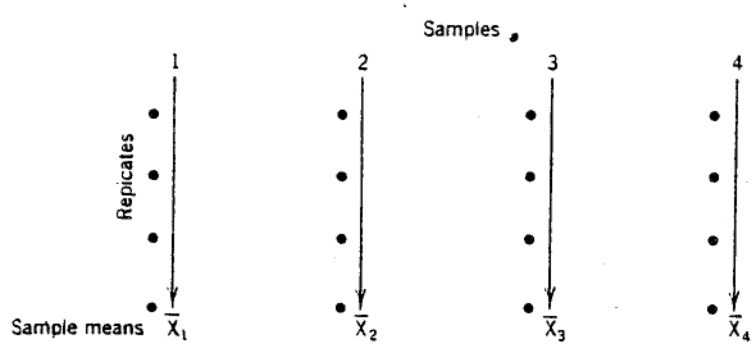
The carving and testing of the properties of dispersion of many well defined populations is generally approached by utilizing methods in the branch of statistics called analysis of variance. Generally, this involves the separation of the total variance of a group of measurements into the various components and sources. The tests of equality relate the simultaneous differences in means and in variances. Currently, two types of analysis of variance are widely used. They are:

1. One-Way Analysis of Variance

A one-way analysis of variance tests the:

a. Variance within each set of replicates b. Variance among the samples

The pattern of summation for one-way analysis of variance is given below:



Pattern of summation in analysis of variance. (a) One-way analysis; summation proceeds down replicates to find sample means.

The standardized format for presentation of results of one-way analysis of variance is given below:

Source of variation	Sum of squares	Degrees of freedom	Mean squares	F test
Among samples	SS_A	$m-1$	MS_A	MS_A/MS_W
Within replications	SS_W	$N-m$	MS_W	
Total variation	SS_T	$N-1$		

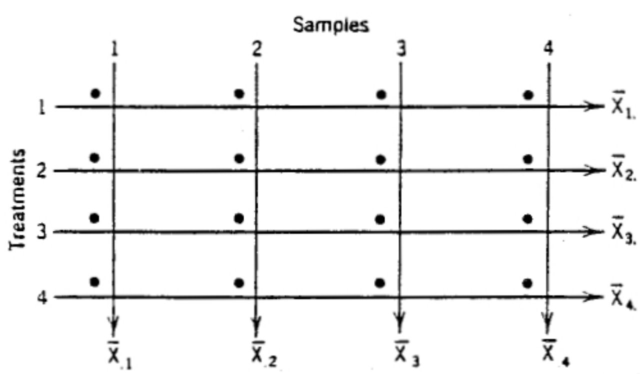
We will use a slightly modified version of the results format in the Examination of the linear regression program (which will be discussed in a later narrative).

2. Two-Way Analysis of Variance

A two-way analysis of variance tests the:

a. Variance among the treatments
b. b. variance among the samples

The pattern for summation of two-way analysis of variance is given below:



Pattern of summation in analysis of variance.

(*b*) Two-way analysis, summation proceeds down treatments to find sample means and also across samples to find treatment means.

The standardized format for presentation of results of two-way analysis of variance is given below:

Source of variation	Sum of squares	Degrees of freedom	Mean squares	F tests
Among samples	SS_A	$m-1$	MS_A	MS_A/MS_E [a]
Among treatments	SS_B	$n-1$	MS_B	MS_B/MS_E [b]
Error	SS_E	$(m-1)(n-1)$	MS_E	
Total variation	SS_T	$N-1$		

[a] Test of significance of differences between samples.
[b] Test of significance of differences between treatments.

We will utilize a slightly modified version of the results format for two-gray analysis of variance in the analysis of the curvilinear regression program (which will be discussed in a later narrative).

<u>Regression Analysis</u>

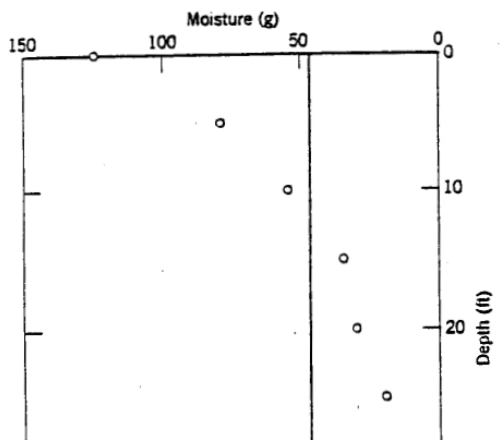
The regression analysis provides a method where the mean can be expressed <u>as a function</u> rather than a single value. This is accomplished by using simultaneous equations and minimizing the deviation of the predicted values from the actual values and by allowing the mean to have a slope which more closely matches the actual data. The regression mean permits the general tendency of the data to be analyzed and predicted. Since the mean is expressed as a function, appropriate values can be evaluated and mean line can be interpolated between data points; extrapolated beyond the sequence of data and estimate characteristics of the population.

The nature of the function is to minimize the deviations about the regression mean. Two of the most utilized regressions are:

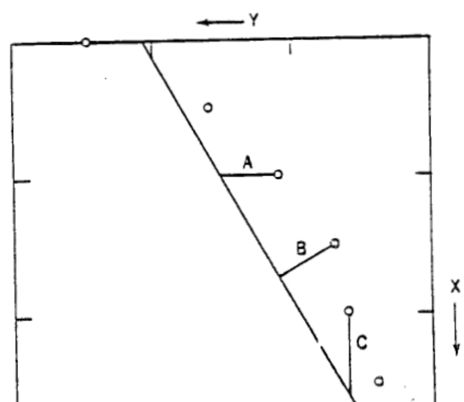
1. linear regression (line fitting)

2. curvilinear regression (curve fitting)

As the names imply, a linear regression expresses the mean as a function which graphs as a straight line. The advantage is that the mean line may have any slope. This advantage can be shown graphically below:



Plot of moisture content (grams water/100 grams dry weight) versus depth below sediment–water interface. Data collected from a core through Recent estuary mud in Louisiana bay. Note that orientation of plot corresponds to correct geologic orientation and not to standard mathematical form.



Possible criteria for minimization of deviations from fitted lines: A, minimization of deviations in moisture content; B, minimization of joint deviations; C, minimization of deviations in depth.

Note the increase in fit to the data for the sloped mean line (right).

However, when a function which graphs as a curved line is fitted to the same data, note the further increase in the appropriateness of the fit, as sown below

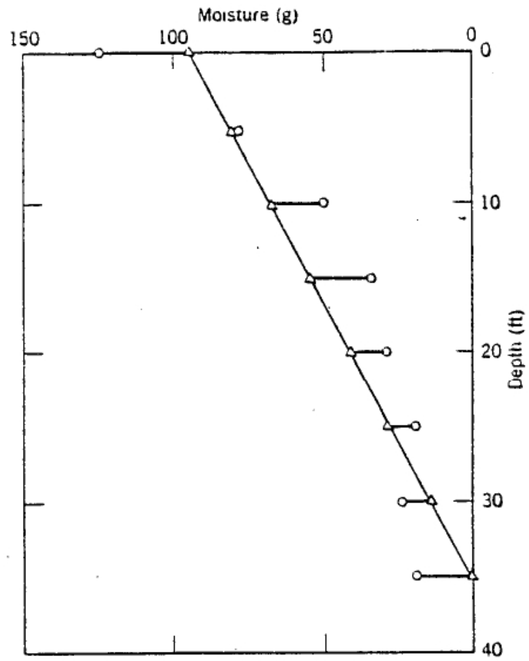


FIGURE 5.7. Observed values of moisture content and estimated values predicted by a straight line fitted by least squares.

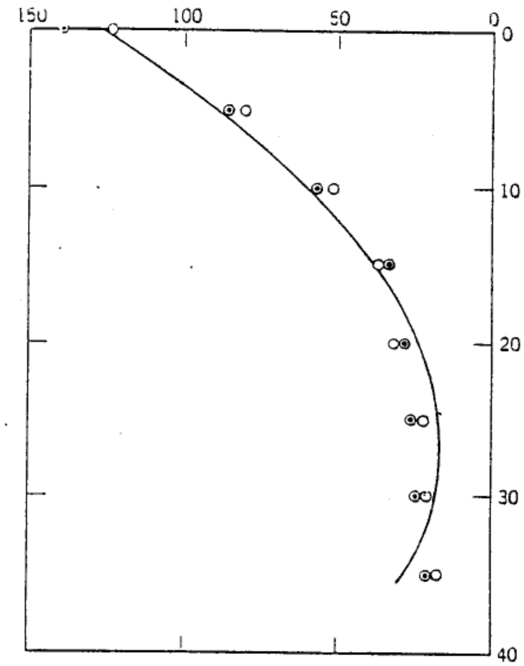


FIGURE 5.12. Second-degree polynomial regression fit to moisture data from Tables 5.7 and 5.10.

1. Testing Linear Regressions

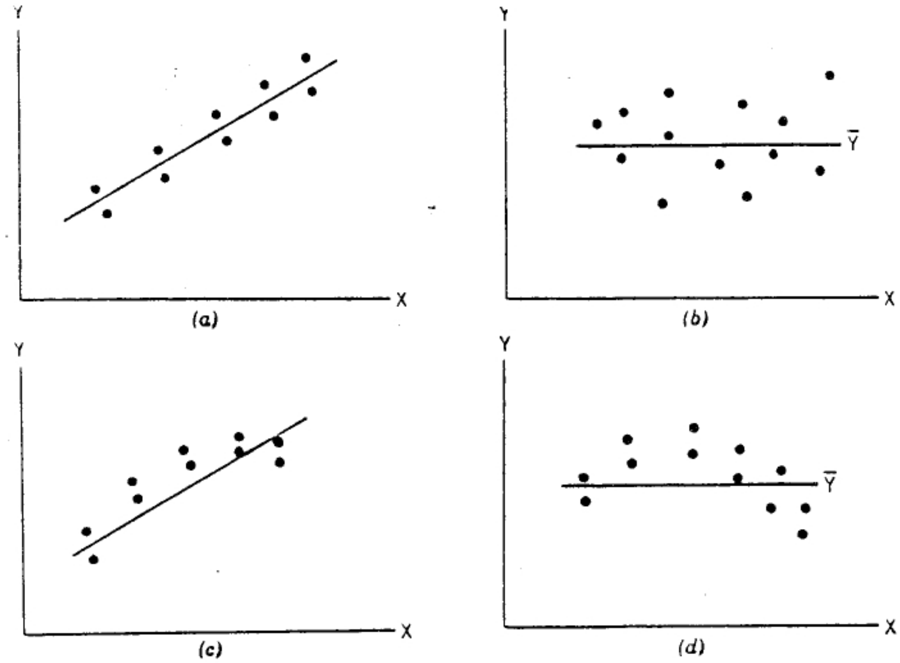
Linear regressions can be tested by analysis of variance procedures as previously described. The format for the analysis of variance is given below:

The Ftest ratio compares the variance about the regression line to the variance of, the population as estimated by the properties of the samples fan the population..

ANOVA for Simple Linear Regression

Source of variation	Sum of squares	Degrees of freedom	Mean squares	F test
Linear regression	SS_R	1	MS_R	MS_R/MS_D
Deviation	SS_D	$n-2$	MS_D	
Total variation	SS_T	$n-1$		

Four possible relationships axis for straight line regressions when they are tested,as shown below:



Possible straight-line regression situations. (*a*) Significant linear regression, no lack of fit. (*b*) Linear regression not significant, no lack of fit. (*c*) Significant linear regression, significant lack of fit. (*d*) Linear regression not significant, significant lack of fit. (After Draper and Smith, 1966.)

F-test ratios will not exceed the critical value for graphs A and B; while Graphs C and D will have F-test ratios in the critical region as the regression is not significant.

We compensate for the "anomalous" fit at graph B by utilizing a quantity called the "goodness of fit," which is defined by the equation:

The goodness of fit is somewhat analogous to the covariance; as it is a ratio of the variance measures of the regression to the total population (as estimated from the samples.) The square root of the goodness of fit is a term called the multiple corre-lation coefficient as described below:

$$R = \sqrt{R^2} = \sqrt{SS_R/SS_T}$$

This definition is algebraically equivalent to the definition of the correlation

$$r = \frac{SS_{xy}}{\sqrt{SS_x \cdot SS_y}}$$

This equation is algebraically equivalent to the correlation coefficient, previously discussed.

2. Testing Curvilinear Regressions

The F-test and correlation coefficients will allow the reviewer to determine the suitability of a linear regression. The linear regression takes the form of a normal equation where:

$$Y_i = \beta_0 + \beta_1 X_i$$

If additional terms of B are to be estimated as a result of testing of the <u>linear</u> regression, then a polynomial expansion (which will graph as a curved line) will be required, as shown below

$$Y_i = b_0 + b_1 X_i + b_2 X_i^2 + b_3 X_i^3 + \cdots + b_m X_i^m$$

The addition of each term of B (i. e. b_2X^2, b_3X^3) allows for increased flexibility to fit the data. However, a loss of degrees of freedom <u>decreases</u> the F-test critical value. This trade off means that while increased flexibility is gained, once the F-test value has been exceeded, no further significant estimation can be made.

Following completion of each curvilinear regression, the regression is tested and the results are displayed as shown below:

TABLE 5.12. ANOVA for Significance of Added Terms in Curvilinear Regression

Source of variation	Sum of squares	Degrees of freedom	Mean square	F test
Linear regression	SS_{R1}	1	MS_{R1}	MS_{R2}/MS_{D2}[a]
Quadratic regression	SS_{R2}	2	MS_{R2}	
Addition by quadratic	SS_{2-1}	1	MS_{2-1}	MS_{2-1}/MS_{D2}[b]
Quadratic deviation	SS_{D2}	$n-3$	MS_{D2}	
Total variation	SS_T	$n-1$		

[a]Tests for significance of the quadratic fit.

[b]Tests for significance of increase of quadratic over linear fit.