A NOTE ON REGRESSION ANALYSIS

By Abraham Wald
Columbia University

1. Introduction. In regression analysis a set of variables y, $x_1, \cdots, x_p$ is considered where y is called the dependent variable and $x_1, \cdots, x_p$ are the independent variables. Let y_α denote the αth observation on y and $x_{i\alpha}$ the αth observation on x_i, (i $= 1, \cdots, p$; $\alpha = 1, \cdots, N$). The observations $x_{i\alpha}$ are treated as given constants, while the observations $y_1, \cdots, y_N$ are regarded as chance variables. The following two assumptions are usually made concerning the joint distribution of the variates $y_1, \cdots, y_N$:

(a) The variates $y_1, \cdots, y_N$ are normally and independently distributed with a common unknown variance σ^2.

(b) The expected value of y_α is equal to $\beta_1 x_{1\alpha} + \cdots + \beta_p x_{p\alpha}$ where $\beta_1, \cdots, \beta_p$ are unknown constants.

In some problems it seems reasonable to assume that the regression coefficients $\beta_1, \cdots, \beta_p$ are not constants, but chance variables. This leads to a different probability model for regression analysis and the object of this note is to discuss certain aspects of this model. In what follows in this note we shall make the following assumptions concerning the joint distribution of the chance variables $y_1, \cdots, y_N; \beta_1, \cdots, \beta_p$.

Assumption 1. For given values of $\beta_1, \cdots, \beta_p$ the joint conditional probability density function of $y_1, \cdots, y_N$ is given by

$$\frac{1}{(2\pi)^{N/2}\sigma^N} \exp\left[-\frac{1}{2\sigma^2}\sum_{\alpha=1}^{N}(y_\alpha - \beta_1 x_{1\alpha} - \cdots - \beta_p x_{p\alpha})^2\right] \tag{1.1}$$

Assumption 2. The regression coefficients $\beta_1, \cdots, \beta_p$ are independently distributed.

Assumption 3. The regression coefficients $\beta_1, \cdots, \beta_r$, $(r \leq p)$, are normally distributed with zero means and a common variance σ'^2.

The purpose of this note is to derive confidence limits for the ratio $\frac{\sigma'^2}{\sigma^2}$. Such confidence limits have been derived by the author [1] for analysis of variance problems assuming that there are only main effects but no interactions. The regression problem treated in the present note is much more general and includes all the analysis of variance problems with or without interactions as special cases.

It should be remarked that Assumptions 2 and 3 do not exclude the case where $\beta_{r+1}, \cdots, \beta_p$ are constants.

2. Derivation of confidence limits for the ratio $\frac{\sigma'^2}{\sigma^2}$.

Let $b_1, \cdots, b_p$ be the sample estimates of $\beta_1, \cdots, \beta_p$ obtained by the method of least squares. We shall denote the difference $b_i - \beta_i$ by ϵ_i, $(i = 1, \cdots, p)$. It is known that for given values of $\beta_1, \cdots, \beta_p$ the conditional joint distribution of $\epsilon_1, \cdots, \epsilon_p$ is normal with zero means and variance-covariance matrix $\| c_{ij} \| \sigma^2$ where

$$(2.1) \qquad \| c_{ij} \| = \| a_{ij} \|^{-1}$$

and

$$(2.2) \qquad a_{ij} = \sum_{\alpha=1}^{N} x_{i\alpha} x_{j\alpha}, \qquad (i, j = 1, \cdots, p).$$

Since the conditional distribution of $\epsilon_1, \cdots, \epsilon_p$ does not depend on the values of $\beta_1, \cdots, \beta_p$, the unconditional distribution of $\epsilon_1, \cdots, \epsilon_p$ is the same as the conditional one, and the set of variates $(\beta_1, \cdots, \beta_p)$ is independently distributed of the set $(\epsilon_1, \cdots, \epsilon_p)$. From this and Assumptions 2 and 3 it follows that $b_1, \cdots, b_r$ have a joint normal distribution and that

$$(2.3) \qquad Eb_i = 0, \qquad (i = 1, \cdots, r)$$

and

$$(2.4) \qquad Eb_i b_j = \left(c_{ij} + \delta_{ij} \frac{\sigma'^2}{\sigma^2} \right) \sigma^2, \qquad (i, j = 1, \cdots, r)$$

where $\delta_{ij} = 0$ for $i \neq j$ and $= 1$ for $i = j$.

We shall denote $\frac{\sigma'^2}{\sigma^2}$ by λ and the elements of the inverse of $\| c_{ij} + \delta_{ij}\lambda \|$ by $d_{ij}(\lambda)$, i.e.,

$$(2.5) \qquad \| d_{ij}(\lambda) \| = \| c_{ij} + \delta_{ij}\lambda \|^{-1}, \qquad (i, j = 1, \cdots, r).$$

Then the quadratic form

$$(2.6) \qquad Q(\lambda) = \frac{1}{\sigma^2} \sum_{j=1}^{r} \sum_{i=1}^{r} d_{ij}(\lambda) b_i b_j$$

has the χ^2 distribution with r degrees of freedom.

It is known that for any given values of $\beta_1, \cdots, \beta_p, b_1, \cdots, b_p$ the quadratic form

$$(2.7) \qquad Q_a = \frac{1}{\sigma^2} \sum_{\alpha=1}^{N} (y_\alpha - b_1 x_{1\alpha} - \cdots - b_p x_{p\alpha})^2$$

has the χ^2 distribution with $N - p$ degrees of freedom provided that the rank of the matrix $\| x_{i\alpha} \|$ is p. Hence Q_a and $Q(\lambda)$ are independently distributed and the ratio

$$(2.8) \qquad F = \frac{N - p}{r} \frac{Q(\lambda)}{Q_a}$$

has the F-distribution with r and $N - p$ degrees of freedom.

Let F_1 and F_2 be two values chosen so that

$$(2.9) \qquad \text{Prob. } \{F_1 \leqq F \leqq F_2\} = c$$

where c is a given positive constant less than 1. Then the set of all values λ for which the inequality

$$F_1 \leq \frac{N - p}{r} \frac{Q(\lambda)}{Q_a} \leq F_2 \tag{2.10}$$

holds forms a confidence set for λ with the confidence coefficient c.

We shall now show that $Q(\lambda)$ is a monotonic function of λ and, therefore, the confidence set determined by (2.10) is an interval. Let $\| g_{ij} \|$, $(i, j = 1, \cdots, r)$, be an orthogonal matrix and let

$$b_i^* = \sum_{j=1}^{r} g_{ij} b_j. \tag{2.11}$$

It then follows from (2.3) and (2.4) that

$$E(b_i^*) = 0, \qquad (i = 1, \cdots, r) \tag{2.12}$$

and

$$E(b_i^* b_j^*) = (c_{ij}^* + \delta_{ij}\lambda)\sigma^2, \qquad (i, j = 1, \cdots, r) \tag{2.13}$$

where

$$c_{ij}^* = \sum_{l=1}^{r} \sum_{k=1}^{r} g_{ik} g_{jl} c_{kl}. \tag{2.14}$$

Let

$$\| d_{ij}^*(\lambda) \| = \| c_{ij}^* + \delta_{ij}\lambda \|^{-1}, \qquad (i, j = 1, \cdots, r) \tag{2.15}$$

and put

$$Q^*(\lambda) = \frac{1}{\sigma^2} \Sigma\Sigma\, d_{ij}^*(\lambda) b_i^* b_j^*.$$

It is easy to verify that $Q^*(\lambda)$ is identically equal to $Q(\lambda)$. Hence, to prove the monotonicity of $Q(\lambda)$, it is sufficient to show that $Q^*(\lambda)$ is a monotonic function of λ. Since no restrictions as to the choice of the orthogonal matrix $\| g_{ij} \|$ are made, we shall choose it so that the matrix $\| c_{ij}^* \|$ becomes diagonal, i.e., $c_{ij}^* = 0$ for $i \neq j$, $(i, j = 1, \cdots, r)$. Then

$$d_{ij}^*(\lambda) = 0 \qquad \text{for } i \neq j \tag{2.16}$$

and

$$d_{ii}^*(\lambda) = \frac{1}{c_{ii}^* + \lambda}. \tag{2.17}$$

Hence

$$Q(\lambda) = Q^*(\lambda) = \frac{1}{\sigma^2} \sum_{i=1}^{r} \frac{b_i^{*2}}{c_{ii}^* + \lambda} \tag{2.18}$$

is a monotonically decreasing function of λ. The confidence set determined by (2.10) is, therefore, an interval.

The upper end point of the confidence interval is the root in λ of the equation

$$\frac{N-p}{r}\frac{Q(\lambda)}{Q_a} = F_1 \tag{2.19}$$

and the lower end point is the root in λ of the equation

$$\frac{N-p}{r}\frac{Q(\lambda)}{Q_a} = F_2. \tag{2.20}$$

If equation (2.20) has no root, the lower end point of the confidence interval is put equal to zero.

REFERENCE

[1] A. Wald, "On the analysis of variance in case of multiple classifications with unequal class frequencies", *Annals. of Math. Stat.*, Vol. 12 (1941).

ON THE SHAPE OF THE ANGULAR CASE OF CAUCHY'S DISTRIBUTION CURVES

By Aurel Wintner

The Johns Hopkins University

1. Let ξ be a *linear* random variable, that is, a random variable capable of values x represented by points of a line $-\infty < x < \infty$, and suppose, for simplicity, that ξ has a density of probability, $f(x)$. Then, subject to provisos of convergence, the series

$$F(x) = \sum_{n=-\infty}^{\infty} f(x+n)$$

represents a periodic function, of period 1, having the following significance: $F(x)$ is the density of probability of the *angular* random variable, say Ξ, which is obtained if all the states

$$\cdots,\quad \xi-2,\quad \xi-1,\quad \xi,\quad \xi+1,\quad \xi+2,\cdots$$

of the linear random variable are identified.

In other words, if a circle of unit circumference rolls from $-\infty$ to ∞ on the ξ-line, then every point of the circumference collects the various densities of probability attached to congruent points of the ξ-line, and a state of Ξ represents a point of the circumference. For a detailed study of the mapping $\xi \to \Xi$ or $f \to F$, cf. [2].

According to Poisson's summation formula, the Fourier constants of the periodic function $F(x)$ can be obtained by restricting u in $g(u)$ to an equidistant sequence of discrete values, where $g(u)$ denotes the Fourier transform of $f(x)$; cf., e.g., [5], p. 78 or [9], pp. 477–478.