A SIGNIFICANCE TEST FOR EXPONENTIAL REGRESSION

By E. S. Keeping

University of Alberta[1]

1. Summary. A general method of testing the significance of nonlinear regression, suggested by Hotelling, is adapted to the regression equations $Y = be^{px}$ and $Y = a + be^{px}$. The values of x are taken to be in arithmetic progression, and the standard deviation of the observed y is supposed constant for all x. This is in contrast to the assumption, implicit in the usual procedure of fitting a straight line to $\log y$, that the standard deviation of $\log y$ is constant.

It will be observed that the distribution of $y_1, y_2, \cdots, y_n$ must be such that the joint probability density for $y_1, y_2, \cdots, y_n$ is a function of $x_1^2 + x_2^2 + \cdots + x_n^2$, and this condition implies the assumption of normality. The null hypothesis is that $be^{px} = 0$ for all x, while the alternative hypotheses are specified by $b \neq 0, p \neq -\infty$.

The method involves the calculation of the volume of a "tube" on a hypersphere in n-dimensional space. An asymptotic expression for the length of the tube is developed, and it is shown that the curvature of the axis is everywhere finite. From this expression, for values of the correlation coefficient R between observed and fitted values of y at least as great as 0.894, a function of R is obtained giving the probability that a random sample would yield at least as great a value of R.

A short table giving R for various significance levels and various sizes of sample is calculated for each of the equations mentioned, and the application to certain experimental data is discussed.

2. Introduction. Some years ago, Hotelling [1] suggested a geometrical method of determining the significance of the correlation coefficient corresponding to a fitted regression of y upon x, when y is a random variable and the values of x are known. Suppose that a curve of the form

$$Y = bf(x, p), \tag{2.1}$$

where b, p are constants and $f(x, p)$ is not identically zero, is fitted to a set of observations $y_1, y_2, \cdots, y_n$, which are assumed to be independently and normally distributed about zero, with the same variance σ^2. The null hypothesis is that $b = 0$, while the alternative hypothesis is that b is not zero. By the principle of least squares, we minimize

$$\Sigma (y_\alpha - Y_\alpha)^2 = \Sigma [y_\alpha - bf(x_\alpha, p)]^2. \tag{2.2}$$

The set of values $y_1, \cdots, y_n$ defines a point in n-dimensional space. The set

[1] Research carried out at the Institute of Statistics, University of North Carolina.

$Y_1, \cdots, Y_n$ also defines a point which lies on the 2-dimensional hypersurface defined parametrically in terms of b, p by the n equations:

$$Y_\alpha = bf(x_\alpha, p), \qquad \alpha = 1, 2, \cdots, n. \tag{2.3}$$

If θ is the angle between the lines joining the origin to $(y_1, \cdots, y_n)$ and $(Y_1, \cdots, Y_n)$,

$$\cos \theta = \Sigma\, y_\alpha Y_\alpha / [\Sigma\, y_\alpha^2 \,\Sigma\, Y_\alpha^2]^{\frac{1}{2}}, \tag{2.4}$$

and this is the correlation coefficient R between the observed and fitted values, calculated without elimination of the mean. The least squares process is thus equivalent to maximizing R, or minimizing θ.

Since by the null hypothesis the point $(y_1, \cdots, y_n)$ lies with uniform probability density anywhere on the surface of a sphere whose centre is the origin, the density function of the projection of the point $(y_1, \cdots, y_n)$ on the unit hypersphere has complete spherical symmetry around the origin. This will also be true if the joint probability density for $y_1, y_2, \cdots, y_n$ is any function of $x_1^2 + x_2^2 + \cdots + x_n^2$, and so is constant on the hypersphere $x_1^2 + \cdots + x_n^2 = 1$. The probability that R is greater than some fixed value R_0 is therefore, for a given Y, proportional to the "volume" of the sphere, in the $(n - 1)$-dimensional spherical space, having centre Y and geodesic radius $\theta_0 = \cos^{-1} R_0$. The total probability that R lies between R_0 and 1 is therefore given by the ratio of the "volume" of the "tube" of geodesic radius θ_0, surrounding the curve formed by the projection of Y, to the total "area" of the unit hypersphere.

Hotelling [1] has shown that the volume of such a tube on a hypersphere in n-dimensional space is equal to the length of the curve multiplied by

$$\pi^{\frac{1}{2}(n-2)} \sin^{n-2} \theta_0 / \Gamma(n/2),$$

provided that the curve is closed, and nowhere has a radius of geodesic curvature less than $\sin \theta_0$, and provided also that portions of the tube corresponding to nonconsecutive arcs of the axial curve do not overlap. If the curve has ends, there will be hemispherical "caps" at the ends of the tube to be added to the total volume.

This geometrical method was applied by D. M. Starkey [2] to the case of periodogram analysis, in which there are additional parameters, so that the projection of Y is not a curve but a surface. The practical application of the method in this case is limited to quite small values of θ_0, because of the approximations necessary in the evaluation of the integrals involved.

The 3-parameter equation

$$Y_\alpha = a + bf(x_\alpha, p) \tag{2.5}$$

is readily reducible, theoretically, to the form treated above. Minimizing $\Sigma\, [y_\alpha - a - bf(x_\alpha, p)]^2$ is equivalent to minimizing $\Sigma\, (y'_\alpha - Y'_\alpha)^2$, where y'_α, Y'_α are the projections of y_α, Y_α on the hyperplane $\Sigma\, y_\alpha = 0$. Since $y'_\alpha = y_\alpha - \bar{y}$ and $Y'_\alpha = b(f_\alpha - \bar{f})$, where f_α stands for $f(x_\alpha, p)$, the angle θ between the lines joining the origin to $(y'_1, y'_2, \cdots, y'_n)$ and $(Y'_1, Y'_2, \cdots, Y'_n)$ is given by

$$\cos\theta = \Sigma\,(y_\alpha - \bar{y})(Y_\alpha - \bar{Y})/[\Sigma\,(y_\alpha - \bar{y})^2\,\Sigma\,(Y_\alpha - \bar{Y})^2]^{\frac{1}{2}} \tag{2.6}$$

and so is equal to the correlation coefficient R between observed and fitted values, calculated in the usual way with elimination of the means. The point $(Y'_1, Y'_2, \cdots, Y'_n)$ lies on a 2-dimensional projection of the 3-dimensional hypersurface defined parametrically by (2.5). If we now project from the origin on to a hypersphere of $n - 2$ dimensions (intrinsically) in the hyperplane $\Sigma\, y_\alpha = 0$, the projection of $(Y'_1, Y'_2, \cdots, Y'_n)$ will be a point $(Y''_1, Y''_2, \cdots, Y''_n)$ lying on a curve on the surface of this hypersphere. The method already given therefore applies in this case, with the appropriate change in the dimensionality of the hypersphere.

The present paper deals with the application to exponential regression. The curves to be fitted are

$$Y = be^{px} \tag{2.7}$$

and

$$Y = a + be^{px}, \tag{2.8}$$

the latter of which will be referred to as the "modified exponential equation." The mathematical difficulties are increased greatly by the additional constant.

3. Formulas for projections of regression curves. We suppose that the fixed values of x are equidistant, and choose units so that

$$x_1 = 1, \qquad x_2 = 2, \qquad \cdots, \qquad x_n = n. \tag{3.1}$$

The corresponding values of Y are

$$Y_1 = bq, \qquad Y_2 = bq^2, \qquad \cdots, \qquad Y_n = bq^n, \tag{3.2}$$

where $q = e^p$. If the projections on the unit hypersphere are denoted by $Y'_1, \cdots, Y'_n$,

$$Y'_\alpha = \lambda q^\alpha, \qquad \Sigma Y'^2_\alpha = 1. \tag{3.3}$$

Hence

$$\lambda^2 = q^{-2}(1 - q^2)(1 - q^{2n})^{-1} \tag{3.4}$$

and

$$Y'_\alpha = q^{\alpha-1}\left(\frac{1 - q^2}{1 - q^{2n}}\right)^{\frac{1}{2}}. \tag{3.5}$$

The element ds of the curve formed by Y' $(0 \le q < \infty)$ is given by

$$\begin{aligned}(ds)^2 &= \Sigma(dY'_\alpha)^2 \\ &= \frac{1 - q^2}{1 - q^{2n}} \sum_{\alpha=1}^{n} \left\{\alpha - 1 + \frac{nq^{2n}}{1 - q^{2n}} - \frac{q^2}{1 - q^2}\right\}^2 q^{2\alpha-4}(dq)^2.\end{aligned} \tag{3.6}$$

This reduces, after some algebraic manipulation, to

$$\text{(3.7)} \qquad \frac{ds}{dx} = (2x)^{-1}\left[\frac{x}{(1-x)^2} - \frac{n^2x^n}{(1-x^n)^2}\right]^{\frac{1}{2}},$$

where $x = q^2$. The length of the projected curve is obtained by integrating from 0 to ∞.

For the modified exponential equation, we have, instead of (3.3),

$$\text{(3.8)} \qquad Y''_\alpha = \lambda(q^\alpha - f/n),$$

where

$$\text{(3.9)} \qquad f = \Sigma\, q^\alpha = q(1-q^n)(1-q)^{-1}.$$

Since $\Sigma\,(Y''_\alpha)^2 = 1$, we obtain

$$\lambda = (g - f^2/n)^{-\frac{1}{2}},$$

where

$$\text{(3.10)} \qquad g = q^2(1-q^{2n})/(1-q^2).$$

The expression for $(ds/dq)^2 = \Sigma\,(dY''_\alpha/dq)^2$ reduces after lengthy algebra to

$$\text{(3.11)} \qquad \left(\frac{ds}{dq}\right)^2 = \frac{1}{(1-q^2)^2} - \left\{\frac{\dfrac{nq^{n-1}}{1-q^n} - \dfrac{1}{n}\cdot\dfrac{1-q^n}{(1-q)^2}}{1+q^n - \dfrac{1}{n}\cdot\dfrac{1+q}{1-q}(1-q^n)}\right\}^2,$$

whence s may be obtained by integration.

4. Lengths of the projected curves. From (3.7)

$$\text{(4.1)} \qquad \begin{aligned} l_n &= \int_0^\infty [x(1-x)^{-2} - n^2x^n(1-x^n)^{-2}]^{\frac{1}{2}}\,dx/(2x) \\ &= \int_0^1 [x(1-x)^{-2} - n^2x^n(1-x^n)^{-2}]^{\frac{1}{2}}\,dx/x, \end{aligned}$$

since the substitution $x = 1/y$ leaves the integral unchanged.

For $n = 2$ the integral is elementary and reduces to $\pi/2$. For $n = 3$ it can be expressed in terms of elliptic integrals of the first and third kinds. For higher values of n a convergent series can be obtained, which, however, converges very slowly for n larger than 5.

Putting $x = (1-u)/(1+u)$, $g = 1 - u^2$, we obtain

$$\text{(4.2)} \qquad \begin{aligned} l_n &= 2\int_0^1 \left[\frac{1-u^2}{4u^2} - \frac{n^2(1-u^2)^n}{\{(1+u)^n - (1-u)^n\}^2}\right]^{\frac{1}{2}} \frac{du}{1-u^2} \\ &= 2\int_0^1 g^{-\frac{1}{2}}\left[\frac{1}{4(1-g)} - \frac{n^2g^{n-1}}{\sum_{t=0}^{2n}\binom{2n}{t}u^t - 2g^n + \sum_0^{2n}\binom{2n}{t}(-u)^t}\right]^{\frac{1}{2}} du, \end{aligned}$$

and this integral may be shown to exist for every finite n. For $n = 3$, we have

$$l_3 = \pi/2\,[1 + 1/4 + 9/256 + 5/512 + 385/262144 + \cdots]$$
$$= 2.037.$$

For $n = 4$, a similar method gives $l_4 = 2.35$, and for $n = 5$ we obtain $l_5 = 2.58$. However, as n increases, the method is more laborious and the integrand does not converge so rapidly, so that an asymptotic expression for l_n is more convenient.

For the case $n = 3$, (3.11) reduces to

$$\frac{ds}{dq} = \frac{\sqrt{3}}{2} \cdot \frac{1}{1 + q + q^2},$$

whence $l_3' = \pi/3$. That this is correct may be seen by visualizing the projection of the regression curve on to a circle in the plane through the origin which is equally inclined to all three axes. Writing $u = \tanh p/2$, we obtain

$$l_n' = \int_0^1 [1 - \varphi^2]^{\frac{1}{2}}\, du/u, \tag{4.3}$$

where

$$\varphi = \left[1 - \frac{4n^2u^2(1 - u^2)^{n-1}}{\{(1 + u)^n - (1 - u)^n\}^2}\right] \div \left[\frac{nu\{(1 + u)^n + 1 - u)^n\}}{(1 + u)^n - (1 - u)^n} - 1\right]. \tag{4.4}$$

It may be shown that the denominator of φ never vanishes and that $\varphi \leq 1$ in the region of integration.

When $n = 4$, the integral (4.3) is elliptic, and we find $l_4' = 1.418$.

$$\text{When} \quad n = 5, \qquad \varphi = \frac{25 - 5u^2 + 15u^4 - 3u^6}{25 + 65u^2 + 35u^4 + 3u^6}.$$

The integral may be evaluated by quadrature. The numerical value is $l_5' = 1.675$. For larger values of n, an approximation is obtained in the next section.

5. Approximations to the length. Putting $x = e^{-2v}$, we have from (4.1)

$$l_n = \int_0^\infty \left(\frac{1}{\sinh^2 v} - \frac{n^2}{\sinh^2 nv}\right)^{\frac{1}{2}} dv. \tag{5.1}$$

For values of v in the range 0 to $1/n$ we may write the integrand as a series and integrate term by term. Thus, if

$$I_1 = \int_0^{1/n} \left(\frac{1}{\sinh^2 v} - \frac{n^2}{\sinh^2 nv}\right)^{\frac{1}{2}} dv,$$

we obtain

$$I_1 = 0.559 - 0.298n^{-2} - 0.0594n^{-4} - \cdots . \tag{5.2}$$

Let

$$\begin{aligned} I_2 &= \int_{1/n}^{\infty} \left(\frac{1}{\sinh^2 v} - \frac{n^2}{\sinh^2 nv}\right)^{\frac{1}{2}} dv \\ &= 2\int_0^{e^{-1/n}} \frac{1}{1-u^2}\left[1 - \frac{n^2u^{2n-2}(1-u^2)^2}{(1-u^{2n})^2}\right]^{\frac{1}{2}} du. \end{aligned} \tag{5.3}$$

The second term in the square bracket is less than 1 at both ends of the range. Also, it has no maximum within the range. Hence the bracket may be expanded and integrated term by term, giving

$$\begin{aligned} I_2 &= 2\int_0^{e^{-1/n}} \frac{1}{1-u^2}\left[1 - \frac{n^2(1-u^2)^2u^{2n-2}}{2(1-u^{2n})^2} - \frac{n^4(1-u^2)^4u^{4n-4}}{8(1-u^{2n})^4} - \cdots\right] du \\ &= I_{21} + I_{22} + I_{23} + \cdots, \end{aligned}$$

where

$$I_{21} = 2\int_0^{e^{-1/n}} \frac{du}{1-u^2} = \log 2n + \frac{1}{12n^2} - \frac{1}{30n^3} + O\left(\frac{1}{n^4}\right). \tag{5.4}$$

Also,

$$\begin{aligned} I_{22} &= -n^2 \int_0^{e^{-1/n}} u^{2n-2}(1-u^2)(1-u^{2n})^{-2}\, du \\ &= -n^2 \int_0^{e^{-1/n}} [(u^{2n-2} - u^{2n}) + 2(u^{4n-2} - u^{4n}) + 3(u^{6n-2} - u^{6n}) + \cdots]\, du, \end{aligned}$$

which, on integrating, expanding the exponentials, and collecting terms, becomes

$$\begin{aligned} I_{22} = &-[(3/2)e^{-2} + (5/4)e^{-4} + (7/6)e^{-6} + (9/8)e^{-8} + \cdots] \\ &- 1/n^2[(19/24)e^{-2} + (71/192)e^{-4} + (61/216)e^{-6} + (379/1536)e^{-8} \\ &+ \cdots] + O(n^{-4}) \\ = &-0.229 - 0.114n^{-2} + O(n^{-4}). \end{aligned} \tag{5.5}$$

Similarly,

$$\begin{aligned} I_{23} &= -\frac{n^4}{4}\int_0^{e^{-1/n}} u^{4n-4}(1-u^2)^3(1-u^{2n})^{-4} du \\ &= -0.029 - 0.029n^{-2} + O(n^{-4}). \end{aligned} \tag{5.6}$$

Later terms in I_2 can be computed in a similar way, but the numerical factors diminish rapidly. Collecting terms from (5.2), (5.4), (5.5), (5.6), we get finally

$$l_n = \log n + 0.990 - 0.358n^{-2} + O(n^{-4}). \tag{5.7}$$

As an indication of the accuracy of this approximation, the value of l_n, neglecting terms of order n^{-4}, has been calculated in Table I for several values of n. It is not, of course, to be expected that the approximation will be very good for small n, although it is actually quite close even for $n = 3$ and $n = 4$.

TABLE I

Length of axis of tube ($Y = be^{px}$)

n	Asymptotic value of l_n	Exact value of l_n
2	1.59	1.57
3	2.05	2.04
4	2.35	2.34
5	2.59	2.59
6	2.77	
7	2.93	
8	3.06	
9	3.18	
10	3.29	
12	3.47	
15	3.70	
20	3.98	
50	4.90	
100	5.60	

For the modified exponential equation, the above method is apparently not practicable with the more complicated integral (4.3). However, it is possible to obtain an approximate expression which will be valid for large n, although the agreement with the numerical values for $n = 3$, 4, and 5 is not very close.

In terms of $v = p/2$, (4.3) may be written

$$l'_n = \int_0^\infty 2(\sinh 2v)^{-1}\left[1 - \left\{\frac{1 - n^2 \sinh^2 v(\sinh nv)^{-2}}{n \tanh v(\tanh nv)^{-1} - 1}\right\}^2\right]^{\frac{1}{2}} dv. \tag{5.8}$$

For values of v between 0 and $1/n$, the integrand may be expanded in powers of v and integrated term by term. The result is

$$I_1 = 0.500 - 1.08n^{-2} + O(n^{-4}). \tag{5.9}$$

For any fixed u between 0 and 1, the function φ in (4.4) tends to $(nu - 1)^{-1}$ as $n \to \infty$, so that the integrand in (4.3) tends to $u^{-1}[1 - (nu - 1)^{-2}]^{\frac{1}{2}}$. However, this approximation is clearly not useful for u near $1/n$. But if we put $u = k/n$, where k is a fixed integer, then, for large n, φ tends to the value

$$\varphi = \frac{e^k - e^{-k} - 4k^2(e^k - e^{-k})^{-1}}{(k-1)e^k + (k+1)^{-k}}. \tag{5.10}$$

For fairly large k, this is very close to $(k - 1)^{-1}$. Thus for $k = 7$, it is 0.1667. Hence from $u = 7/n$ to 1 we can approximate φ asymptotically by $(nu - 1)^{-1}$, and obtain

$$I_2 \sim \int_{7/n}^{1} [1 - (nu - 1)^{-2}]^{\frac{1}{2}} u^{-1} \, du = \int_{w_1}^{w_2} [1 - \operatorname{sech} w] \, dw,$$

where $nu - 1 = \cosh w$ and w_1, w_2 are the values of w corresponding to $u = 7/n$ and $u = 1$ respectively. Hence

$$(5.11) \qquad \begin{aligned} I_2 &\sim w_2 - w_1 - \tan^{-1} \sinh w_2 + \tan^{-1} \sinh w_1 \\ &= -1.952 + \log n + (1/4)n^{-2} + O(n^{-3}) \quad (n \geq 7). \end{aligned}$$

It remains to integrate between tanh $(1/n)$ and $7/n$. From $1/n$ to $7/n$, an approximation may be obtained by quadrature, the ordinates being calculated from (5.10) for values of k between 1 and 7 inclusive. This gives, by Simpson's rule, a value 1.573. A small correction may be made for the integral between tanh $(1/n)$ and $1/n$. Since $\tanh (1/n) = 1/n - (1/3)n^{-3}$, approximately, and since for $u = 1/n$, $u^{-1}[1 - \varphi^2]^{\frac{1}{2}} = 0.475n$ approximately, this integral will be $0.158n^{-2}$, neglecting terms of higher order.

Hence the final expression for the length of the axis of the tube is

$$(5.12) \qquad l_n' = \log n + 0.121 - 0.67n^{-2} + O(n^{-3}).$$

Table II gives a few numerical values of l_n', neglecting $O(n^{-3})$.

TABLE II

Length of axis of tube ($Y = a + be^{px}$)

n	Asymptotic value of l_n'	Correct value of l_n'
3	1.14	1.05
4	1.47	1.42
5	1.70	1.68
6	1.89	
7	2.05	
8	2.19	
9	2.31	
10	2.42	
15	2.83	
20	3.12	

6. Curvature of the projected curve. It was shown by Hotelling [1] that to avoid difficulties connected with local overlapping, or "kinking", of the tube surrounding the projected curve, it is necessary and sufficient that $\sin \theta \leq \rho$, where θ is the geodesic radius of the tube and ρ the radius of geodesic curvature of its axis. In this section we show that the radius of curvature is always finite and greater than $1/\sqrt{5} = 0.447$. The statement by Hotelling (*loc. cit.*, p. 452), that the radius of curvature of the projected curve corresponding to $Y = be^{px}$ becomes zero at $p = \pm\infty$, appears to be in error.

The radius of curvature ρ with which we are dealing is defined by

$$\begin{aligned}\rho^{-2} &= \Sigma_\alpha(d^2Y'_\alpha/ds^2)^2\\ &= (ds/dp)^{-6}[(ds/dp)^2\Sigma(d^2Y'_\alpha/dp^2)^2 + (d^2s/dp^2)^2\Sigma(dY'_\alpha/dp)^2\\ &\qquad - 2ds/dp\cdot d^2s/dp^2\Sigma(dY'_\alpha/dp\cdot d^2Y'_\alpha/dp^2)].\end{aligned} \tag{6.1}$$

Since

$$\Sigma(dY'_\alpha/dp)^2 = (ds/dp)^2, \quad \Sigma(dY'_\alpha/dp\cdot d^2Y'_\alpha/dp^2) = ds/dp\cdot d^2s/dp^2,$$

this reduces to

$$\rho^{-2} = (ds/dp)^{-4}[\Sigma(d^2Y'_\alpha/dp^2)^2 - (d^2s/dp^2)^2]. \tag{6.2}$$

In the present problem, $Y'_\alpha = \lambda e^{\alpha p}$, so that

$$d^2Y'_\alpha/dp^2 = e^{\alpha p}(d^2\lambda/dp^2 + 2\alpha\, d\lambda/dp + \alpha^2\lambda),$$

where

$$\lambda^{-2} = \Sigma e^{2\alpha p} = e^{2p}(e^{2np} - 1)(e^{2p} - 1)^{-1}.$$

After some reduction we obtain

$$\begin{aligned}\Sigma(d^2\, Y'_\alpha/dp^2)^2 &= \frac{3}{16}\left[\frac{1}{\sinh^2 p} - \frac{n^2}{\sinh^2 np}\right]^2\\ &\qquad + \frac{1}{8}\left[\frac{3 + 2\sinh^2 p}{\sinh^4 p} - n^4\,\frac{3 + 2\sinh^2 np}{\sinh^4 np}\right].\end{aligned} \tag{6.3}$$

From equation (3.7) we have, in terms of $p = \log q$ as parameter,

$$ds/dp = \tfrac{1}{4}[(\sinh p)^{-2} - n^2(\sinh np)^{-2}], \tag{6.4}$$

whence

$$d^2s/dp^2\cdot ds/dp = \tfrac{1}{4}[-\cosh p(\sinh p)^{-3} + n^3\cosh np(\sinh np)^{-3}]. \tag{6.5}$$

Therefore, from (6.2),

$$\begin{aligned}\rho^{-2} = 3 &+ \left[\frac{6 + 4\sinh^2 p}{\sinh^4 p} - n^4\,\frac{6 + 4\sinh^2 np}{\sinh^2 np}\right]\left[\frac{1}{\sinh^2 p} - \frac{n^2}{\sinh^2 np}\right]^{-2}\\ &- 4\left[\frac{\cosh p}{\sinh^3 p} - \frac{n^3\cosh np}{\sinh^3 np}\right]^2\left[\frac{1}{\sinh^2 p} - \frac{n^2}{\sinh^2 np}\right]^{-3}.\end{aligned} \tag{6.6}$$

Now as $p \to 0$, the right hand side tends to $3 - [6(n^2 + 1)]/[5(n^2 - 1)]$. For $n = 2$ this is 1, as it should be, and as $n \to \infty$ it approaches the value 9/5. It may be shown also that for $n \geq 4$, $\rho^{-2} \to 5 - 2e^{-2|p|}$ as $p \to \pm\infty$, so that $1/\rho^2 \approx 5$ for moderate values of p.

If $u = (\sinh^2 p)^{-1} - n^2 (\sinh^2 np)^{-1}$, the expression for ρ^{-2} can be written as

$$\rho^{-2} = 3 - u'^2 u^{-3} + u'' u^{-2}. \tag{6.7}$$

Hence

$$d(\rho^{-2})/dp = u^{-4}[3u'^3 - 4uu'u'' + u^2 u'''] . \tag{6.8}$$

By expanding in powers of e^{-p} it can be shown that the terms in e^{2p} and the constant term vanish, so that

$$d(\rho^{-2})/dp = O(e^{-2p}), \qquad p > 0,$$

for $n \geq 4$. Hence ρ^{-2} has no maximum or minimum at any point p, apart from the minimum at $p = 0$. The radius of curvature is therefore finite and remains between $1/\sqrt{5}$ and $[(5n^2 - 5)/(9n^2 - 21)]^{\frac{1}{2}}$, i.e., between 0.447 and 0.745, for any $n \geq 4$.

The condition for no local overlapping at any point of the tube is, therefore, $\sin\theta \leq 0.447$, or equivalently, $\cos\theta \geq 0.894$, where θ is the geodesic radius of the tube.

For the modified exponential curve, (6.2) still holds, with Y'_α replaced by $Y''_\alpha = (g - f^2/n)^{-\frac{1}{2}} (e^{\alpha p} - f/n)$. We now have

$$(ds/dp)^2 = (4 \sinh^2 p)^{-1} \left[1 - \frac{1 - n^2 \sinh^2 (p/2)(\sinh np/2)^{-2}}{n \tanh (p/2)(\tanh np/2)^{-1} - 1} \right]. \tag{6.9}$$

I have not been able to obtain an explicit expression for the curvature of the axis of the tube, similar to (6.6). However, for small values of p, Y''_α and ds/dp may be expressed in series of powers of p, and I find after much algebraic calculation that when $p \to 0$,

$$1/\rho^2 \to \frac{19n^4 - 212n^2 + 544}{7n^4 - 56n^2 + 112}. \tag{6.10}$$

For $n = 3$ this reduces to 1, as it should, since in this case the curve is an arc of a unit circle.

As $n \to \infty$, $\rho \to (7/19)^{\frac{1}{2}} = 0.607$, so that the radius of curvature at the centre of the axial curve of the tube lies between 0.607 and 1 for all values of $n > 3$.

To find the curvature at the ends of the axial curve we need the limit of $1/\rho$ as $p \to \pm\infty$. Since the curve is symmetrical about $p = 0$, it is sufficient to consider $p \to \infty$. For $n \geq 5$, it may be shown that, as $p \to \infty$,

$$1/\rho^2 \to \frac{5n^5 - 35n^4 + 78n^3 - 108n - 12}{n(n-1)(n-2)^3}. \tag{6.11}$$

As $n \to \infty$, $\rho \to 1/\sqrt{5} = 0.446$, as for the simpler case treated above. For $n = 6$, ρ has the value 0.519.

Intuitively, one would expect for $Y = a + be^{px}$, as we have shown for $Y = be^{px}$, that the curvature of the projection would vary monotonically between the centre and the ends. A proof of this would, however, be desirable. Assuming that this statement is true, we have as the condition for no local overlapping that $\cos\theta \geq 0.894$.

It is of interest to see how much of the length of the tube will have a radius of curvature near to the minimum value 0.477, which corresponds to $p = \pm\infty$. Now, for $n \geq 5$, and for large p, equation (6.6) can be written

$$\rho^{-2} = 5 - 2e^{-2|p|} + O(e^{-6|p|}), \tag{6.12}$$

so that for $|p| > 1$, the first two terms may be considered a fair approximation. If we take $|p| = 2\log 2 = 1.386$, $\rho^{-2} = 39/8$, or $\rho = 0.453$, approximately. The length of the part of the tube for which the radius of curvature lies between 0.447 and 0.453 is

$$2\int_0^{1/4} (1 - u^2)^{-1}[1 - n^2u^{2n-2}(1 - u^2)^2(1 - u^{2n})^{-2}]^{\frac{1}{2}}\, du,$$

which equals 0.511 approximately for $n \geq 5$. Since for $n = 5$, the whole length of the tube is 2.59, nearly one-fifth of this length has a radius of curvature between 0.447 and 0.453.

As n increases, the ratio of this part to the total length diminishes to zero, but for $n = 100$ it is still about 1/11. Hence it appears that local overlapping may be serious for values of $\cos\theta$ appreciably less than the critical value.

Nonlocal overlapping will not occur. It is necessary for such overlapping that the tube should bend around so that two points P_1 and P_2 of the axis are at a geodesic distance apart less than twice the radius of the tube even though they are separated by a considerably greater distance than this, measured along the axis.

If P_1 and P_2 correspond to values q_1 and q_2 of the parameter q, the square of the distance in Euclidean n-space between them is given by

$$D = \sum_\alpha (\lambda_2 q_2^\alpha - \lambda_1 q_1^\alpha)^2, \tag{6.13}$$

where

$$\lambda_i^2 = q_i^{-2}(1 - q_i^2)(1 - q_i^{2n})^{-1}, \qquad i = 1, 2.$$

If we transform to polar coordinates $\theta_1, \cdots, \theta_{n-1}$ on the unit hypersphere and let P_1 be the point $(0, 0, \cdots, 0)$ and P_2 the point $(\alpha, 0, \cdots, 0)$, then $D^2 = 2 - 2\cos\alpha$ and the geodesic distance between P_1 and P_2 is α. If, therefore, D is the distance from a fixed point q_1 to a variable point q,

$$D^2 = 2 - 2\,\frac{1 - q^n q_1^n}{1 - qq_1}\left\{\frac{(1 - q^2)(1 - q_1^2)}{(1 - q^{2n})(1 - q_1^{2n})}\right\}^{\frac{1}{2}}, \tag{6.14}$$

and a minimum value of D^2 corresponds to a maximum value of

$$\cos^2\alpha = \left(\frac{1 - q^n q_1^n}{1 - qq_1}\right)^2 \frac{(1 - q^2)(1 - q_1^2)}{(1 - q^{2n})(1 - q_1^{2n})}. \tag{6.15}$$

The ends of the axis of the tube are at a geodesic distance $\pi/2$ apart. One end of the axis is at the point where the positive x_1 coordinate axis cuts the unit hypersphere and the other end is at the point where the positive x_n axis cuts it. The axis of the tube lies wholly on that part of the surface of the hypersphere for which all coordinates are positive, and so cannot spiral around the end points or form an equatorial spiral around the sphere in the middle.

If there is nonlocal overlapping there must be at least three distinct roots of (6.15), considered as an equation in q corresponding to a given q_1 and a given α (less than twice the geodesic radius of the tube). Two of the roots will represent neighbouring points on the axis, one on each side of q_1, and the others, if they exist, correspond to points on nonlocal portions of the axis. If the point q_1 is at a geodesic distance less than α from either end of the axis of the tube, the existence of *two* distinct roots would imply nonlocal overlapping.

Since the tube is symmetrical about the middle of its axis (at $q = 1$) we can assume $0 \leq q_1 \leq 1$. Then q can take any real value from 0 to ∞. By the condition for the absence of *local* overlapping, $\cos^2 \alpha > 0.360$.

If $q_1 = 0$, the equation for q becomes

$$1 + q^2 + q^4 + \cdots + q^{2n-2} - \sec^2 \alpha = 0, \tag{6.16}$$

and this, by Descartes' rule of signs, has at most one real root. There is therefore no nonlocal overlapping at the ends of the tube, for any value of n. This is also true at the middle, where $q_1 = 1$.

For any finite n the geodesic distance β of q_1 from the end $q = 0$ is given by $1 - q_1^2 = (1 - q_1^{2n}) \cos^2 \beta$. If there is to be no nonlocal overlapping, the equation in q,

$$\left(\frac{1 - q^n q_1^n}{1 - qq_1}\right)^2 \left(\frac{1 - q^2}{1 - q^{2n}}\right) = \frac{\cos^2 \alpha}{\cos^2 \beta}, \tag{6.17}$$

should possess only one real positive root if $\beta < \alpha$ and only two real positive roots if $\beta > \alpha$. If $\beta = \alpha$, one root is $q = 0$. The equation in this case becomes

$$1 + q^2 + \cdots + q^{2n-2} = (1 + qq_1 + \cdots + q^{n-1}q_1^{n-1})^2. \tag{6.18}$$

Writing $y_1 = 1 + q^2 + \cdots + q^{2n-2}$ and $y_2 = (1 + qq_1 + \cdots + q^{n-1}q_1^{n-1})^2$, it is clear that y_1 and y_2 and all their derivatives are nonnegative for all q, that y_1 and y_2 are never zero, and that at $q = 0$, $y_2' > 0$. The curve of y_2 as a function of q starts together with that of y_1 at $q = 0$ and remains above the latter for an interval of $q > 0$. When q tends to infinity, $y_2/y_1 \to q_1^{2n-2}$, which is less than 1, so that the curves must eventually cross again. A real root of $y_1 = y_2$ greater than zero therefore exists. We shall now show that this root is unique.

Since $y_1 < y_1^2 < y_2$ for $0 < q < q_1$, we can confine our attention to the case $q > q_1$.

If $q_1 = 1$, (6.18) clearly has no real root for $q > 0$. Moreover, y_2 is a continuous function of q_1. Hence, if for any q_1, between 0 and 1, two real roots exist, there must be some q_1 and some corresponding q such that (6.18) has a *double* root.

It is sufficient therefore to show that such a double root cannot exist. The conditions are

$$(6.19) \qquad y_1 - y_2 = 0, \qquad y_1' - y_2' = 0.$$

Writing $y_1' = [n - q^2(1 + q^2 + \cdots + q^{2n-2})]/(1 - q^2)$ and $y_2' = \sqrt{y_2}[n - qq_1(1 + qq_1 + \cdots + q^{n-1}q_1^{n-1})]/(1 - qq_1)$, the second condition of (6.19) gives

$$(6.20) \qquad \frac{n - q^2 - q^4 - \cdots - q^{2n}}{n - qq_1 - q^2q_1^2 - \cdots - q^nq_1^n} \cdot \frac{1 - qq_1}{1 - q^2} = \sqrt{y_2}.$$

From the first condition,

$$\sqrt{y_2} = \frac{y_1}{\sqrt{y_2}} = \frac{1 + q^2 + \cdots + q^{2n-2}}{1 + qq_1 + \cdots + q^{n-1}q_1^{n-1}}.$$

Substituting for $\sqrt{y_2}$ in (6.20), subtracting 1 from both sides, removing a common factor $q^2 - qq_1$, cross-multiplying and collecting terms, we arrive at an equation of the form

$$A + Bqq_1 + Cq^2q_1^2 + \cdots + Zq^{n-2}q_1^{n-2} = 0,$$

where all the coefficients are positive. This can obviously not be satisfied for positive q and q_1.

In the more general case of (6.17) the equation is $y_1 = cy_2$, where $c = \cos^2 \beta/\cos^2 \alpha$, and it is readily verified that the above argument holds for $c > 1$. If $c < 1$, the curve for cy_2 is below that for y_1 at $q = 0$ and at $q = \infty$, so that if a real root exists at all there will be at least two roots. These roots cannot coincide, since if they did we should have $y_1 = cy_2$ and $y_1' = cy_2'$ simultaneously, which is ruled out by the above argument. The same argument also shows that there cannot be more than two roots. That there are at least two follows from the fact that when $y_1 = \sec^2 \beta$, say at $x = b$, $cy_2 = \sec^2 \beta \sec^2 \alpha$, so that $y_1/cy_2 = 1/\sec^2 \alpha < 1$. The curve for y_1 is therefore below that of cy_2 at $x = b$.

For the modified exponential curve the expression for $\cos^2 \alpha$ corresponding to (6.15) is

$$(6.21) \qquad \cos^2 \alpha = \left\{\frac{1 - q^n q_1^n}{1 - qq_1} - \frac{(1 - q_1^n)(1 - q^n)}{n(1 - q_1)(1 - q)}\right\}^2 \Big/ ff_1,$$

where

$$f = \frac{1 - q^{2n}}{1 - q^2} - \frac{1}{n}\left(\frac{1 - q^n}{1 - q}\right)^2$$

and f_1 is the same expression with q_1 instead of q.

When $q_1 = 1$,

$$(6.22) \qquad \cos^2 \alpha = \frac{3}{n} \cdot \frac{n - 1}{n + 1}\left(\frac{1 - q^n}{1 - q}\right)^2 \Big/ f,$$

which may be written

$$\frac{1+q^n}{1+q} = \frac{1-q^n}{1-q}\frac{1}{n}\{1 + 3(n-1)\sec^2\alpha/(n+1)\}$$

or

$$(1+q)(1+q+q^2+\cdots+q^{n-1})\left(\frac{1}{n} + \frac{3(n-1)\sec^2\alpha}{n(n+1)}\right) - (1+q^n) = 0.$$

Since in this equation there are only two changes of sign, the factor $\frac{1}{n} + \frac{3(n-1)\sec^2\alpha}{n(n+1)}$ being less than 1 for admissible values of α, there cannot be more than two real positive roots. Hence there is no nonlocal overlapping near the middle of the tube.

When $q_1 = 0$,

$$\cos^2\alpha = \frac{n}{n-1}\left\{1 - \frac{1-q^n}{n(1-q)}\right\}^2 \Big/ f. \tag{6.23}$$

It may be shown that the derivative of the right-hand side, considered as a function of q, is negative for all values of $q > 0$. Since the right-hand side is equal to 1 for $q = 0$ and to $(n-1)^{-2}$ for $q = \infty$, there is just one real positive root for any admissible value of $\cos^2\alpha$. Therefore no nonlocal overlapping is possible at the ends of the tube.

Moreover, it is readily shown that $\cos^2\alpha$ in (6.22) has no maximum or minimum for any value of q except 0 or 1. That is, the geodesic distance from the midpoint of the tube to a variable point P of the axis increases monotonically as P moves away towards either end. The same conclusion follows from (6.23) as P moves away from the end of the tube towards the middle. This circumstance, which of course holds also for the simple exponential tube, suggests that the possibility of nonlocal overlapping is effectively ruled out.

7. Probability formulas and tables. The "volume" of a tube of geodesic radius θ surrounding the projected curve will be given by

$$l_n \pi^{\frac{1}{2}(n-2)} \sin^{n-2}\theta/\Gamma(\tfrac{1}{2}n), \tag{7.1}$$

where l_n is a function of n evaluated in Section 5. For any value of $n > 2$ there will also be hemispherical "caps" at the ends of the curve. The "volume" of a complete cap of radius θ surrounding a given point on the hypersphere is

$$\begin{aligned} &\frac{2\pi^{\frac{1}{2}(n-2)}}{\Gamma[\frac{1}{2}(n-2)]}\int_0^{2\pi\sin(\theta/2)}\int_0^{\theta/2}\cos\varphi\sin^{n-3}\varphi\,d\varphi\,ds \\ &\qquad = 2\pi^{\frac{1}{2}n}\sin^{n-1}(\theta/2)/\Gamma(\tfrac{1}{2}n), \end{aligned} \tag{7.2}$$

and we may consider this as the sum of the two hemispherical caps at the ends.

The probability that a random sample point will lie within the tube is therefore given by

$$(7.3) \qquad p(\theta) = \frac{l_n}{2\pi} \sin^{n-2} \theta + \sin^{n-1} (\theta/2).$$

In terms of the correlation coefficient R, the probability of obtaining by chance a value of R at least as great as R_0 is

$$(7.4) \qquad p(R_0) = \frac{l_n}{2\pi} (1 - R_0^2)^{\frac{1}{2}(n-2)} + \{\tfrac{1}{2}(1 - R_0)\}^{\frac{1}{2}(n-1)}.$$

This, of course, is true only for values of $R_0 > 0.894$. It will often happen, however, that when the data suggest an exponential trend the correlation between observed and fitted values will be high.

The value of R_0 corresponding to an assigned significance level can be calculated from equation (7.4) for particular values of n. A few such values are given in Table III, where it will be noted that in each column the last entry is below the critical value.

TABLE III

Values of correlation coefficient corresponding to certain significance levels ($Y = be^{px}$)

n	Significance level			
	.05	.01	.001	.0001
3	.990	.9995	>.9999	>.9999
4	.938	.987	.999	.9999
5	.876	.958	.991	.9996
6		.923	.976	.992
7		.887	.956	.983
8			.935	.970
9			.912	.955
10			.889	.939
12				.906
15				.859

For the modified exponential, the probability of obtaining by chance a value of the correlation coefficient R between observed and computed values of y at least as great as R_0, on the null hypothesis that $b = 0$ in the parent population, is given by

$$(7.5) \qquad p(R_0) = \frac{l_n'}{2\pi} (1 - R_0^2)^{\frac{1}{2}(n-3)} + \{\tfrac{1}{2}(1 - R_0)\}^{\frac{1}{2}(n-2)},$$

for $R_0 > 0.894$. This is the same formula as (7.4) with $n - 1$ instead of n, because of the loss of one dimension in projection. Also, l_n' is now given by (5.12), instead of (5.7). A short table, computed on the basis of this formula and assuming that no overlapping exists, is appended as Table IV.

TABLE IV

Correlation coefficient corresponding to certain significance levels ($Y = a + be^{px}$)

n	Significance level .05	.01	.001	.0001
4	.983	.999	.9999	.9999
5	.918	.983	.998	.9998
6	.849	.949	.989	.998
7		.910	.972	.991
8		.872	.951	.981
9			.904	.952
10			.881	.935
11				.918
12				.901

Note that for Table III the correlation coefficient is computed *without* elimination of the means, whereas for Table IV the correlation coefficient is computed in the usual way.

8. Methods of fitting curves. As n increases, smaller values of R become significant at any given level, and for moderately large n these values of R, at the lower levels, pass out of the region for which the calculation is valid. However, the table may be useful in deciding whether an assumed exponential regression is plausible, when the number of sample points available is fairly small.

Tables for use in fitting exponential curves may be found in Glover's *Tables* [3] but unfortunately these tables cover a very limited range of values of p (from 0 to 0.0953, e^p between 1.0 and 1.1). The exact solution by least squares is laborious. Values of b and p are calculated from the equations

$$\begin{aligned} \sum_{\alpha} ye^{p\alpha} &= b\Sigma e^{2p\alpha}, \\ \Sigma \alpha ye^{p\alpha} &= b\Sigma \alpha e^{2p\alpha}. \end{aligned} \tag{8.1}$$

Rough approximations to b and p may be found by fitting a straight line to the values of log y, and these approximations may be improved by Seidel's method.

Villars [4] has recently given some approximate methods of fitting the modified exponential equation $y = a + be^{px}$. In the first method the n observations (n even) are divided into two groups, one including the 1st, 3rd, 5th, etc., and the other the 2nd, 4th, 6th, etc. The relationship between the expectations of corresponding members of the two groups is

$$\nu_j = h + m\mu_j, \qquad j = 1, 2, \cdots, n/2, \tag{8.2}$$

where

$$\mu_j = a + be^{(2j-1)p}, \tag{8.3}$$

$$\nu_j = \mu + be^{2jp},$$

and so $m = e^p$, $h = a(1 - e^p)$. Hence if a straight line through the origin is fitted to the observed u and v values, where $u_j = y_{2j-1}$, $v_j = y_{2j}$, both variables being subject to error, the slope of this line will give an estimate of e^p and its intercept will give an estimate of $a(1 - e^p)$. An estimate of b can then be found from (8.5), or alternatively both a and b can be found from (8.4) and (8.5), where

$$Na + b\,\Sigma m^\alpha = \Sigma y, \tag{8.4}$$

$$a\,\Sigma m^\alpha + b\,\Sigma m^{2\alpha} = \Sigma y m^\alpha. \tag{8.5}$$

An alternative method, also given by Villars, is applicable whether n is odd or even. It consists in treating y_j and y_{j+1}, $j = 1, 2, \cdots, n - 1$, as pairs of corresponding u and v values, although, since each u except the last appears also as a v, the pairs are clearly not independent.

A systematic method of calculating the exact least squares solution, starting with an approximate value of p (or of $q = e^p$), has been presented by W. J. Spillman [7]. This method utilises tables of q^x for x between 2 and 20 and for q at intervals of .01 between 0 and 1.

There is, of course, no point in fitting an exponential equation by least squares, or by any more or less equivalent method, unless there is some reason to believe that the underlying assumption of approximate uniformity in the variance of y is valid. If log y is approximately normal with constant standard deviation, as seems to be true for many data in the field of economics, the usual procedure of fitting a straight line to the logarithms of the values of y is clearly justified. On the other hand, if the standard deviation of y is constant, the effect of this procedure is to give undue weight to the smaller values of y. Some data exist on fertilizer trials in which the assumption of constant standard deviation of y seems reasonable, and which suggest an exponential, or modified exponential, trend.

9. Numerical illustrations. The data in Table V, referring to the mean girths y of rubber trees in inches after various levels x of fertilizer treatment, I owe to the courtesy of Mr. H. Fairfield Smith.

TABLE V

x	y	Y
0	20.518	20.526
1	21.138	21.109
3	21.734	21.804
5	22.218	22.144
7	22.286	22.311

It will be observed that the values of x are not equidistant. This makes the calculations more awkward.

If the fitted equation is $Y = a + bq^x$, where $q = e^p$, the least squares equations for a, b, q are

$$
\begin{aligned}
&5a + b[1 + q + q^3 + q^5 + q^7] = y_0 + y_1 + y_3 + y_5 + y_7, \\
&a[1 + q + q^3 + q^5 + q^7] + b[1 + q^2 + q^6 + q^{10} + q^{14}] \\
&\qquad = y_0 + y_1q + y_3q^3 + y_5q^5 + y_7q^7, \qquad (9.1)\\
&a[1 + 3q^2 + 5q^4 + 7q^6] + b[q + 3q^5 + 5q^9 + 7q^{13}] \\
&\qquad = y_1 + 3y_3q^2 + 5y_5q^4 + 7y_7q^6.
\end{aligned}
$$

Approximate values for a, b, and q were found by Cowden's method [6]. A curve was drawn by eye between the plotted points and a trial value of a calculated from three equidistant ordinates, Y_0, Y_1, Y_2, by the formula

$$a = \frac{Y_0 Y_2 - Y_1^2}{Y_0 + Y_2 - 2Y_1}. \qquad (9.2)$$

Values of $Y - a$ were then plotted on semi-logarithmic paper, and the value of a was adjusted by trial so that a straight line fitted the points reasonably well. From this straight line, values of b and q were obtained, b being the ordinate at $x = 0$, and q^7 the ratio of the ordinates at $x = 7$ and $x = 0$. In this way the following approximate values were calculated:

$$a_0 = 22.5, \qquad b_0 = -2.0, \qquad q_0 = 0.70.$$

One application of Seidel's method (solving linear equations in δa, δb, δc) gave the improved values

$$a = 22.47, \qquad b = -1.945, \qquad q = 0.7000.$$

Using these values, the calculated Y of Table V were obtained. The correlation coefficient between observed y and computed Y is 0.99735, which corresponds to $n = 5$. If, ignoring the slight deviation from uniformity in the x-intervals, we use Table IV, we find that P is slightly >0.001. The null hypothesis of no effect of the fertilizer is decisively rejected.

Villars [4] gave an illustration of the fitting of an exponential curve to data referring to a certain property of a rubber latex. By his first method he obtained as the regression equation

$$Y = 1.0009 - 0.2877e^{-0.2236x}, \qquad (9.3)$$

where $2x = t + 1$ of his formula (4.1). By the second method he obtained the equivalent of

$$Y = 1.0000 - 0.2811e^{-0.2218x}. \qquad (9.4)$$

If the correlation coefficients between Y and y are computed for both equations, the values turn out to be 0.9560 and 0.9572, respectively, so that the second method gives a slightly better fit. The number of observations, however, is large enough (sixteen) for such a coefficient of correlation to be very highly significant, with reference to the null hypothesis.

For the purposes of illustration, we will use only the first six of Villars' observations, given in the first two columns of Table VI.

TABLE VI

x	y	Y (method 1)	Y (method 2)
1	0.776	0.7715	0.7742
2	0.852	0.8555	0.8415
3	0.850	0.8826	0.8754
4	0.869	0.8914	0.8924
5	0.939	0.8942	0.9010
6	0.904	0.8951	0.9053

By Villars' first method the values of Y given in column 3 were calculated, corresponding to the equation $Y = 0.8955 - 0.3851e^{-1.1329x}$. The correlation coefficient between y and Y is $R = 0.871$, so that from Table IV the departure from the null hypothesis is barely significant. Villars' second method gives the values in column 4, corresponding to $Y = 0.9088 - 0.2693e^{-0.6870x}$, and in this case $R = 0.906$. The fit is therefore appreciably better, and the regression appears to be significant at a level about midway between .01 and .05.

10. Acknowledgement. In conclusion, the author would like to express his great indebtedness to Professor Harold Hotelling for having suggested the problem and for helpful advice and criticism.

REFERENCES

[1] HAROLD HOTELLING, "Tubes and spheres in n-spaces and a class of statistical problems," *Am. Jour. of Math.*, Vol. 61 (1939), p. 440.

[2] D. M. STARKEY, "The distribution of the multiple correlation coefficient in periodogram analysis," *Annals of Math. Stat.*, Vol. 10 (1939), p. 327.

[3] J. W. GLOVER, *Tables of Applied Mathematics*, George Wahr, Ann Arbor, Mich., 1923, pp. 468–481.

[4] D. S. VILLARS, "A significance test and estimation in the case of exponential regression," *Annals of Math. Stat.*, Vol. 18 (1947), p. 596.

[5] D. S. VILLARS AND T. W. ANDERSON, "Significance tests for normal bivariate distributions," *Annals of Math. Stat.*, Vol. 14 (1943), p. 141.

[6] D. J. COWDEN, "Simplified methods of fitting certain types of growth curves," *Jour. Am. Stat. Assn.*, Vol. 42 (1947), p. 585.

[7] W. J. SPILLMAN, *Use of the Exponential Yield Curve in Fertilizer Experiments* (Technical Bulletin No. 348), U. S. Department of Agriculture, Washington, D. C.