

ON FIDUCIAL INFERENCE¹

BY D. A. S. FRASER

*University of Toronto*²

1. Introduction. The subject of fiducial probability was introduced thirty years ago by R. A. Fisher. In the original paper [8] entitled "Inverse probability" Fisher discussed the importance of the maximum likelihood method and then produced a *fiducial distribution* for a parameter in roughly the following manner. Let T be a maximum likelihood estimate of a parameter θ . The distribution function for T given θ , $F(T | \theta)$, has a uniform distribution on the interval $[0, 1]$. Differentiating partially with respect to T gives the probability density function for T given θ :

$$\left| \frac{\partial}{\partial T} F(T | \theta) \right|$$

Differentiating partially with respect to θ gives a function treated as a density function for "the fiducial distribution of a parameter θ for a given statistic T ." From this density function, "fiducial limits" for the parameter θ given T can be calculated.

As an illustration Fisher treated the correlation coefficient r for sampling from a normal bivariate population having correlation ρ . The supporting interpretation for the fiducial method in this example seems to me very much like a present-day confidence argument. This, I gather, led Professor Neyman in 1934 [14] to present his theory of confidence intervals as an *extension* of the fiducial method. Both Fisher and Neyman have since emphasized that the theories are different and the recent literature stands in testimony to the large separation now existing between them.

Today I shall review some of the problems that have been analyzed by the fiducial method and discuss briefly some of the results obtained for these problems; also, I shall put forward a mathematical framework³ within which I feel fiducial probability has a clear frequency interpretation for a large class of problems. A natural beginning is Fisher's [2] statement: "By contrast, the fiducial argument uses the observations only to change the logical status of the parameter from one in which nothing is known of it, and no probability statement about it can be made, to the status of a random variable having a well defined distribution." Such statements have perturbed many mathematical statisticians.

Received November 18, 1960; revised March 16, 1961.

¹ An address presented on August 24, 1960, under the title "Fiducial probability," at the Annual Meeting of the Institute of Mathematical Statistics in Stanford, California. The address was prepared on the invitation of the IMS Committee on Special Invited Papers.

² Present address: Bell Telephone Laboratories, Murray Hill, N. J.

³ The development and the proof will be found in [10].

2. A simple example yielding a frequency interpretation. Consider a sample of n observations from a normal distribution with unknown mean μ and known variance. A sufficient statistic is the sample mean \bar{x} . For simplicity suppose the known variance to be such that \bar{x} is normally distributed with mean μ and variance 1. The information concerning frequency distribution can be described in several ways. First, it can be described by means of the frequency distribution of \bar{x} ; see Fig. 1. Second, it can be described in terms of the frequency distribution of the error $g = \bar{x} - \mu$; see Fig. 2. Borrowing from notions in the theory of measurement we might then say that μ was "measurable" in the sense that \bar{x} which can be observed is *in error*, or fluctuates, in known frequency form about μ . As a third way consider the following. In the light of the available information concerning frequency distribution, let μ^* designate possible values for the parameter relative to an observed \bar{x} ; see Fig. 3. The statistical problem admits free translation on the \bar{x} axis. Consider a very large number of samples from normal distributions within the specifications of this example. In each case translate the sample mean to the value in Fig. 3. The parameter values will be correspondingly translated. Simple mathematics then shows that the frequency distribution of these translated means μ^* is normal with center at \bar{x} and with unit scale parameter. There is thus a frequency distribution of parameter values μ^* that might have produced the observed \bar{x} . A probability statement of the form

$$\Pr \{ \bar{x} - 1.96 < \mu^* < \bar{x} + 1.96 \} = 95 \%$$

can then be made with \bar{x} fixed and with μ^* treated as a variable designating possible values for the parameter. The asterisk on the μ can even be omitted provided we keep the interpretation just given. Translation of any kind of repeated sampling (n fixed) so that sample means are moved to the observed \bar{x} will yield fixed frequency (95% in the case just given) for the event, the population means falling in any prescribed interval about \bar{x} . The process leading to the above probability statement might be compared with a gambling game in which the dice are rolled but concealed from view; bets made; then the dice exposed. With honesty, or with perfect concealment, such a game would be equivalent to one in which the bets are made before the dice were rolled.

In this example, the freedom of translation has, I feel, produced a precise frequency interpretation for fiducial probability.

3. A generalization of the simple example. Let \mathfrak{X} be a sample space for a sufficient statistic x ; let G be a group of transformations on \mathfrak{X} with typical element h ; let Ω be a parameter space with parameter θ . Suppose that the spaces \mathfrak{X} , G , Ω are identical and hence are groups, and that the distribution of x for any parameter value θ is obtained by using θ as a transformation (left group multiplication) applied to a variable g with a *fixed* distribution; g will be referred to as the *error variable* and its distribution as the *error distribution*. With this formulation the specification, i.e., the family of possible distributions, is invariant under each transformation belonging to the group G , indeed $(hx) = (h\theta)g$ for h in G shows

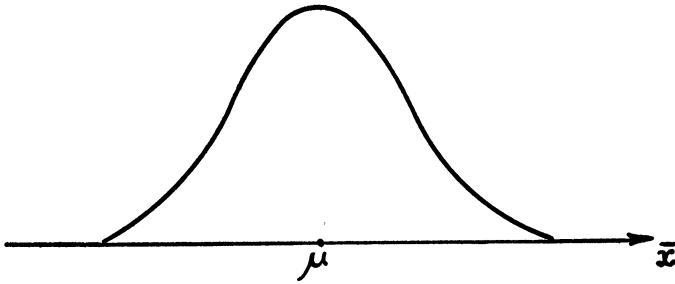


FIGURE 1

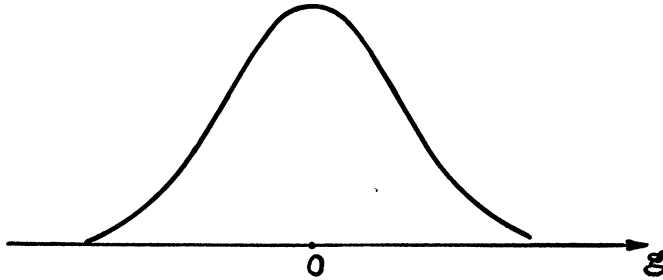


FIGURE 2

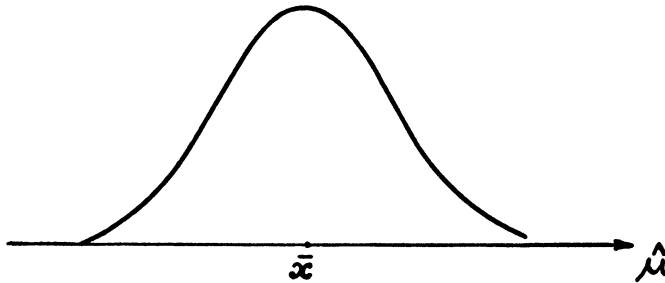


FIGURE 3

that the variable hx has parameter $h\theta$ when x has parameter θ . Further details on this generalization may be found in [10].

The equation $x = \theta g$ can be simply manipulated to produce

$$g = \theta^{-1}x.$$

Since g has a fixed frequency distribution this equation shows that $\theta^{-1}x$ is a *pivotal quantity* (a function of the parameter and the sufficient statistic that has a fixed frequency distribution). A simple analysis [10] then shows that $\theta^{-1}x$ is

essentially unique among invariant pivotal quantities. The frequency information in the specification of the problem can then be compactly formulated in terms of the fixed distribution for the *error variable* g . From this latter formulation the distribution of x given θ can be obtained from the equation $x = \theta g$ and the distribution of θ given x from the equation $\theta = xg^{-1}$. A simple interpretation of this second equation is in terms of possible values $\hat{\theta}$ for the parameter: let x be an observed value from the sufficient statistic; consider the sufficient statistic with distribution determined by some parameter value; use a variable transformation to transform the variable sufficient statistic into the fixed value x ; the parameter value will correspondingly be transformed into a frequency distribution—the *fiducial distribution* given by equation $\theta = xg^{-1}$.

Suppose now that the distributions can be described by means of density functions. The natural “carrying” measure on a group is the invariant or Haar measure.⁴ Let μ designate the *left Haar measure* on G ; it has the property: $\mu(hH) = \mu(H)$ for all h in G and all $H \subset G$. A related measure is *right Haar measure*; designate it by ν ; it has the property $\nu(Hh) = \nu(H)$ for all h in G and all $H \subset G$. The *modular function* $\Delta(h)$, defined by $\mu(Hh) = \Delta(h)\mu(H)$ (which holds for all H), gives the relationship between left and right Haar measures: $d\mu = \Delta d\nu$.

Let $p(g)$ be the probability density function for the error variable g with respect to left Haar measure; accordingly, the probability element for g is

$$(1) \quad p(g) d\mu(g).$$

Simple analysis [10] then shows that the probability element for x given θ is

$$(2) \quad p(\theta^{-1}x) d\mu(x),$$

and for θ given x is

$$(3) \quad p(\theta^{-1}x)\Delta(x) d\nu(\theta).$$

This last expression describes a standard conditional distribution for θ given x within the mathematical model just presented. The freedom to make transformations can thus be used to remove the rigidity that results when seemingly undue emphasis is placed on the “origin” in the coordinate system.

4. Sampling from a normal distribution. Consider a sample of size n from a normal distribution with unknown mean and variance. The sample mean and standard deviation (\bar{x}, s) form a sufficient statistic for the parameter (μ, σ) , the mean and standard deviation of the normal distribution. For most physical problems from which this statistical problem might have been abstracted, the origin and the unit for measurement are arbitrary or conventional; a natural group of transformations then is that involving location and scale changes. The multiplication of this group then yields the representation

⁴ See, for example, Chap. XI in Halmos, [11].

$$(\bar{x}, s) = (\mu, \sigma)g = (\mu, \sigma) \left(\frac{z}{n^{\frac{1}{2}}}, \frac{\chi}{(n-1)^{\frac{1}{2}}} \right),$$

which expresses the distribution of (\bar{x}, s) in terms of relocation (μ) and rescaling (σ) of the distribution in the standardized case; here z and χ are independent variables, standardized normal and chi on $n - 1$ degrees of freedom respectively. The formula for group multiplication is $(a, b)(c, d) = (a + bc, bd)$. By working from this fixed *error* distribution for g it is straightforward to calculate the conditional distribution of (\bar{x}, s) given (μ, σ) , or the conditional distribution of (μ, σ) given (\bar{x}, s) , the latter being the fiducial distribution. These are ordinary conditional distributions resulting from the freedom of location and scale.

A frequency interpretation for the fiducial distribution of μ given (\bar{x}, s) can proceed as follows: let (\bar{x}, s) designate the observed values; consider a long sequence of parameter-observation combinations in which the parameter may stay constant or may vary; in each case relocate and rescale to bring the sample mean to the fixed \bar{x} and the sample standard deviation to the fixed s ; these transformations then carry the parameter values into a frequency distribution which simple analysis shows to be Student's distribution located at \bar{x} and scaled by $s/n^{\frac{1}{2}}$. This is the fiducial distribution of μ given (\bar{x}, s) .

The general expressions for probability elements in the preceding section can easily be specialized. Left Haar measure under the location-scale group has measure element $d\bar{x} ds/s^2$ and right Haar measure has element $d\bar{x} ds/s$; the modular function is $\Delta(\bar{x}, s) = 1/s$. The probability element for (\bar{x}, s) given (μ, σ) has the form

$$(4) \quad \frac{n^{\frac{1}{2}}}{(2\pi)^{\frac{1}{2}}\sigma} \exp \left[-\frac{n}{2\sigma^2} (\bar{x} - \mu)^2 \right] \frac{1}{\Gamma(\frac{1}{2}(n-1))} \exp \left[-\frac{(n-1)s^2}{2\sigma^2} \right] \cdot \left(\frac{(n-1)s^2}{2\sigma^2} \right)^{\frac{1}{2}(n-3)} \cdot \frac{2s(n-1)}{2\sigma^2} s^2 \cdot \frac{d\bar{x} ds}{s^2}.$$

(Note the rearrangement so that left Haar measure is used for the measure element.) The probability element for (μ, σ) given (\bar{x}, s) is then easily seen to be

$$(5) \quad \frac{n^{\frac{1}{2}}}{(2\pi)^{\frac{1}{2}}\sigma} \exp \left[-\frac{n}{2\sigma^2} (\bar{x} - \mu)^2 \right] \frac{1}{\Gamma(\frac{1}{2}(n-1))} \exp \left[-\frac{(n-1)s^2}{2\sigma^2} \right] \cdot \left(\frac{(n-1)s^2}{2\sigma^2} \right)^{\frac{1}{2}(n-3)} \cdot \frac{2s(n-1)}{2\sigma^2} s^2 \cdot \frac{1}{s} \cdot \frac{d\mu d\sigma}{\sigma}$$

This probability element for (μ, σ) given (\bar{x}, s) admits a precise frequency interpretation in terms of possible parameter values corresponding to the particular (\bar{x}, s) . Consequently it can be integrated to produce marginal frequency distributions with similar interpretations for "variables" such as $\mu, \sigma, \mu + 1.96\sigma$, always of course, given (\bar{x}, s) . These fiducial distributions correspond to those given by Fisher (e.g., in his *Statistical Methods and Scientific Inference* [2]).

The use of location and scale transformations has led to a mathematical

framework for a problem that Fisher treated in a physical setting. With a few deletions I quote from Fisher [9]: "A complementary doctrine . . . violating . . . the principles of deductive logic is to accept a general symbolical statement such as

$$\Pr \{(\bar{x} - ts) \leq \mu \leq (\bar{x} + ts)\} = \alpha$$

as rigorously demonstrated and yet, when numerical values are available for the statistics \bar{x} and s , so that on substitution of these and use of 5 per cent. value of t , the statement would read

$$\Pr \{92.99 < \mu < 93.01\} = 95 \%,$$

to deny to this *numerical* statement any validity. This evidently is to deny the syllogistic process of making a substitution in the major premise of terms which the minor premise establishes as equivalent. . . . [This is used to support] the assertion that if μ stands for some objective constant of nature, or property of the real world, such as the distance of the sun, its probability of lying between any named numerical limits is necessarily either 0 or 1 and we cannot know which unless the true distance is known to us. The paradox . . . requires that we should wilfully misinterpret the probability statement so as to pretend that the population to which it refers is not defined by our observations and their precision but is absolutely independent of them. As this is certainly not what any astronomer means and is not in accordance with the origin of the statement, it seems rather like an acknowledgment of bankruptcy to pretend that it is."

The framework I have suggested puts emphasis on the error variable " g " and on the freedom from an artificial origin in the coordinate system. In this framework there seems to me to be little basis for criticizing the fiducial method.

5. A further generalization. The model in Section 3 can sometimes be applied *within* more general problems if there is an ancillary statistic. Let (x, a) be an *exhaustive statistic* for estimating the parameter θ ; that is,

(i) The conditional distribution given the statistic (x, a) does not depend on θ .

(ii) No reduction can be made in (x, a) ; that is, no non-trivial function of (x, a) exists that satisfies (i). Suppose now that a is an *ancillary statistic*, a statistic having a fixed distribution regardless of the value of θ . In this situation x can be interpreted as a sufficient statistic for θ *given* the value of the ancillary statistic a . In a sense a describes a situation in which one is caught, and in which x is sufficient for θ .

For some problems of this type it may be possible to apply the model in Section 3 to the conditional problem involving the distribution of x given the value of the ancillary statistic a .

6. The problem of location and scale. In 1938 E. J. G. Pitman [15] gave an extensive treatment of interval estimation for problems of location, of scale, and of location and scale. These problems lend themselves to the methods in the

preceding sections and the fiducial distributions can be obtained very simply. Consider the problem of location and scale; the other problems can be treated similarly.

Let (x_1, \dots, x_n) be a sample of n from the distribution having density function

$$\frac{1}{\sigma} f\left(\frac{x - \mu}{\sigma}\right),$$

where f is a specified function and μ, σ are the parameters of location and scale. The density function for a sample of n is

$$\frac{1}{\sigma^n} \prod_{i=1}^n f\left(\frac{x_i - \mu}{\sigma}\right).$$

To have the transformation properties for the fiducial method, statistics of location and scale are needed; the sample mean and standard deviation are a simple and convenient choice; any other pair would serve, however. The remaining information in the sample (x_1, \dots, x_n) can be described by the relative spacing of the sample values,

$$a = \left(\frac{x_1 - x}{s}, \dots, \frac{x_n - \bar{x}}{s}\right),$$

which is called a *configuration statistic* and describes the "shape" of the sample without reference to its location and its scale. (The elements in the expression for the statistic above satisfy two constraints and as a result any designated pair of elements could be omitted from the expression.) The invariance of the configuration statistic under location and scale changes shows easily that it has a fixed distribution independent of (μ, σ) and hence is ancillary; its distribution will however depend on the form of the density function f . The statistic (\bar{x}, s) is conditionally sufficient. The combination (x, a) here is not exhaustive in the sense that the reduction can be made from the ordered observations (x_1, \dots, x_n) to the unordered observations $\{x_1, \dots, x_n\}$; this has, however, no effect on the argument here.

The approach put forth in Section 5 is to examine the problem conditionally given the ancillary statistic a . For this, the joint density for (\bar{x}, s, a) given (μ, σ) is needed. The Jacobian of the transformation from (x_1, \dots, x_n) to (\bar{x}, s, a) where two of the elements of a are omitted turns out to depend on (\bar{x}, s) through a factor s^{n-2} . From this, it follows that the conditional probability element for (\bar{x}, s) given (a, μ, σ) has the form

$$(6) \quad \frac{c}{\sigma^n} \prod_{i=1}^n f\left(\frac{x_i - u}{\sigma}\right) s^{n-2} \cdot s^2 \cdot \frac{d\bar{x} ds}{s^2};$$

the measure element is arranged to exhibit left Haar measure. The formulas of Section 3 and the Haar measure results in Section 4 then produce the following fiducial element for (μ, σ) given (\bar{x}, s, a) :

$$(7) \quad \frac{c}{\sigma^n} \prod_{i=1}^n f\left(\frac{x_i - \mu}{\sigma}\right) s^{n-2} \cdot s^2 \cdot \frac{1}{s} \frac{d\mu}{\sigma} \frac{d\sigma}{\sigma}.$$

The frequency interpretation of this fiducial element is in terms of possible values for (μ, σ) relative to (\bar{x}, s) , all given the value of the ancillary statistic a .

7. The relationship of fiducial distributions and prior distributions.⁵ Consider a statistical problem of the form introduced in Sections 3 or 5, and suppose that there is an *a priori* distribution for the parameter θ having a density function with respect to Haar measure; let the *a priori* probability element be

$$n(\theta) d\nu(\theta)$$

with respect to right Haar measure. The analysis of Sections 3 or 5, which ignores prior distributions, produces a fiducial probability element

$$kp(\hat{\theta}^{-1}x) d\nu(\hat{\theta})$$

for the parameter given the sufficient or conditionally sufficient statistic x . The prior distribution can be combined with the observational results in two ways:

By a Bayes argument. The formal joint probability element for x and θ is

$$n(\theta) d\nu(\theta)p(\theta^{-1}x) d\mu(x)$$

which yields the following *a posteriori* element for θ given x :

$$(8) \quad kn(\theta)p(\theta^{-1}x) d\nu(\theta).$$

By a joint distribution argument. Consider the joint distribution of the prior θ and the possible values $\hat{\theta}$ given the observation x . Introduce the condition $\theta = \hat{\theta}$ with respect to right Haar measure. The resulting conditional distribution for θ is just that produced by the Bayes argument in the preceding paragraph.

Thus within the transformation framework of Sections 3 and 5 the information about θ extracted from the observations by means of the fiducial argument can be combined in a logical manner with prior information and be entirely consistent with the Bayes approach.

Some other results on the relationship of fiducial probability and prior probability are the following. Lindley [12] has proved for a real-valued parameter that the fiducial distribution is a Bayes posterior distribution if and only if the parameter is essentially a location parameter. D. R. Brillinger in unpublished research at Princeton University has proved that the fiducial distribution is a Bayes posterior distribution if the transformations on the sample space form an r -dimensional Lie group and the sample and parameter spaces are r -dimensional manifolds. In the framework of Sections 3 and 5, setting $n(\theta) \equiv c$ in the formula for the Bayes posterior distribution (8) gives the formula for the fiducial probability element. Thus the fiducial distribution is a Bayes posterior distribution if

⁵ For the details of the analysis in this section, see [10].

the sample and parameter spaces are identical to the transformation group and the distributions are given by means of density functions with respect to right Haar measure.

8. Combining fiducial distributions from separate systems.⁶ Suppose there are two systems concerned with a parameter θ and suppose that each is of the form introduced in Sections 3 or 5. Let x be a sufficient statistic, and $p(g)$ be the density function for the random "error" in a first system, and let y , $q(h)$ be the corresponding ingredients of the second system. Consider the following three methods of combining the information from these two systems.

By direct combination of fiducial distributions. The first system produces the fiducial element $p(\theta_1^{-1}x) d\nu(\theta_1)$; the second system the fiducial element $q(\theta_2^{-1}y) d\nu(\theta_2)$. A reasonable method of combining these is by reference to the condition $\theta_1 = \theta_2$ re right Haar measure. The resultant fiducial element for θ is

$$kp(\theta^{-1}x)q(\theta^{-1}y) d\nu(\theta)$$

By a Bayes argument. If the fiducial distribution from the first system is used as a prior distribution for the second system, the resulting posterior distribution is just that obtained by the first method of analysis.

By an overall fiducial argument. If the combined observational system admits a sufficient statistic then the combined system has the form in Sections 3 and 5 and yields a fiducial distribution. A modest amount of analysis shows that it is just the distribution obtained by the first two methods.

If the combined system does not have a sufficient statistic it at least has a conditionally sufficient statistic given an ancillary statistic, since the orbits in the product space for (x, y) generated by the transformations in the group may be taken as values of an ancillary statistic and the positions on an orbit may be taken as the values of the conditionally sufficient statistic. The fiducial distribution then obtained by the method suggested in Section 5 is just that obtained by the other methods.

In the Pitman location-and-scale problem consider separate groups of observations and treat them as separate systems. The results described above then show that the fiducial distributions from the separate groups can be combined by either of the first two methods yielding the overall fiducial distribution based on the ancillary statistic. Thus the ancillary statistic itself is generated by a general fiducial argument.

9. Fiducial probability and statistical inference. The results concerning frequency interpretation and freedom of handling of fiducial probabilities have, I feel, some general implications for statistical inference. A formal statistical problem contains certain basic information concerning possible frequency functions for the variables being observed. This has been called the *specification* by Fisher. If the specification has the transformation properties of Sections 3 or 5, then the

⁶ For the details of the analysis in this section, see [10].

fiducial argument gives a frequency distribution of possible parameter values given the data. In this framework then fiducial probability gives an answer to the question D. R. Cox [5] in his 1958 paper felt that statistical inference should answer: "What do the data tell us about θ ?"

In addition, many statistical problems pose some question concerning the parameter or concerning some function of the parameter. The interpretation of fiducial probability in this paper suggests to me that the answering of such questions can be and perhaps should be a separate part of the analysis which would use from the observations only the fiducial distribution; in fact, accepting the fiducial distribution as a frequency distribution of possible values for the parameter allows the answering of questions concerning the parameter even in cases where the transformation properties do not hold for these questions. Some support for this suggestion may be found in the result in the preceding section, that the fiducial distribution can be combined in a logical manner with the prior distribution and still be consistent with Bayes. A statistical inference would then involve combining with logic and judgment the fiducial information and any other information—of frequency form, or of a restriction on the parameter range, or of personal probability form. Marginal distributions of parameters, even if related to fixed points in the sample space, would have a frequency interpretation.

10. The Behrens-Fisher problem. The specification of the Behrens-Fisher problem describes two samples from normal populations with unknown parameters. The problem is to estimate or make tests on the difference in population means. For a first system let (\bar{x}_1, s_1) be the sufficient statistic for (μ_1, σ_1) based on a sample size n_1 and for the second system (\bar{x}_2, s_2) for (μ_2, σ_2) based on a sample size n_2 .

The specification for each system admits scale and location changes and these would often be reasonable for the problem in a physical setting. For the first system repeated sampling from the random error variable yields the following fiducial distribution for μ_1 :

$$\mu_1 = \bar{x}_1 + t_1 s_1 / \sqrt{n_1}$$

where t_1 is the variable and it has Student's distribution with $n_1 - 1$ degrees of freedom. Similarly for the second system repeated sampling from the random error variable yields

$$\mu_2 = \bar{x}_2 + t_2 s_2 / \sqrt{n_2}$$

where t_2 is the variable and it has Student's distribution with $n_2 - 1$ degrees of freedom. These distributions together provide a distribution for (μ_1, μ_2) and it has a frequency interpretation as derived from transformation freedom.

From the joint distribution for (μ_1, μ_2) the marginal distribution of $\mu - \mu_2$ can be derived in a straightforward manner; percentage points for this marginal distribution can be obtained from Sukhatme's tables.

In reply to criticisms concerning the error rate of a test based on this distribution Fisher has stated "it is irrelevant to the purposes of the test for the experimenter is not concerned with repeated sampling from the same population." In the transformation framework proposed here the fiducial distribution is generated by repeated sampling from the error variables and has a frequency interpretation given the x 's and s 's.

11. Confidence intervals versus fiducial intervals; an example. In 1939 B. L. Welch [18] considered a simple statistical problem and examined some properties of a best confidence region and a fiducial interval obtained from the use of an ancillary statistic. Let (x_1, x_2) be a sample of two from the uniform distribution on the interval $\theta \pm \frac{1}{2}$. The statistic $z_2 = (x_2 - x_1)/2$ is easily seen to have a fixed distribution and hence is ancillary. The statistic $z_1 = (x_2 + x_1)/2$ is then conditionally sufficient for θ .

A fiducial interval with probability 95% is $z_1 \pm 95\%(0.5 - |z_2|)$; it embodies 95% of the permissible range for θ . A best unbiased 95% confidence interval is

$$z_1 \pm \min \{ [0.5 - |z_2|], [0.5 + |z_2| - (0.5 - 0.95/2)^{\frac{1}{2}}] \}.$$

The confidence interval embodies the full permissible range for θ when that range $(1 - 2|z_2|)$ is small.

The fiducial interval is also a confidence interval and consequently from the confidence point of view falls short of being optimum. Welch plots for each interval the probability that it covers a value Δ away from θ . Over most of the range of Δ the confidence interval has a smaller probability, and nowhere larger.

There is however another side to the comparison. When the permissible range for θ is small the confidence interval embraces not 95% but the *full* range of possible values for θ . It is not hard to see what is happening: when the range of permissible values is short the confidence interval takes the *full* range on the grounds that in probability there may be another occasion when the range of permissible values is larger and *less* than 95% can be chosen and still maintain the long run 95% average. Is the long run average more important than the specialized knowledge of the particular situation?

The comparison between the intervals thus involves a conflict between a general principle and specific knowledge for data in hand. In his 1958 paper Cox emphasized the second of these. Recent papers by Buehler [4] and Wallace [17] are concerned with this sort of conflict and use notions of relevant subsets and strong exactness. There are certainly good grounds, I feel, for ignoring a general principle when it is in conflict with data-specific knowledge. My preference weighs heavily in favour of the fiducial interval for Welch's example.

D. R. Brillinger, in unpublished research at Princeton University, has proved that a fiducial region chosen to be invariant in the transformation framework is also a confidence region. An advantage of the frequency interpretation in terms of an error variable is that it is applicable when the fiducial region has its form based on the variable being observed, the case not covered by Brillinger's result.

12. The Creasy-Fieller problem. The problem of finding an interval estimate for the ratio of the means of two normal distributions was examined by Miss M. A. Creasy [6] and E. C. Fieller [7] in papers presented to the Royal Statistical Society in 1954. The papers drew attention to differences between the fiducial and the confidence methods and provoked interesting discussion [6] [7] bearing on these differences.

The specification in its simplest form describes independent variables x and y that are normally distributed with means μ and ν and with known variance. The problem is to focus on the parameter $\alpha = \mu/\nu$ and find some sort of interval estimate.

Fieller used the pivotal quantity

$$\frac{y - \alpha x}{(1 + 2\alpha + \alpha^2)^{\frac{1}{2}}}$$

and from it obtained regions for the parameter α by manipulations standard to the confidence method. The regions are certainly confidence regions; Fieller, however, used the term fiducial for them, but there seems to be some doubt whether the method of derivation fits the fiducial pattern implicit in Fisher's work.

Miss Creasy proceeded differently. She combined the separate fiducial distributions for μ and ν to obtain a joint fiducial distribution and then integrated to obtain a marginal fiducial distribution for $\alpha = \mu/\nu$. Fiducial limits were then calculated from this distribution.

The pivotal quantity used by Fieller is not compatible with the location transformations natural to the variables x and y . His regions have a confidence frequency interpretation but they do not seem to me to have a frequency interpretation *given* x and y .

From the transformation point of view Miss Creasy's procedure extracts what the data have to say in a frequency sense about the parameters μ and ν . This frequency information is then taken at its face value and used to obtain the marginal distribution of the parameter of interest. The fiducial distribution has a frequency interpretation in terms of error variables and its derivation seems to fit the pattern implicit in Fisher's work.

In this simplest form of the problem, an interval obtained by Miss Creasy's method is always contained in the interval at the same frequency level obtained by Fieller's method. As a result, if the intervals are evaluated from the point of view of repeated sampling from the same population, then the probability with which the Creasy interval covers the value of the parameter must be less than the corresponding probability for Fieller's confidence interval and hence less than the Creasy interval's fiducial probability.

The general tenor of the discussion following the Creasy-Fieller papers favoured the approach used by Fieller—it had the backing of confidence theory.

13. Regression analysis. Consider linear regression analysis with normally distributed errors. Let $\alpha, \beta_1, \dots, \beta_p$ designate the regression parameters and σ

the scale parameter of the error distribution. Also let a, b_1, \dots, b_p designate the least-squares fitted regression coefficients and s the standard deviation about regression. The fitted coefficients $a, \mathbf{b} = (b_1, \dots, b_p)$ have a multivariate normal distribution with means $\alpha, \boldsymbol{\beta} = (\beta_1, \dots, \beta_p)$ and covariance matrix of specified form but scaled by σ^2 ; let $n(a, \mathbf{b} \mid \alpha, \boldsymbol{\beta}; \sigma)$ designate its probability density function. The standard deviation s has a χ type of distribution scaled by σ ; let $\chi(s \mid \sigma)$ designate its density function. A natural group of transformations involves location changes for each regression coefficient and a scale change for the regression coefficients and error scale-parameter. The probability element for a, \mathbf{b}, s given $\alpha, \boldsymbol{\beta}$, can be written

$$n(a, \mathbf{b} \mid \alpha, \boldsymbol{\beta}, \sigma) \chi(s \mid \sigma) \cdot s^{p+2} \cdot \frac{da \, d\mathbf{b} \, ds}{s^{p+2}}$$

where the measure element is arranged so as to exhibit left Haar measure. Formula 3 in Section 3 then gives the fiducial probability element for $\alpha, \boldsymbol{\beta}, \sigma$ given a, \mathbf{b}, s :

$$n(a, \mathbf{b} \mid \alpha, \boldsymbol{\beta}, \sigma) \chi(s \mid \sigma) \cdot s^{p+2} \cdot \frac{1}{s^{p+1}} \frac{d\alpha \, d\boldsymbol{\beta} \, d\sigma}{\sigma}$$

This distribution for $\alpha, \boldsymbol{\beta}, \sigma$ can be described by the equations

$$\alpha = a + \sigma w_0, \quad \beta_i = b_i + \sigma w_i, \quad \sigma = s \sqrt{f/\chi}$$

in terms of variables w_0, \dots, w_p, χ , the w 's having a multivariate normal distribution with means 0 and with covariance matrix equal to the inverse of the matrix of inner products of the structural vectors corresponding to the parameters $\alpha, \beta_1, \dots, \beta_p$, and χ being distributed as a χ variable with the usual 'error' degrees-of-freedom.

In certain cases of nonlinear regression the space for the mean of the basic variables may be contained in a linear subspace of slightly higher dimension. It then seems reasonable to calculate the fiducial distribution for the regression coefficients in the linear subspace and then condition it according to permissible parameter values. If the sample size is large enough most such problems become approximately linear and the fiducial method can be applied directly.

If the error distribution form is something other than the normal, the least squares estimates will in general not be sufficient. The transformation method can, however, *still be* used but in relation to an ancillary statistic—a value of the ancillary statistic corresponding to an orbit under the transformation group.

14. The correlation coefficient. Consider estimation problems concerned with sampling from a bivariate normal distribution. Let $\bar{x}, \bar{y}, s_{11}, s_{12}, s_{22}$ be the sufficient statistic estimates for the parameters $\mu, \nu, \sigma_{11}, \sigma_{12}, \sigma_{22}$.

In the original paper [8] on fiducial probability Fisher produced a fiducial distribution for the correlation coefficient $\rho = \sigma_{12}/(\sigma_{11}\sigma_{22})^{\frac{1}{2}}$ by the method outlined in Section 1 of this paper. In the discussion following the Creasy and Fieller papers [6] [7] Fisher produced a fiducial distribution for (μ, ν) . Mauldon

[13] obtained several alternative pivotal quantities and hence several alternative fiducial distributions for the parameters. Quenouille [3] in his recent book, *The Fundamentals of Statistical Reasoning*, gives a set of rules for handling fiducial probabilities and applies them to this bivariate problem. There seems to be a general lack of consistency among these results. Various transformation groups can be used on the basic sample space for sampling from the bivariate normal distribution. One group that might be natural in certain situations is formed from the following kinds of transformation: scale and location changes for x ; changes in y that are of linear regression form on x ; changes in scale of deviations from the linear regression. The resulting model is of the form discussed in Section 3 and a fiducial distribution of the five parameters can be derived. From this joint distribution a marginal fiducial distribution for the correlation coefficient can be obtained; it will have a frequency interpretation in terms of repeated sampling from the error variable.

The transformation group just given can be used with the roles of x and y interchanged—that is, with regression of x on y . From this, a marginal fiducial distribution for the correlation coefficient can be obtained and it will in general be different from the fiducial distribution mentioned in the preceding paragraph. This second fiducial distribution also has a frequency interpretation but in terms of a different kind of random error transformation. These two distributions appear symmetrically one with respect to the other and hence must be different from the fiducial distribution in Fisher's original fiducial paper [9]. The only frequency interpretation for the original fiducial distribution that I can see is the frequency interpretation that customarily goes with confidence intervals.

This multiplicity of fiducial distributions for the correlation coefficient perhaps reflects certain inadequacies of the correlation coefficient, and hence of the bivariate normal for the strong statements of the fiducial method. Additional structure in the form of suitable or reasonable transformations on the sample space gives some unity in terms of relative positioning of points and can yield a fiducial distribution with a frequency interpretation that in effect states where the parameter might be relative to the observation. Perhaps this additional information from the physical situation is necessary before we can make the strong statements of fiducial probability.

15. Addendum. Prof. Allan Birnbaum has given a paper at this meeting describing some results of his on the topic "informative inference." Consider the case of two parameter values θ_1 and θ_2 and distributions that admit symmetry of the likelihood ratio.

There is then a unique decomposition into *simple experiments* ordered by "more informative than." These simple experiments are symmetrical. This model lends itself nicely to the transformation argument of fiducial probability. A value for the ancillary statistic corresponds to the particular simple experiment one finds oneself in. Within that simple experiment the permutation group on two elements can be used and a fiducial distribution on the two parameter values

derived. The ratio of fiducial probabilities turns out to be the ratio of the likelihoods. This fiducial distribution has a frequency interpretation given the observation.

This is an example of the use of fiducial probability when the sample space is finite. The method of Sections 3 and 5 is of course valid for finite groups and thus extends, in this direction, the usual usage of fiducial probability.

16. Bibliography. Tukey [16] gives a bibliography of fiducial probability; a more extensive bibliography appeared in the handout for the 1958 Wald lectures given by Tukey at the August 1958 meeting of the Institute of Mathematical Statistics.

17. Acknowledgments. Professor L. J. Savage and Professor J. W. Tukey read a preliminary write-up of this paper and offered extensive and penetrating comments. Some of these I was able to incorporate into the paper; others, due to differences in our viewpoints I preferred not to include. I thank them with deep appreciation for their help and for the time and consideration they gave this paper.

Mr. David Brillinger and Mr. Collin Mallows also read the preliminary write-up and offered helpful comments. I appreciate their help, and thank them for it.

REFERENCES

- [1] R. A. FISHER, *Contributions to Mathematical Statistics*, John Wiley and Sons, New York, 1950.
- [2] R. A. FISHER, *Statistical Methods and Scientific Inference*, Oliver and Boyd, London, 1956.
- [3] M. H. QUENOUILLE, *The Fundamentals of Statistical Reasoning*, Charles Griffin and Co., London, 1958.
- [4] ROBERT J. BUEHLER, "Some validity criteria for statistical inferences," *Ann. Math. Stat.*, Vol. 30 (1959), pp. 845-863.
- [5] D. R. COX, "Some problems connected with statistical inference," *Ann. Math. Stat.*, Vol. 29 (1958), pp. 357-372.
- [6] M. A. CREASY, "Limits for the ratio of means," *J. Roy. Stat. Soc.*, Vol. 16, Ser. B (1954), pp. 186-194 (and discussion pp. 204-222).
- [7] E. C. FIELLER, "Some problems in interval estimation," *J. Roy. Stat. Soc.*, Vol. 16, Ser. B (1954), pp. 175-185 (and discussion pp. 204-222).
- [8] SIR RONALD FISHER, "Inverse probability," *Proc. Camb. Philos. Soc.*, Vol. 26 (1930), pp. 528-535. (Reprinted as paper 22 in Fisher [1].)
- [9] R. A. FISHER, "Statistical methods and scientific induction," *J. Roy. Stat. Soc.*, Ser. B, Vol. 17 (1955), pp. 69-78.
- [10] D. A. S. FRASER, "The fiducial method and invariance," *Biometrika*, December (1961).
- [11] P. R. HALMOS, *Measure Theory*, D. Van Nostrand and Co., New York, 1950.
- [12] D. V. LINDLEY, "Fiducial distribution and Bayes' theorem," *J. Roy. Stat. Soc.*, Ser. B, Vol. 20 (1958), pp. 102-107.
- [13] J. G. MAULDON, "Pivotal quantities for Wishart's and related distributions, and a paradox in fiducial theory," *J. Roy. Stat. Soc.*, Vol. 17, Ser. B (1955), pp. 79-85.
- [14] J. NEYMAN, "On two different aspects of the representative method," *J. Roy. Stat. Soc.*, Vol. 97 (1934), pp. 558-573.

- [15] E. J. G. PITMAN, "The estimation of the location and scale parameters of a continuous population of any given form," *Biometrika*, Vol. 30, pp. 391-421.
- [16] JOHN W. TUKEY, "Some examples with fiducial relevance," *Ann. Math. Stat.*, Vol. 28 (1957), pp. 687-695.
- [17] DAVID L. WALLACE, "Conditional confidence level properties," *Ann. Math. Stat.*, Vol. 30 (1959), pp. 864-876.
- [18] B. L. WELCH, "On confidence regions and sufficiency, with particular reference to parameters of location," *Ann. Math. Stat.*, Vol. 10 (1939), pp. 58-69.