

But the series $\sum (m+1)x^m$ converges for $|x| < 1$, therefore \bar{t}_∞ is finite.

COROLLARY 3.1. *A necessary and sufficient condition for the process to be divergent is that \bar{t}_∞ shall be finite.*

The result of (1.1) follows immediately.

COROLLARY 3.2. *For a birth and death process with no lower absorbing barrier $P(t_\infty < \infty)$ is either zero or 1.*

PROOF. If \bar{t}_∞ is finite then, from Theorem 2, we have for all $t > \bar{t}_\infty$

$$P(t_\infty > t) \leq \frac{\bar{t}_\infty}{t}$$

But $(\bar{t}_\infty/t) \rightarrow 0$ as $t \rightarrow \infty$ so that

$$(3.9) \quad \lim_{t \rightarrow \infty} P(t_\infty < t) = 1, \quad \text{or equivalently} \quad \lim_{t \rightarrow \infty} \sum_n p_n(t) = 0.$$

It follows immediately from Theorem 3 that, if $P(t_\infty < \infty)$ is not zero, then \bar{t}_∞ is finite, so that the probability must be 1.

4. Acknowledgements. I wish to thank Professor Casper Goffman for his assistance and advice during the direction of this work; I wish also to thank the referees for their helpful suggestions and for drawing my attention to references [1], [3] and [5].

REFERENCES

- [1] M. S. BARTLETT, *Stochastic Processes*, Cambridge University Press, 1955, p. 88.
- [2] H. CRAMÉR, *Mathematical Methods of Statistics*, Princeton University Press, 1946, p. 182.
- [3] R. L. DOBRUSIN, *Uspehi Matem. Nauk* (N.S.) 7 (1952) No. 6 (52) (185-191). Abstract in *Math. Rev.*, Vol. 14, p. 567.
- [4] W. FELLER, *Probability Theory and Its Applications*, John Wiley & Sons, New York, 1950, pp. 371-373.
- [5] S. KARLIN AND J. MCGREGOR, "Representation of a class of stochastic processes," *Proc. Nat. Acad. Sci.*, Vol. 41, pp. 387-391.
- [6] G. E. H. REUTER AND W. LEDERMANN, "On the differential equations for the transition probabilities of Markov processes with enumerably many states," *Proc. Cambridge Philos. Soc.*, Vol. 49 (1953), pp. 247-262.

A REGRESSION ANALYSIS USING THE INVARIANCE METHOD

BY D. A. S. FRASER

University of Toronto

1. Summary. The invariance method is applied to a regression problem for which the "errors" have a rectangular distribution. The invariance method can also be applied to produce good estimates for the regression problem when the "errors" form a sample from any fixed distribution.

Received November 4, 1955; revised November 26, 1956.

2. Introduction. The invariance method is discussed in, for example, Blackwell and Girshick [1]. We summarize briefly its form for estimation. Let θ be a parameter that indexes the probability distributions and let there be a group of transformations s on the sample space that leaves the class of probability distributions unchanged. Suppose that the group of transformations is such that any of the probability distributions can be transformed into any other. This implies that the risk function for any invariant procedure is constant valued. Let $m(x)$ label the invariant subsets on the sample space for x . If s^* is the transformation on the parameter space corresponding to the transformation on the sample space, then it is easily seen that minimum risk estimator $f(x)$ may be found from any invariant estimator $f_0(x)$ by finding a transformation s_m^* for each m such that

$$E_{\theta_0}\{W(s_m^*f_0(x), \theta_0) \mid m(x) = m\}$$

is minimized for any fixed θ_0 where $W(f, \theta)$ is an invariant loss function and $f(x) = s_{m(x)}^*f_0(x)$.

3. A regression problem. This problem was suggested by Prof. E. G. Olds. Let Y_1, \dots, Y_n be real valued random variables with the following structure:

$$(1) \quad Y_i = \sum_1^r \beta_j x_{ij} + U_i,$$

where U_1, \dots, U_n are independent random variables and each is uniformly distributed with mean 0 and known range δ . In vector notation we have $\mathbf{Y} = \sum \beta_j \mathbf{x}_j + \mathbf{U}$. The x_{ij} are given numbers and the β_j are known regression coefficients. The problem is to obtain good estimates of the regression coefficients. In this section we find invariant estimators with minimum variances.

To simplify the notation we consider the case having $r = 2$ and

$$(2) \quad Y_i = \alpha + \beta x_i + U_i,$$

where the x_i have been adjusted so that $\sum x_i = 0$. Let the loss function for an estimate (f, g) of (α, β) be a weighted sum of the squared errors:

$$(3) \quad W(f, g; \alpha, \beta) = p(f - \alpha)^2 + q(g - \beta)^2 \quad 0 \leq p, q.$$

As a class of transformations consider

$$(4) \quad \{y'_i = y_i + a + bx_i (i = 1, \dots, n) \mid (a, b) \in R^2\}.$$

It is straightforward to see that this class of transformations is an invariant class. Also the induced group of transformations on the parameter space is easily seen to be

$$(5) \quad \{\alpha' = \alpha + a, \beta' = \beta + b \mid (a, b) \in R^2\}.$$

Any $(\alpha, \beta) \in R^2$ can be transformed into any other point $(\alpha', \beta') \in R^2$; hence the group leaves no set invariant. Now, restricting ourselves to invariant estimators, we find that an estimator $(f(\mathbf{y}), g(\mathbf{y}))$ for (α, β) has the value

$$(6) \quad [f(\mathbf{y}^0) + a, g(\mathbf{y}^0) + b]$$

at the point

$$(y_1^0, \dots, y_n^0) + a(1, \dots, 1) + b(x_1, \dots, x_n).$$

For convenience in describing (6) we introduce new coordinates in R^n , say w_1, \dots, w_n , using $(1, \dots, 1)$ and (x_1, \dots, x_n) as the unit points for the first two coordinates w_1, w_2 and any $(n - 2)$ orthogonal vectors for the remaining coordinates. Then letting f^*, g^* be the functions f, g expressed in terms of the new coordinates, we obtain from (6):

$$(7) \quad \begin{aligned} f^*(a, b, w_3, \dots, w_n) &= a + f^*(0, 0, w_3, \dots, w_n), \\ g^*(a, b, w_3, \dots, w_n) &= b + g^*(0, 0, w_3, \dots, w_n). \end{aligned}$$

We need minimize the risk only for a single parameter value, say $\alpha = 0, \beta = 0$, and it will be uniformly minimized. The risk when $\alpha = 0, \beta = 0$ is

$$(8) \quad \begin{aligned} & k \int_C [p(a + f^*(0, 0, w_3, \dots, w_n))^2 \\ & \quad + q(b + g^*(0, 0, w_3, \dots, w_n))^2] da db dw_3, \dots, dw_n \\ & = kp \int_C [a + f^*(0, 0, w_3, \dots, w_n)]^2 da db dw_3, \dots, dw_n \\ & \quad + kp \int_C [b + g^*(0, 0, w_3, \dots, w_n)]^2 da db dw_3, \dots, dw_n, \end{aligned}$$

where k is the constant value of the Jacobian from y_1, \dots, y_n to w, \dots, w and C is the set of values of (a, b, w_3, \dots, w_n) corresponding to the "cube" $[-\delta/2, \delta/2]^n$ in the coordinates y_1, \dots, y_n . It is easily seen that the values for $f^*(0, 0, w_3, \dots, w_n)g^*(0, 0, w_3, \dots, w_n)$ which minimize the risk are such that $[-f^*(0, 0, w_3, \dots, w_n), -g^*(0, 0, w_3, \dots, w_n)]$ is the center of gravity of the set

$$C(w_3, \dots, w_n) = \{(a, b) \mid (a, b, w_3, \dots, w_n) \in C\},$$

the w_3, \dots, w_n section of C . This choice of f^*, g^* produces the minimum risk invariant estimate for α, β .

The determination of the values of the functions $f^*(0, 0, w_3, \dots, w_n), g^*(0, 0, w_3, \dots, w_n)$ can, however, be simplified. If we change the sign of the coordinates of all points in $C(w_3, \dots, w_n)$, then the center of gravity of the new set will have as coordinates the minimizing values $f^*(0, 0, w_3, \dots, w_n), g^*(0, 0, w_3, \dots, w_n)$, determined in the paragraph above. This altered set has, however, a simple interpretation. It is the set of points (a, b) such that the cube C , shifted to have center at $(a, b, 0, \dots, 0)$, contains the point $(0, 0, w_3, \dots, w_n)$. Similarly, the value of the estimator,

$$[a' + f^*(0, 0, w_3, \dots, w_n), b' + g^*(0, 0, w, \dots, w)],$$

or $[f^*(a', b', w_3, \dots, w_n), g^*(a', b', w_3, \dots, w_n)]$, is the center of gravity of the points (a, b) for which the cube C , shifted to have center at $(a, b, 0, \dots, 0)$, contains the point $(a', b', w_3, \dots, w_n)$. However, it is equivalent to state that the estimate of (α, β) is the center of gravity of the points (a, b) for which the line $y = a + bx$ is a possible regression line for the observed points $(x_1, y_1), \dots, (x_n, y_n)$, i.e., for which $y = a + bx$ is within $\delta/2$ vertically of each point $(x_1, y_1), \dots, (x_n, y_n)$.

It is of interest to note that the estimates of α and β are \bar{y} and $\sum yx_i / \sum x_i^2$ plus corrections which depend only on the deviations from the usual regression line. This is essentially the invariance requirement.

The methods of Section 3 up to formulas (7) and (8) may be applied in much the same manner to any regression problem for which the errors are a sample from some given fixed distribution.

REFERENCE

- [1] DAVID BLACKWELL AND M. A. GIRSHICK, *Theory of Games and Statistical Decisions*, John Wiley & Sons, New York, 1954.

ON DISCRETE VARIABLES WHOSE SUM IS ABSOLUTELY CONTINUOUS¹

BY DAVID BLACKWELL

University of California, Berkeley

1. Summary. If $\{Z_n\}$, $n = 1, 2, \dots$ is a stationary stochastic process with D states $0, 1, \dots, D - 1$, and $X = \sum_1^\infty Z_k/D^n$, Harris [1] has shown that the distribution of X is absolutely continuous if and only if the Z_n are independent and uniformly distributed over $0, 1, \dots, D - 1$, i.e., if and only if the distribution of X is uniform on the unit interval. In this note we show that if $\{Z_n\}$, $n = 1, 2, \dots$ is any stochastic process with D states $0, 1, \dots, D - 1$ such that $X = \sum_1^\infty Z_n/D^n$ has an absolutely continuous distribution, then the conditional distribution of $R_k = \sum_{n=1}^\infty Z_{k+n}/D^n$ given Z_1, \dots, Z_k converges to the uniform distribution on the unit interval with probability 1 as $k \rightarrow \infty$. It follows that the unconditional distribution of R_k converges to the uniform distribution as $k \rightarrow \infty$. Since if $\{Z_n\}$ is stationary the distribution of R_k is independent of k , the result of Harris follows.

2. Proof of the theorem.

THEOREM. If $\{Z_n\}$, $n = 1, 2, \dots$ is a sequence of random variables, each assuming only values $0, 1, \dots, D - 1$ such that $X = \sum_1^\infty Z_n/D^n$ has an absolutely continuous distribution, and

$0 < \lambda \leq 1$, then $U_k(\lambda) = P(\sum_1^\infty Z_{k+n}/D^n < \lambda \mid Z_1, \dots, Z_k) \rightarrow \lambda$ with probability 1 as $k \rightarrow \infty$.

Received July 18, 1956.

¹ This paper was prepared with the partial support of the Air Research and Development Command under contract AF 41(657)-29, with the USAF School of Aviation Medicine.