STATISTICAL ESTIMATION OF A CRACK DAMAGE PARAMETER IN MANUFACTURING PROCESSES*

By ALVIN D. WIGGINS (*University of California, Davis*)

Abstract. A stochastic model of the production of cracks in certain types of linear manufacturing processes is constructed. Maximum likelihood is used to estimate a parameter, representing crack damage, of the resulting probability law.

1. Introduction. The purpose of this paper is to construct a mathematical model of the production of defects in certain types of manufacturing processes, to propose a certain parameter as a measure of defect or crack damage, and to present an estimate of this parameter using sampling data. The specific type of process which we have in mind is one which can be considered in some sense as temporally sequential in character. The following are examples of the type of process which can qualify for treatment in terms of the mathematical model developed herein:

i) extrusion processes;
ii) acetylene or arc-welding processes;
iii) hot- or cold-rolling of steel.

For purpose of concreteness only, we shall develop the mathematical model in terms of application to welding, keeping in mind that extrusion or rolling of steel could equally well have been used.

2. Derivation. We shall suppose that a manufacturing process involving welding is such that it produces a random distribution of cracks of varying lengths within the weld. It will be helpful if we view this random distribution as a two-stage process. In the first stage we view only the endpoints of the cracks, more specifically the left-hand endpoints, as being distributed in some random fashion along the length of the weld. In the second stage we associate with each left-hand endpoint a crack, the length of which is a random variable. We now imagine an axis of real numbers, having a fixed origin, and lying parallel to the weld (assuming that the two edges of the metal to be joined are straight).

The mathematical abstraction of the problem begins by projecting the left-hand endpoints of the cracks perpendicularly onto the axis of real numbers. Such points we shall call *E-points* for brevity. We next project the *lengths* of the cracks onto the real axis in the same fashion. The perpendicular projection of a crack length onto the real axis results in an interval of real numbers, the left-hand endpoint of which is an E-point. Such an interval will be called a *segment*.

Let L denote the total length of weld produced. Consider now the real axis, choose arbitrarily (subject to a mild restriction to be made later) a set of m positive numbers t_i, $i = 1, \cdots, m$, satisfying $0 < t_1 < \cdots < t_m < L$ (see Fig. 1), and let the $\{t_i\}$ remain

* Received August 19, 1972; revised version received August 13, 1973.

fixed throughout the remainder of the analysis. We now ask the following m questions: "In how many segments is t_i ($i = 1, \cdots, m$) contained?", where the word "segment" has the meaning assigned to it above. Note that, operationally, the choice of a t_i corresponds exactly to making a transverse cut of a weld, of polishing, etching, and inspecting the face of a specimen for cracks. The question: "In how many segments is t_i contained?" is operationally equivalent to the question: "How many cracks appear in the weld face located at t_i?"

Corresponding to the first stage of randomness mentioned above we make

Assumption 1. The locations of all E-points in the interval $[0, L)$ are independent identically distributed random variables having the uniform distribution on the interval $[0, L)$:

$$P\{X \le x\} = F(x) = 0, x < 0,$$
$$= x/L, \quad 0 \le x < L,$$
$$= 1, \quad L \le x.$$

In order that this assumption be reasonable, we require that the physical parameters underlying the welding process remain unchanged throughout that portion of the weld which will be subjected to sampling and statistical analysis. Corresponding to the second stage of randomness mentioned above we make

Assumption 2. Each crack has a length ("segment"), and these lengths are independent identically distributed random variables having the distribution $P\{Y \le y\} = G(y)$, where $G(y) = 0$ for $y \le 0$, and $G(y) = 1$ for $y \ge z > 0$.

Assumption 3. If $(X_1, Y_1), \cdots, (X_M, Y_M)$ are the beginnings and lengths, respectively, of any number M of cracks, then the random variables $X_1, Y_1, \cdots, X_M, Y_M$ are totally independent. From now on we shall say that a crack beginning at X and having length Y is detected at the point t_i if the interval $(X, X + Y)$ contains t_i.

The quantity z in Assumption 2 requires some explanation. We suppose that the manufacturer, being duly cognizant of contractual obligations, has taken pains to remove, or in some way eliminate, as many of the defects of manufacture as possible. Assumption 2 means that he has been successful in removing all cracks of length greater than z, and unsuccessful in removing any cracks of length z or less.

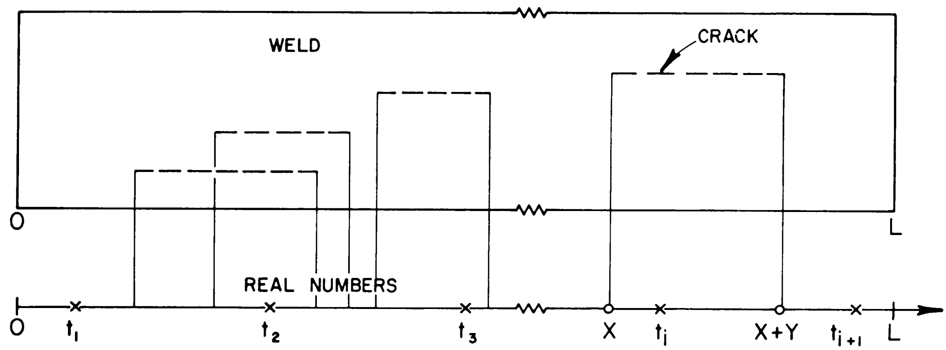

FIGURE 1

We proceed now to the calculation of the joint probability law of the random variables denoting the number of cracks detected at each point t_i. It was hinted earlier that a mild restriction be placed on the otherwise arbitrary choice of the set $\{t_i, i = 1, \cdots, m\}$ of sampling points. This restriction consists of the assumption that $t_i - t_{i-1} > z$. Under this assumption it is obvious that no crack can cross more than one of the points t_i. Suppose first that we choose an observed value of the random variable X according to the probability distribution F. Next choose an observed value of the random variable Y according to the probability distribution G. The observed value x of X (E-point) will fall in some subinterval, say $[t_{i-1}, t_i)$. The observed value y of Y (segment), when added to x, may either cross t_i or it will cross none of the t's. For a single performance of the foregoing experiment let C_ϕ denote the event consisting of no observed crossings of any of the t's, and C_i the event consisting of the crossing of t_i. The probabilities $p_\phi = P(C_\phi)$, and $p_i = P(C_i)$ for $i = 1, 2, \cdots, m$ are computed as follows:

$$\begin{aligned} p_i &= P\{t_i - z < X < t_i \quad \text{and} \quad X + Y > t_i\} \\ &= \int_{t_i - z}^{t_i} \int_{t_i - x}^{z} dF(x)\, dG(y) = L^{-1} \int_{t_i - z}^{t_i} [G(z) - G(t_i - x)]\, dx \\ &= L^{-1}[z - z + \mu_G(z)] = L^{-1}\mu_G(z), \qquad i = 1, \cdots, m. \end{aligned} \tag{2.1}$$

Since

$$p_\phi + \sum_{i=1}^{m} p_i = 1,$$

we have

$$p_\phi = 1 - \sum_{i=1}^{m} p_i - 1 = mL^{-1}\mu_G(z). \tag{2.2}$$

Now let K_i be a random variable denoting the number of crossings of t_i (number of cracks in the ith face), $i = 1, \cdots, m$. The conditional probability of the event $K_1 = k_1, \cdots, K_m = k_m$, given that altogether M cracks have occurred, is obviously the multinomial probability

$$\begin{aligned} P = P\{K_i = k_i, i = 1, \cdots, m\} &= \frac{M!}{k_1! \cdots k_m!\,(M - k)!} p_1^{k_1} \cdots p_m^{k_m} p_\phi^{M-k} \\ &= \frac{M!}{\left(\prod_{i=1}^{m} k_i!\right)(M - k)!} \left[\frac{\mu_G(z)}{L}\right]^k \left[1 - \frac{m\mu_G(z)}{L}\right]^{M-k}, \end{aligned} \tag{2.3}$$

where

$$k = \sum_{i=1}^{m} k_i .$$

We now propose, as a measure of crack damage, the use of the parameter

$$\rho = (M/L)\mu_G(z).$$

Note that ρ is a dimensionless number, since M/L has the dimension of reciprocal length and $\mu_G(z)$ has the dimension of length. The parameter ρ can be thought of as the fraction of total weld length containing cracks. Note that this fraction can theoretically exceed 1, due to the possibility of separate cracks existing side by side.

3. Estimation of ρ. Rewriting (2.3) in terms of ρ, we have

$$P = \frac{M!}{\left(\prod_{i=1}^{m} k_i!\right)(M-k)!}\left(\frac{\rho}{M}\right)^k\left(1-\frac{m\rho}{M}\right)^{M-k}, \tag{3.1}$$

and this becomes maximum for

$$\rho = \hat{\rho} = k/m = \sum_{i=1}^{m} k_i/m = \bar{k}. \tag{3.2}$$

Thus the maximum likelihood estimate of ρ is the average number of cracks per face, detected in m specimen faces, subject to the condition that the specimen faces are separated by more than z. It is easily shown that $\hat{\rho} = \bar{k}$ is an unbiased estimate of ρ, with variance

$$\sigma_{\hat{\rho}}^2 = (\rho/m)(1 - m\rho/M).$$

If the total length of welded seam to be produced is very large, one might consider a passage to the limit, with $M \to \infty$, $L \to \infty$, $M/L \to \lambda$. In this case the multinomial probability law (2.3) becomes asymptotically the product of m independent identical Poisson probability laws with parameter $\rho = \lambda\mu_G(z)$.

REFERENCES

[1] H. Cramér, *Mathematical methods of statistics*, Princeton University Press, Princeton, 1946
[2] W. Feller, *An introduction to probability theory and its application*, 2nd ed., Wiley, New York, 1957