TUMBLER MORTALITY

George W. Brown
Statistical Laboratory, Iowa State College
AND
Merrill M. Flood
Princeton, New Jersey

I. INTRODUCTION

1. *The statistical problem.* Life tables are constructed for certain types of equipment, under specified conditions of use, from records of experience with the equipment in service. Items such as electric lamps, glass tumblers, and silk stockings, under the normal conditions of use, in effect may be said to "die" when they "burn out," "crack," or "run." The data and analysis necessary to obtain an estimate of the mortality distribution, life expectancy, and other similar characteristics of such equipment, are exactly analogous to the more familiar techniques applied in the case of human mortality experience. This paper presents the results of an analysis of a service test that was conducted in order to estimate the mean lengths of life for each of two types of glass tumbler when used in a particular cafeteria, and discusses statistical techniques that proved to be well-suited for the treatment of the problem. Technological considerations of a model for tumbler breakage are given, leading to familiar mortality curves of Makeham-Gompertz type.

2. *The service test.** A fixed number of tumblers of each of two types, called "annealed" and "toughened," were kept in service at all times in the test cafeteria. At the end of each week, each broken tumbler was replaced by a new one of the same type. A record was kept of the date each tumbler was introduced into service and of the week each broken tumbler was removed from service. The test was continued for 78 weeks employing 60 annealed tumblers and 120 toughened tumblers.

II. THE SERVICE TEST ANALYSIS

1. *Discussion.* What complicates the analysis of this experiment is the fact that the (necessarily) finite nature of the experiment leads to truncated samples. This unpleasant feature might be avoided by ex-

* The service test was conducted by H. Scholtz and A. Basham, of the Preston Laboratories. The work was sponsored by the Libbey Glass Company, the Federal Glass Company, and the Preston Laboratories. Acknowledgment is due to T. Collins and H. Black, H. H. Blau and Conrad Stone, and F. W. Preston, as the representatives of these three companies who provided impetus and guidance for the test program.

cluding from the analysis any generations not completely extinct by the end of the experiment. The procedure would, however, be extremely wasteful in the annealed case, and would be fatal in the toughened case, since practically no data would be retained. A further difficulty is the fact that each subsequent generation of replacements has its history truncated relatively earlier than the preceding generations. Hence, except for direct estimation of the stable replacement rate by averaging weekly breakages over a *sufficiently long* experiment, any estimates *must* be based on an assumed analytical form for the mortality distribution.

The method described below makes effective use of all the data and supplies essentially optimum estimates of the parameters, provided only that the conditions of handling of the tumblers may be considered homogeneous from week to week. Mean lengths of life were estimated as 8.7 and 60.4 weeks for the annealed and toughened tumblers, respectively.

2. *Statistical method.* Consider a batch of $\overline{N}_\alpha$ tumblers, introduced together into the experiment at the beginning of the α^{th} week. Let $n_{\alpha i}$ be the number of tumblers of this group broken in the i^{th} week following their introduction. Then

$$N_{\alpha i} = \overline{N}_\alpha - \sum_{j<i} n_{\alpha j}$$

is the number of individuals of this group surviving $i-1$ weeks, hence eligible for possible breakage in the i^{th} week. Then $n_{\alpha i}$ is an observation on a sample of $N_{\alpha i}$ individuals distributed according to the binomial distribution with probability equal to the conditional probability of breakage in the i^{th} week of service, assuming survival of $i-1$ weeks of service. The aggregate of all such samples for successive values of i, terminating either when $N_{\alpha i}=0$ or when the last week of the experiment terminates the observations, contains all the available information on the history of the generation in question. Moreover the successive $n_{\alpha i}$, for given $N_{\alpha i}$, are mutually independent in probability.

Let the mortality distribution be $f(x)dx$, i.e., the probability of breakage in the i^{th} week since introduction is given by

$$P_i = \int_{i-1}^{i} f(x)dx.$$

Then it is easily seen that the conditional probability relevant to the samples considered above is given by

$$P_i^c = \frac{\int_{i-1}^{i} f(x)dx}{\int_{i-1}^{\infty} f(x)dx} = \frac{\int_{i-1}^{i} f(x)dx}{1 - \int_{0}^{i-1} f(x)dx}.$$

Assuming homogeneity of experimental conditions from week to week the samples from the various generations may be pooled by summation, so that

$$N_i = \sum_{\alpha} N_{\alpha i} \quad \text{and} \quad n_i = \sum_{\alpha} n_{\alpha i}$$

provide a sample on the binomial distribution with probability P_i^c.

The method of "maximum likelihood" furnishes the key to the problems of estimation and testing concerned with the distribution $f(x)dx$. The quantities P_i^c depend directly on the form of $f(x)$ and on the parameters involved. The joint probability of obtaining the observations n_i for given N_i can be evaluated as a function of the P_i^c, hence as a function of the basic parameters. The parameter values which maximize the joint probability are essentially optimum estimates of the parameters. The ratio of this maximum joint probability to the maximum obtainable with arbitrary P_i^c measures the "goodness of fit," therefore supplies a criterion for testing the form of $f(x)$. Other statistical tests can be based on maxima obtained by holding some of the parameters fixed and maximizing with respect to the remainder.

Significance tests based on the maximum likelihood method are, except in certain special instances, generally based on asymptotic approximations. In the cases under consideration here no appreciable loss results from replacing the likelihood ratio by the asymptotically equivalent χ^2 corresponding to the Pearson test for goodness of fit, i.e.,

$$\chi^2 = \sum_i \frac{(n_i - N_i P_i^c)^2}{N_i P_i^c (1 - P_i^c)}.$$

Let $f(x)$ depend on parameters $\theta_1, \theta_2, \cdots, \theta_p$, and assume that samples are included from a total of k basic intervals, with $k > p$. Denote by $\bar{\theta}$ the parameters values which *minimize* $\chi^2(\theta_1, \theta_2, \cdots, \theta_p)$. Then the $\bar{\theta}$ values are estimates of the corresponding parameters, and

$$\chi^2_{k-p} = \chi^2(\bar{\theta}_1, \bar{\theta}_2, \cdots, \bar{\theta}_p)$$

is distributed approximately according to the χ^2 distribution with $k-p$ degrees of freedom, provided $f(x)$ actually has the functional form assumed. Thus an excessive value of χ_{k-p}^2 would indicate a poor fit to the assumed form. Now consider a fixed value θ_1, and let $\bar{\bar{\theta}}_2, \cdots, \bar{\bar{\theta}}_p$ represent the new values which minimize χ^2 with θ_1 fixed.

$$\chi^2_{k-p+1}(\theta_1) = \chi^2(\theta_1, \bar{\bar{\theta}}_2, \cdots, \bar{\bar{\theta}}_p)$$

has, approximately, the χ^2 distribution with $k-p+1$ degrees of freedom, and

$$\chi^2_1(\theta_1) = \chi^2_{k-p+1} - \chi^2_{k-p}$$

is independent of χ_{k-p}^2, and distributed like χ^2 with one degree of freedom. In this manner, the standard error of estimate for $\bar{\theta}_1$ may itself be estimated, by varying θ_1 around $\bar{\theta}_1$, determining $\bar{\bar{\theta}}_2, \cdots, \bar{\bar{\theta}}_p$ for the given θ_1, and extracting the square root of the quantity $\chi_1^2(\theta_1)$. This procedure corresponds to the frequently used "fiducial argument."

3. *Application of the Method.* In the present instance there was theoretical justification for trying the incomplete gamma function distribution[1]

$$f(x)dx = \frac{\lambda^\mu}{\Gamma(\mu)} x^{\mu-1} e^{-\lambda x} dx.$$

No obvious direct calculation exists, except in the special case $\mu = 1$, for determining the parameter values which best fit the data. Fortunately the incomplete gamma function has been tabled (K. Pearson, *Biometrika*, 1922), in the form

$$I(u, p) = \frac{1}{\Gamma(p+1)} \int_0^{u\sqrt{p+1}} e^{-v} v^p dv.$$

It is therefore possible to carry out an empirical minimization of χ^2 by successive approximations. In the neighborhood of the minimum one can proceed by minimizing with respect to λ for fixed μ, then minimizing with respect to μ using the λ value obtained by the previous process, etc. The process is laborious, since each minimization requires the computation of several complete trials, but feasible, for lack of a better procedure.

In analyzing the annealed tumblers, the basic unit of one week was preserved, up to 20 weeks, pooling weeks 21–25 and weeks 26–30 because of the small sample sizes in these groups. The one week sub-

[1] A discussion, with applications, of the renewal theory based on this distribution function may be found in a paper by A. W. Brown entitled "A Note on the Use of a Pearson Type III Function in Renewal Theory" published in *The Annals of Mathematical Statistics*, Vol. XI, December 1940, pps. 448–453.

division was unnecessarily fine for the toughened tumblers, because of their considerably longer life, so that the data were grouped according to five-week units, except for the pooling of weeks 56–65 and 66–75. The resultant reduction in labor more than justified the trivial loss of information associated with the grouping.

The accompanying sample computations (Tables 1, 2) show the

TABLE 1

ANNEALED CASE ($\lambda = 0.183$, $\mu = 1.60$)

Week	Exposed N	Broken n	n/N	F_i	P_i^c	σ	δ	δ/σ
1	549	23	.042	.0410	.041	.0086	.001	.12
2	521	38	.073	.1127	.075	.0115	−.002	−.17
3	470	48	.102	.1934	.091	.0133	.011	.83
4	415	40	.096	.2763	.103	.0149	−.007	−.47
5	371	36	.097	.3564	.111	.0163	−.014	−.86
6	331	42	.127	.4317	.117	.0177	.010	.56
7	285	35	.123	.5008	.122	.0194	.001	.05
8	247	33	.134	.5638	.126	.0211	.008	.38
9	210	26	.124	.6199	.129	.0231	−.005	−.22
10	182	17	.093	.6698	.131	.0250	−.038	−1.52
11	163	22	.135	.7145	.134	.0267	.001	.04
12	139	16	.115	.7537	.136	.0291	−.021	−.72
13	123	8	.065	.7876	.138	.0311	−.073	−2.35
14	114	19	.167	.8174	.140	.3026	.027	.83
15	95	17	.179	.8432	.141	.0357	.038	1.06
16	78	13	.167	.8656	.143	.0396	.024	.61
17	65	8	.123	.8850	.144	.0436	−.021	−.48
18	56	7	.125	.9016	.145	.0470	−.020	−.43
19	49	12	.245	.9161	.146	.0504	.099	1.96
20	37	5	.135	.9284	.147	.0582	−.012	−.21
21–25	31	17	.548	.9680	.553	.0893	−.005	−.06
26–30	14	10	.714	.9860	.563	.1325	.151	1.14

$$\chi^2 = \Sigma(\delta/\sigma)^2 = 18.3$$

$F_i = \int_0^i f(x)dx$ (From tables of the incomplete gamma function)

$$\sigma \sqrt{\frac{P_i^c(1-P_i^c)}{N}}$$

$$P_i^c = \frac{F_i - F_{i-1}}{1 - F_{i-1}}$$

$$\delta = \frac{n}{N} - P_i^c$$

calculations corresponding to the optimum parameter values for the two types of tumblers. Figures 1 and 3 show the corresponding theoretical mortality distribution and the theoretical curve of conditional probability by weeks, for the annealed tumblers; the curves of Figures 2 and 4 are the corresponding curves for the toughened tumblers, by five-week groups.

The small minimum χ^2 values obtained indicate that in both cases the gamma distribution fits the data very well. In the annealed case, the value of 18.3 on 20 degrees of freedom is within the 50% point, that is, under random sampling from gamma function distributions the value of χ^2_{20} would exceed the observed value more than 50% of the time. In the toughened case the value of 7.0 on 11 degrees of freedom would be exceeded about 80% of the time under random sampling.

TABLE 2

TOUGHENED CASE ($\lambda=.0257$, $\mu=1.55$)

Weeks	Exposed N	Broken n	n/N	F_i	P_i^c	σ	δ	δ σ
1– 5	241	7	.029	.0279	.028	.0106	.001	.09
6–10	223	11	.049	.0757	.049	.0145	.000	.00
11–15	208	10	.048	.1315	.060	.0165	−.012	− .73
16–20	188	9	.048	.1907	.068	.0184	−.020	−1.09
21–25	170	17	.100	.2505	.074	.0201	.026	1.29
26–30	146	10	.068	.3093	.078	.0222	−.010	− .45
31–35	132	13	.098	.3660	.082	.0239	.016	.67
36–40	110	9	.082	.4199	.085	.0266	−.003	− .11
41–45	94	6	.064	.4704	.087	.0291	−.023	− .79
46–50	86	9	.105	.5179	.090	.0309	.015	.49
51–55	73	5	.068	.5625	.093	.0338	−.025	− .74
56–65	60	14	.233	.6411	.181	.0497	.052	1.05
66–75	45	10	.222	.7073	.184	.0578	.038	.66

$$\chi^2=\Sigma(\delta/\sigma)^2=7.0$$

$F_i \int_0^i f(x)dx$ (From tables of the incomplete gamma function) $\qquad \sigma\sqrt{\dfrac{P_i^c(1-P_i^c)}{N}}$

$$P_i^c=\frac{F_i-F_{i-1}}{1-F_{i-1}} \qquad \delta=\frac{n}{N}-P_i^c$$

Standard errors of estimate for the mean length of life were approximated by a minimization procedure involving fixed values for the mean. Thus, in the annealed case, the optimum values were $m=8.7$, $\sigma=6.9$, yielding $\chi^2=18.3$. Choosing $m=9.2$ led to a minimum χ^2 of 23.4 at $\sigma=7.6$ and the choice of $m=8.2$ led to a minimum χ^2 of 22.4 at $\sigma=6.4$. Now $\sqrt{23.4-18.3}=\sqrt{5.1}=2.3$ and $\sqrt{22.4-18.3}=\sqrt{4.1}=2.0$, so that the increment of .5 week is equivalent to about 2 standard errors, or the estimated standard error is $\frac{1}{4}$ week. Similarly, in the toughened case the optimum values were $m=60.4$, $\sigma=48.5$, with a χ^2 of 7.0. Minimum χ^2 values corresponding to $m=65$ and $m=55$ were 8.4 and 8.5 respectively, leading to the conclusion that a five-week interval is about 1.2 standard errors, in other words, the estimated standard error is of the order of four weeks.

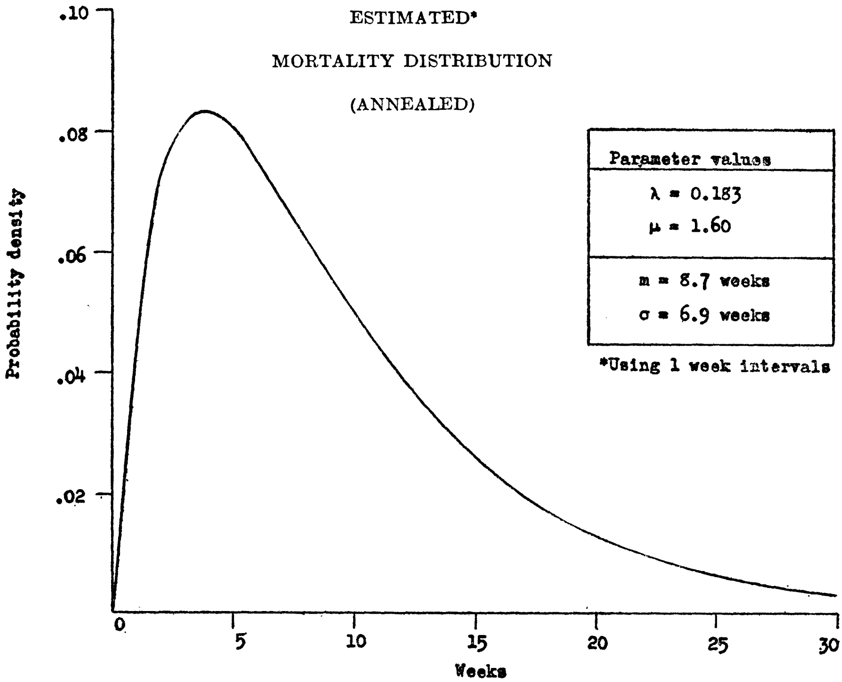

FIGURE 1

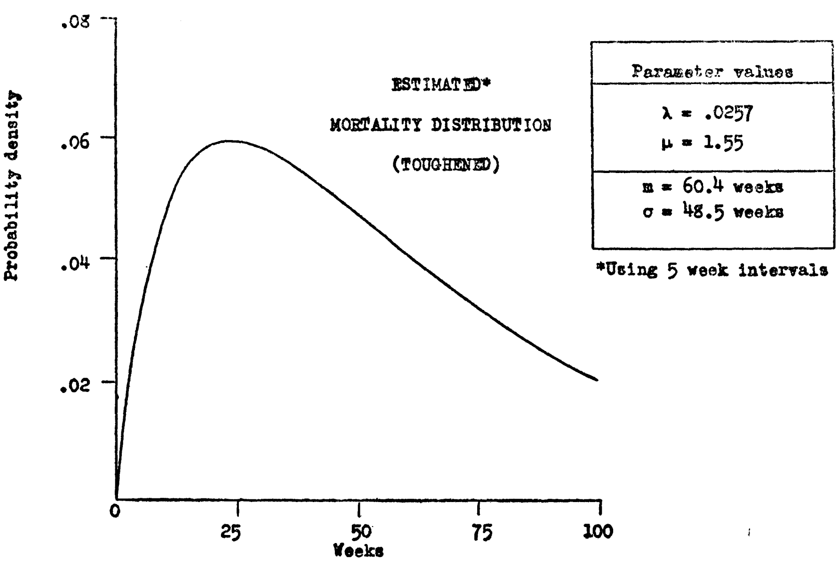

FIGURE 2

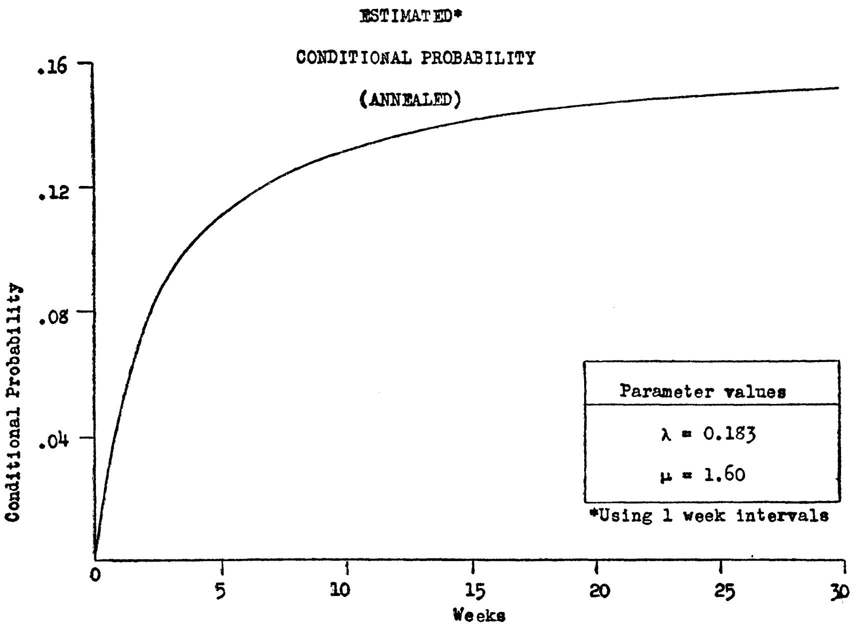

FIGURE 3

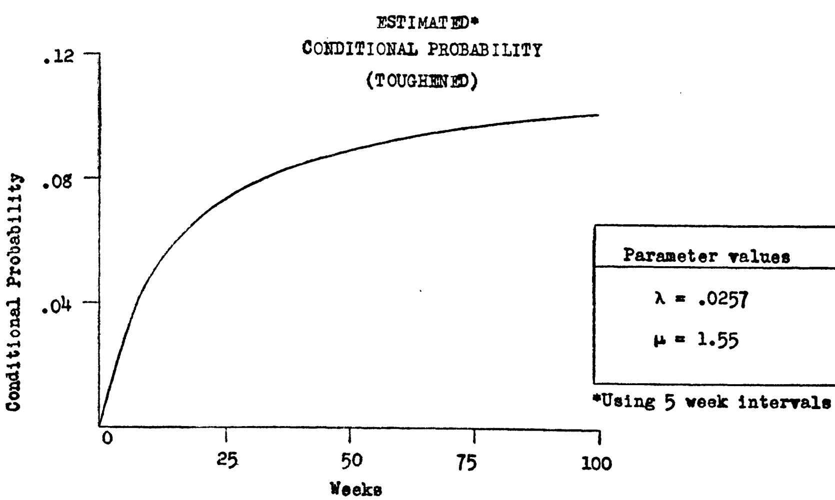

FIGURE 4

4. *Summary of test results.* Technological considerations suggested the *incomplete gamma* distribution,

$$f(x)dx = \frac{\lambda^\mu}{\Gamma(\mu)} x^{\mu-1}e^{-\lambda x}dx \qquad (\mu > 0, \lambda > 0)$$

closely related to the Pearson Type III curve, for the mortality distribution. Statistical analysis indicates that the data for both types of tumbler are consistent with this form of distribution.

The mean and standard deviation of the above distribution are given by $m=\mu/\lambda$ and $\sigma=\sqrt{\mu}/\lambda$. Estimates of the parameters for the two populations are presented in Table 3 following.

TABLE 3

ESTIMATES OF PARAMETERS

Tumbler	m	σ	λ	μ
Annealed	8.7	6.9	.183	1.60
Toughened	60.4	48.5	.0257	1.55

The estimated mean lengths of life, 8.7 weeks for the annealed, and 60.4 weeks for the toughened tumblers, are subject to uncertainties measured by standard errors of approximately 0.25 and 4.0 weeks, respectively.

The close agreement between the estimates of the exponent μ for the two cases is of some interest. The analysis does not indicate that the difference is statistically significant. Certainly the slight difference is of no practical importance. Further investigation is required to determine whether or not the observed value may be more generally applicable.

5. *Tumbler mortality distribution.* The technical considerations behind the choice of the form of the mortality distribution to be used in the analysis of the service test data are discussed in Part III, following.

Evidence is given to indicate that tumblers may be expected to break in service in accordance with the mortality distribution

$$dF = - e^{+\alpha_1}T^{\alpha_0-1}e^{-\alpha_1 T}(\alpha_1 T - \alpha_0)dT$$

where

$$T = e^{-\beta t},$$

and where α_0, α_1, and β have a definite technological significance. It is noted that this distribution resembles the incomplete gamma distribution somewhat, and represents the familiar Makeham-Gompertz formula if β is negative.

It is likely that the three-parameter distribution would fit the service test data at least as well as did the two-parameter incomplete gamma distribution that was used in the actual analysis. No attempt was made to test this possibility. The estimation methods that were used are equally applicable to either form of distribution, but would involve much more tedious computations for the three-parameter form of distribution than for the simpler incomplete gamma form.

III. TUMBLER MORTALITY DISTRIBUTIONS

1. *Technological background.*[2] Tumblers break in service for a variety of reasons. One primary cause of breakage is impact at the rim. The force of impact required to break a tumbler with a blow at the rim will depend upon the thickness of the glass, and upon very many other factors but in particular it has been found to depend upon the extent of abrasion of the rim surface.

There is good evidence that tumblers "age" in service, and in particular that they are abraded around the rim sufficiently to weaken their resistance to breakage on impact. For the purpose of this report the theoretical model for breakage is based on the assumption that tumblers in service are subjected continually to wear and impact, but that the *rate of wear* at any time during the service life is proportional to the "unwear" at that time.

For example, suppose that abrasion of the rim were the only form of wear involved, and that each point along the rim of a specific tumbler at a definite time may be classified either as abraded or unabraded. If liability to abrasion is the same for each rim point of each tumbler in service then the fraction of the rim which may be expected to be abraded at time t is

$$A(t) = 1 - e^{-\beta t},$$

where β is the proportionality factor in the rate of abrasion. The assumed exponential aging function $A(t)$ is illustrated by this rim abrasion example, but obviously applies much more generally to aging processes of all sorts which satisfy the basic assumption that *rate of wear is proportional to unwear.*

Returning to the rim example, it seems reasonable to assume that the probability that a tumbler will break, between times t and $t+\Delta t$, under a blow at the rim is

$$[\mu_1(1 - A(t)) + \mu_2 A(t)]\mu_0 \Delta t \equiv [k_0 - k_1 e^{-\beta t}]\Delta t,$$

[2] We are indebted to F. W. Preston and J. Glathart, of the Preston Laboratories, for assistance with the construction of the general technological model adopted for this discussion.

where μ_1 and μ_2 are the probabilities that a blow on the unabraded and abraded portions, respectively, will break the tumbler, and where $\mu_k\Delta t$ is the probability of a blow during the interval Δt. There will certainly be different kinds of abrasion in practice, with different effects on the probabilities of breakage, but the simple model described is the only one considered in the present report and may very possibly represent reality to a good approximation.

2. *Mortality distributions.* The technological assumptions of Section 1, preceding, lead immediately to a mortality distribution. The probability $q(t)$ that a tumbler survives to at least time t satisfies the equation

$$q(t + \Delta t) = q(t)[1 - (k_0 - k_1e^{-\beta t})\Delta t].$$

Thus

$$q(t) = e^{k_1/\beta}e^{-k_0t}e^{(-k_1/\beta)e^{-\beta t}}.$$

The mortality distribution is determined by the equation

$$df = - dq = [e^{k_1/\beta}e^{-k_0t}e^{(-k_1/\beta)e^{-\beta t}}(k_0 - k_1e^{-\beta t})]dt.$$

For convenience, set

$$T = e^{-\beta t}, \quad \alpha_0\beta = k_0, \quad \text{and} \quad \alpha_1\beta = k_1.$$

Then

$$df = - e^{\alpha_1}T^{\alpha_0-1}e^{-\alpha_1T}(\alpha_1T - \alpha_0)dT \equiv F(T)dT,$$

and

$$q(t) = e^{\alpha_1}T^{\alpha_0}e^{-\alpha_1T} \equiv Q(T).$$

A special case of interest is that in which a blow on the unabraded portion has zero probability of breakage; in this case $\mu_1=0$, so $k_0=k_1$ and $\alpha_0=\alpha_1\equiv\alpha$, whence

$$F(T)dT = - \alpha e^{\alpha}T^{\alpha-1}e^{-\alpha T}(T - 1)dT$$

and

$$Q(T) = e^{\alpha}T^{\alpha}e^{-\alpha T}.$$

3. *Discussion.* The mortality distribution, written in terms of the original technological quantities t, β, μ_0, μ_1, and μ_2, is

$$df = - \frac{\mu_0}{\beta} e^{(\mu_0-\mu_1)\mu_0/\beta}T^{(\mu_0\mu_2/\beta)-1}e^{[-(\mu_2-\mu_1)\mu_0/\beta]T}[(\mu_2 - \mu_1)T - \mu_2]dT$$

where

$$T \equiv e^{-\beta t} \quad \text{and} \quad dT = -\beta e^{-\beta t} dt.$$

It is to be noted also that the quantities α_1 and α_0 are related to the technological quantities by the relations:

$$\alpha_1 = \frac{(\mu_2 - \mu_1)\mu_0}{\beta} \quad \text{and} \quad \alpha_0 = \frac{\mu_0 \mu_2}{\beta} \cdot$$

Appropriate tumbler tests, such as the cafeteria service test discussed in Part II, preceding, provide statistical data which may be analyzed to yield estimates of the technological quantities $\mu_1\mu_0$, $\mu_2\mu_0$, and β. For statistical purposes, it is more convenient to first estimate the distribution parameters α_0, α_1, and β which then in turn determine the estimated values of the technological quantities. Computational procedures which may be used for the estimation of such parameters are described in detail in Part II.

Actually, the analysis of the cafeteria service test data was carried out on the assumption that the mortality distribution was an incomplete gamma distribution. The similarity between the two distributions is shown when df is written in the form

$$df = \beta e^{\alpha_1} t^{\alpha_0 [-\beta t / \ln t]} e^{-\alpha_1 t [e^{-\beta t} / t]} \left[\alpha_1 e^{-\beta t} - \alpha_0\right] dt.$$

If the quantities which are shown in braces were all nearly constant over the range of values of t involved in the test then the technological distribution would be essentially the same as the incomplete gamma distribution which was used in Part II.

An extensive discussion of the application of the Makeham-Gompertz survival function to problems of industrial replacement is given by Kurtz,[3] who shows empirically that a number of types of industrial property can by fitted in this manner or with the Pearson Type I distribution.

4. *The incomplete gamma distribution.* An interesting technological model for tumbler breakage is obtained[4] if it is assumed that a tumbler breaks when it is bumped the n^{th} time, and that the number of bumps is proportional to the time in service. In this case, it is appropriate to use the Poisson distribution for the number of bumps the tumbler receives, to determine the survival function $q(t)$ as follows

[3] Edwin B. Kurtz, *Life Expectancy of Physical Property*, New York (1930).

[4] We are indebted to C. P. Winsor for calling this technological model to our attention.

$$q(t) = 1 - e^{-\lambda t} \sum_{i=n}^{\infty} (\lambda t)^i / i!.$$

In this expression for $q(t)$ the quantity

$$e^{-\lambda t} (\lambda t)^i / i!$$

represents the probability that the tumbler will be bumped exactly i times in time t. The mortality distribution dg is obtained directly from the survival function $q(t)$ by differentiation, thus

$$dg = - dq = \frac{\lambda^n}{(n-1)!} t^{n-1} e^{-\lambda t} dt.$$

This distribution is exactly the incomplete gamma distribution, provided the critical number n of bumps required to break a tumbler is limited to non-negative integral values.

It is interesting to note from Part II, preceding, that the estimated values for the parameter n, there not limited to integral values, were 1.55 and 1.60 for the two types of tumblers tested. This provides some evidence that the simple bump model is not adequate, even though the incomplete gamma function did fit the experimental data quite well.