The Analysis of Exponentially Distributed Life-Times with Two Types of Failure

By D. R. Cox

Birkbeck College, University of London

[Received April, 1959]

SUMMARY

A NUMBER of alternative probability models are considered for the interpretation of failure data when there are two or more types of failure. Some of the statistical techniques that can be used for such data are illustrated on an example discussed recently by Mendenhall and Hader.

1. Introduction

Mendenhall and Hader (1958) have recently given an interesting account of a model for the analysis of failure-time distributions when there are two, or more, types of failure. They illustrate their theory by analysing some data on the failure-times of radio transmitter receivers; the failures were classed into two types, those confirmed on arrival at the maintenance centre and those unconfirmed. In the present paper their example is used to illustrate and distinguish between a number of models that can be used for this type of data.

The essential feature of the problem is that we have independent individuals exposed to risk, and that on failure an individual is withdrawn from risk. We observe, for example, that individual number one fails after life-time t, and that the failure is say of the first type: this means that we know the time at which failure of the first type occurs, but only that failure of the second and other types had not occurred by time t. In many applications, including Mendenhall and Hader's, the sample contains individuals that have not failed at the end of the period of the observation. Thus in their example no receivers were operated after 630 hr.

Data like this arise in several fields in addition to industrial life-testing. For example, in medical and actuarial work the estimation and comparison of death rates from a particular cause requires corrections for deaths from other causes. In particular Seal (1954) and Elveback (1958) have discussed the more theoretical aspects of this in connection with actuarial work and given numerous references. Sampford (1954) has dealt with similar problems in bioassays. In tensile strength testing there may be two or more types of failure, for example jaw breaks and fractures in the centre of the test specimen; here the observation is load on failure, not life on failure. A further interesting application is in experimental psychology. Audley (1957) has interpreted latency measurements in learning experiments by postulating independent Poisson processes of A-responses and B-responses; the first process for which an event occurs is considered to determine the nature (A or B) of the response and the time at which it occurs.

We shall consider in the present paper a number of probability models that can be used for these problems; the models become identical in special cases. The choice between the models is a question partly of which fits the data best, and partly of which set of assumptions is the most reasonable physical representation of the process. If one type of model fits much better than another, information may be obtained about the underlying mechanism.

For the most part we consider here systems in which the frequency distributions involved are exponential. This is a serious restriction, even though it is known that in some applications distributions close to the exponential are obtained. There are two situations in which an exponential distribution would be expected on general grounds. First failure may be due to an external point occurrence, for example an accident, arising randomly in an age-independent way. Secondly there may be many more or less independent causes of failure, when the observed failure time is the smallest of a number of independent random variables; under some rather special conditions the resulting distribution will be exponential. On the other hand, if, for example, there is a single process of wear going on at a fairly steady rate, the distribution of failure-time may be far from exponential.

2. Some Probability Models

We consider first two models that involve exponential distributions for the component life-times. We suppose that there are two types of failure; the generalization when there are more than two is straightforward.

Model A. Independent Poisson risks.—Imagine that failures of Types I and II occur independently in Poisson processes with parameters λ_1 and λ_2; that is we have random variables T_1, T_2 independently distributed with p.d.f.'s $\lambda_1 e^{-\lambda_1 t}$, $\lambda_2 e^{-\lambda_2 t}$. The random variable T_i can be interpreted as the time of failure from cause i, if the other cause of failure were inoperative. The observed type and time of failure is determined by the smaller of T_1 and T_2.

The main properties of this model are that:

(i) the probability that a failure is of type I is $\lambda_1/(\lambda_1+\lambda_2)$, independently of the time at which failure occurs;

(ii) the p.d.f. of failure time of all individuals, and of those whose type of failure is given, is $(\lambda_1+\lambda_2)\, e^{-(\lambda_1+\lambda_2)t}$.

Model B. Single risk.—Suppose that individuals are of two types, the chance that an individual is of the first type being θ. An individual of type I is subject only to the risk of failure of the first type and the p.d.f. of failure-time is $h_1(t)=\alpha_1 e^{-\alpha_1 t}$; similarly the p.d.f. of failure-time of the second type is $h_2(t)=\alpha_2 e^{-\alpha_2 t}$. The probability that failure is of the first type given that it occurred at time t is

$$\frac{\theta h_1(t)}{\theta h_1(t)+(1-\theta)\, h_2(t)} \tag{1}$$

and is independent of t if and only if $\alpha_1=\alpha_2$, when the process is completely equivalent to Model A with $\alpha_1=\alpha_2=\lambda_1+\lambda_2$, $\theta=\lambda_1/(\lambda_1+\lambda_2)$.

Model B is the one fitted by Mendenhall and Hader (1958). Non-standard statistical problems arise only if there are individuals who have not failed at the end of the period of observation. Bartlett (1953) has considered a similar model in connection with the estimation from cloud chamber tracks of the mean life-times of unstable particles, when two types of particle are present.

Model B generalizes in a straightforward way when the distributions $h_i(t)$ are not exponential.

We consider now a very general model involving arbitrary distributions.

Model C. General independent risks.—Suppose that the random variables T_1, T_2 of Model A are independently distributed with continuous distribution functions $F_1(t)$, $F_2(t)$. Then the probability, say $g_1(t)\,\delta t + o(\delta t)$, that failure occurs between $(t,\ t + \delta t)$ and is of the first type is given by

$$g_1(t) = F_1'(t)[1 - F_2(t)] \tag{2}$$

and similarly

$$g_2(t) = F_2'(t)[1 - F_1(t)]. \tag{3}$$

Now in this model the probability that a failure is of type I, given that it occurs at t, is equal to $g_1(t)/[g_1(t) + g_2(t)] = \pi_1(t)$, say. This probability does not involve t if and only if

$$\frac{F_1'(t)}{1 - F_1(t)} = \psi \frac{F_2'(t)}{1 - F_2(t)} \tag{4}$$

for some constant ψ; that is, if and only if

$$1 - F_1(t) = [1 - F_2(t)]^{\psi}. \tag{5}$$

The probability that a failure is of type I is $\psi/(1 + \psi)$. Armitage (1959) has used this condition in a study of the comparison of different survivor curves. Equation (4) shows that the conditional failure rates (forces of mortality) of the two types of failure are in constant ratio for all t. Our final model involves this condition.

Model D. Independent proportional risks.—Here we assume the conditions of Model C, plus the requirement that (4) is satisfied.

Thus in the model time of failure and type of failure are statistically independent, and the p.d.f. of time of failure among all individuals, and among individuals whose type of failure is given, is the same, namely $g_1(t) + g_2(t)$.

3. Comparison and Interpretation of Models

Obviously the form of model B in which the $h_i(t)$ are arbitrary, can be used to fit any data of the type considered here, namely data in which only one failure-time is observed for each individual. It is convenient to begin this section by showing that model C also is all-embracing in the same way. For let

$$G(x) = \int_0^x [g_1(t) + g_2(t)]\,dt \tag{6}$$

be the probability of failure before x. Then it follows from (2) and (3), or can be seen directly, that

$$1 - G(x) = [1 - F_1(x)][1 - F_2(x)]. \tag{7}$$

If we substitute in (2) we get that

$$\frac{F_1'(x)}{1 - F_1(x)} = \frac{g_1(x)}{1 - G(x)} \tag{8}$$

from which

$$1 - F_1(x) = \exp\left\{-\int_0^x \frac{dG_1(u)}{1 - G(u)}\right\}, \tag{9}$$

where $G_1(x)$ is the integral of $g_1(x)$. There is a similar equation for $F_2(x)$. The distribution functions $F_i(x)$ determine and are determined by the functions $g_i(t)$.

The equality of the two sides of (8) is seen alternatively by noting that both are equal to the conditional failure rate from failure of type I at age t. In fact (9) is the actuarial life-table analysis of the system. For $dG_1(u)/[1 - G(u)]$ is the conditional failure rate for Type I failures, and equation (9) is that expressing a distribution function in terms of the conditional failure rate.

It follows that no data of the present type can be inconsistent with model C. In order to test the applicability of the model it would be essential to have a different type of data; for example it may be possible to get information about the values of both T_1 and T_2 for the same individual, or to divide the individuals into rational subgroups.

In some applications the single-risk model B can be ruled out on general grounds, it being clear that all individuals are subject to both types of risk. In other cases, as in the application to cloud chamber tracks, there are two distinct types of individual and the single-risk model is the appropriate one.

Model A, with independent Poisson risks, is the simplest, and is of course implied in Markovian birth-death and allied processes. Harris *et al.* (1950), Moran (1953) and Meier (1955) have studied some of the associated estimation problems. The generalization with independent proportional risks, model B, is likely to arise in two main circumstances. First there may be similar processes of wear and ageing underlying the two types of failure. A second, quite different, possibility is that there is "really" only one type of failure, and that when failure occurs, a random event with constant probability $\psi/(1 + \psi)$ determines the type to which the failure is assigned.

Now in Mendenhall and Hader's example, the classification of failures is based on whether or not the failure is confirmed on arrival at the maintenance centre. Various probability models may be worth considering in situations like this, and the choice between them must depend on detailed knowledge not available to me in the present case. The following points may, however, be made:

(a) It seems reasonable to regard each individual as subject to both risks, and hence not to use model B.

(b) The two types of failure may not correspond to physically different processes of wear. A possible model is then the one mentioned in the preceding paragraph, i.e. model D, with $\psi/(1 + \psi)$ interpreted as the probability that when a failure has occurred it is classed as confirmed.

(c) In some circumstances, failures of only one type may arise completely randomly. For example, the unconfirmed failures may have been due to temporary misuse of the equipment. This suggests the use of model C with $F_2(t)$ exponential, and $F_1(t)$ general.

4. Some Statistical Procedures

A first step in the analysis of this sort of data is to examine whether the proportion of failures of Type I varies in time; if it does vary models A and D are excluded. A simple test is to group the times of failure and to test by χ^2 the hypothesis that the probability $\pi_1(t)$ that a failure is of Type I is constant.

A test more sensitive against alternatives that $\pi_1(t)$ changes monotonically in time is as follows; let $t_{11}, \ldots, t_{1r_1}$ be the times of failure of the first type, $t_{21}, \ldots, t_{2r_2}$ those of the second type, individuals that do not fail being ignored. Then under the null hypothesis H_0 that $\pi_1(t)$ is independent of t, $\bar{t}_{1.} = \Sigma\, t_{1i}/r_1$ is the observed value of a random variable T_1, having the distribution of the mean of a sample of size r_1 drawn randomly without replacement from the finite population $\{t_{11}, \ldots, t_{2r_2}\}$ of size $r_1 + r_2$. Thus under H_0

$$E(T_1) = \frac{\sum_{i,j} t_{ij}}{r_1 + r_2} = \bar{t}_{..}, \tag{10}$$

$$V(T_1) = \frac{r_2}{r_1(r_1 + r_2)(r_1 + r_2 - 1)} \sum_{i,j} (t_{ij} - \bar{t}_{..})^2. \tag{11}$$

The random variable T_1 is nearly normally distributed when r_1 and r_2 are large; an exact test can in principle always be made. The test is equivalent to the permutation t test for comparing two sample means and is optimum (Cox, 1958) when the alternative hypothesis is that $\pi_1(t)/[1 - \pi_1(t)] = g_1(t)/g_2(t) = \alpha e^{\beta t}$.

If this test is used it needs to be supplemented by inspection of the data to check that no large changes, not contributing to $T_1 - E(T_1)$, get overlooked.

Table 1

Frequency Distributions of Times of Failure of Radio Transmitters (after Mendenhall and Hader)

(*hr.*)	*Type I* (*confirmed*)	*Type II* (*unconfirmed*)	*Total*	*Fitted frequency* (*exponential*)
		(*Observed frequency*)		
0–	26	15	41	55·9
50–	29	15	44	47·7
100–	28	22	50	40·0
150–	35	13	48	34·3
200–	17	11	28	29·0
250–	21	8	29	24·6
300–	11	7	18	20·9
350–	11	5	16	17·7
400–	12	3	15	15·0
450–	7	4	11	12·7
500–	6	1	7	10·8
550–	9	2	11	9·2
600–629	6	1	7	4·8
Not failed at 630 hr.	—	—	44	46·4
Total	218	107	369	369·0

Example.—Table 1 shows the data of Mendenhall and Hader's example, given in full in their paper, grouped by time of failure. The first test of the constancy of $\pi_1(t)$ is the standard χ^2 test applied to the 12×2 contingency table formed from the first two columns of Table 1. We get $\chi^2_{12} = 9 \cdot 37$. For the more sensitive test based on (10) and (11), we have from the grouped data $\bar{t}_{1.} = 229 \cdot 7$. The mean and standard deviation of this under the null hypothesis are $218 \cdot 5$ and $6 \cdot 18$, so that the departure from expectation is $1 \cdot 81$ times the standard error; the corresponding multiple using the ungrouped data is $2 \cdot 05$. There is thus some evidence, by no means conclusive, that $\pi_1(t)$ increases with time. In fact the proportion of failures of Type I is reasonably constant at about 2/3 up to about 400 hr and then increases to about 4/5.

Since there is some doubt about the reality of this effect it is worth continuing the analysis in two parts, one taking $\pi_1(t)$ constant, the other not.

If we assume $\pi_1(t)$ to be constant, the next step is to test whether the Poisson form, model A, applies, or whether the more general model D with non-exponential distributions must be fitted. Suppose that the failure-times, without regard to type of failure, are $t_1, \ldots, t_r$ and that there are s individuals that have not failed at the end of their periods of observation $t_1', \ldots, t_s'$. If these are sampled froma Poisson process with parameter λ, the maximum likelihood estimate of λ is (Epstein and Sobel, 1953)

$$\hat{\lambda} = \frac{r}{\Sigma\, t_i + \Sigma\, t_j'}. \tag{12}$$

If the frequency distribution is grouped, and censored at or near a fixed value, fitted frequencies can be calculated and a χ^2 goodness of fit test applied.

Example.—The estimate of $\hat{\lambda}$ from the total column of Table 1 is given by $50\hat{\lambda} = 0 \cdot 1646$ and the fitted frequencies are given in the last column. We get $\chi^2_{12} = 16 \cdot 66$, which is well short of the tabulated 10 per cent. point. Note however that, whereas there is excellent agreement with the fitted frequencies above 200 hr, there is a systematic departure up to 200 hr., the departure being of similar form for type I and for Type II failures. The adoption of a finer grouping for the small failure-times does not throw more light on this.

If $s = 0$, so that we have an uncensored sample of failure-times, a test of the exponential form more sensitive against smooth departures can be obtained in a number of ways. One of the simplest, and one that does not depend unduly on the small observations, which are often subject to severe errors of recording, etc., is that based on Sherman's statistic S, the ratio of the mean deviation to the mean. Sherman (1957) has provided tables of its distribution for $r \leqslant 21$; in Sherman's notation $S = 2\tilde{\omega}_n$, $r = n + 1$. In larger samples

$$S = \frac{\sum_{i=1}^{r} |\, t_i - \bar{t}\,|}{r\bar{t}} \tag{13}$$

can be taken approximately normally distributed with mean $e^{-1}[1 - (2r)^{-1}]$ and variance $0 \cdot 05908/r - 0 \cdot 01237/r^2$ (Bartholomew, 1954).

This test, at any rate in its simple form, cannot be used on the example because of the censoring. One way of getting a "smooth test" for this is by postulating a functional form reducing to the exponential as a special case, to estimate parameters by maximum likelihood and to test for departure from exponential form by comparison against a large-sample standard error. The most suitable functional form seems to be the Weibull type, with distribution function

$$1 - \exp\{-(\gamma t)^{1+\beta}\}, \tag{14}$$

the null hypothesis being that $\beta = 0$.

However the test obtained like this is quite cumbersome, even when the approximation that β is small is used. Moreover it depends rather critically on the small observations. We shall therefore not go into details here.

Instead, in the special case when censoring is at a fixed point, $t_j' = t_c$, $t_i < t_c$, we can find a simple test by the method of Cochran (1955). In this a weighted sum of deviations between observed and expected frequencies is compared with its large-sample standard error. Let f_i be the observed frequency and m_i the fitted frequency in the i^{th} group. We need to choose weights g_i determining a test statistic

$$L_g = \Sigma\, g_i(f_i - m_i). \tag{15}$$

Cochran gives the large-sample variance $V(L_g)$ when the g_i are preassigned and also proves that if the g_i are chosen as functions of the sample to maximize $L_g^2/V(L_g)$, then the maximized ratio is equal to the χ^2 goodness-of-fit criterion. Here, however, we want to choose "smooth" weights, and a reasonable procedure is to take a family of weights with a small number of adjustable parameters, chosen to maximize the above ratio; the maximized value will, under the null hypothesis, be distributed as χ_k^2, where k is the dimensionality of the family of weights.

Example.—Number the 14 groups of Table 1 from 1, . . . , 14, counting the group of censored observations as number 14. How shall the deviations be weighted? One might first consider linear weighting $g_i = i$; however the estimating equation for the unknown parameter is such that this weighting amounts very nearly to testing the difference between fitted and observed frequencies for the censored observations. This may indeed be one right thing to test, if, for example, it is required to examine the adequacy of the exponential curve for extrapolation, but it would best be done directly by giving all groups, except the last, zero score. In the example it is clear that good agreement with the null hypothesis would be obtained.

One way of selecting a family of "smooth" weights is to take a general linear combination of the second and third degree orthogonal polynomials for 14 equally-spaced points, $\xi_2' + \lambda\xi_3'$ say. The resulting ratio $L_g^2/V(L_g)$, maximized with respect to λ, can be tested as χ_2^2. It is clear from Table 1 that $\xi_2' + \lambda\xi_3'$ does not correlate well with the deviations and that a significant value of χ^2 will not result. Indeed if one had no prior knowledge of the pattern of deviations expected, a many-parameter family of weights would be needed to produce a large $L_g^2/V(L_g)$, and it is unlikely that very high significance would result. From a practical point of view, it would hardly be justifiable to pursue the point far.

An extremely rough argument is that if we were to fit a k parameter family of weights, which accounted for all the real variation present, the remaining degrees of freedom in the total χ^2 would on the average contribute unity each. Thus with a total $\chi_{12}^2 = 16\cdot 66$, we can expect roughly that if fitting a k parameter family of weights leaves only random variation, the remaining $12 - k$ degree of freedom contribute $12 - k$, thus giving for testing significance $\chi_k^2 = 16\cdot 66 - (12 - k) = 4\cdot 66 + k$. Now at the 5 per cent. level this value of χ_k^2 is significant only if $k = 1$, i.e. only if the weights are preassigned. In general, if we found that $k \leqslant k_0$ for significance, we would have to consider whether the observed pattern of deviations is such as would, on prior grounds, have been included in a k_0 parameter pattern of weights for picking out systematic deviations.

In the present case the observed departures from the fitted distribution have little effect on the mean value but would be important if, for example, one wished to test whether a log-normal distribution is a better fit than an exponential. Bias in recording small values also could produce such a departure from the exponential distribution.

To sum up, there is an interesting suggestion of a departure from the exponential form, in the lower part of the distribution, but this cannot be regarded as firmly established.

Thus Model A cannot definitely be ruled out, even though there is evidence against it on two counts.

In the present example there is no point in trying to fit a more elaborate distribution, but such fitting would be necessary if it were required to estimate the mean of the complete distribution of failure-times. Distributions that might be fitted are

(a) a Weibull distribution, (14);

(b) a Γ distribution;

(c) a log-normal distribution.

The first two, but not the third, reduce to the exponential as a special case, and it is known (Irwin, 1942) that the log-normal and the exponential are hard to distinguish from data. Moreover, a Weibull or Γ distribution with coefficient of variation less than one is very difficult to distinguish from a log-normal distribution. The main differences between the families occur when there is a high proportion of very short failure-times, as in a Weibull distribution, (14), with $\beta < 0$, or in a Γ distribution with infinite ordinate at the origin. Such a distribution is appreciably different from a log-normal distribution, both near the origin and in having a relatively shorter tail. A consequence of the long tail of the log-normal distribution is that the conditional failure-rate, after increasing to a maximum, tends to zero for very long failure-times (Gumbel, 1958). This is often implausible on general grounds. In the example the departure is in the direction of the log-normal. Note that the region of small values which provides much information for discriminating between distributional types is the region of least importance in determining the mean failure-time.

There are sometimes theoretical reasons for expecting (a), which is an extreme-value distribution, but if the object of the analysis is not to discriminate between different stochastic models, but is simply to estimate mean values, it is reasonable to choose, among distributions that fit the data equally well, that for which the parameters are most easily estimated from censored data. Very often we shall then be free to choose between (a), (b), and (c). However, if the object is to examine distributional form, and especially if the conditional failure rate is of prime concern, the log-normal may be excluded on account of the property noted in the last paragraph.

The Weibull distribution and the log-normal are easily fitted graphically to censored data, the Γ distribution less easily. Efficient numerical estimation is much the easiest for the log-normal, since methods developed for the normal distribution (Gupta, 1952;

Sarhan and Greenberg, 1956, 1958; Blom, 1958) can be applied to the log-failure-times. It may be advisable to censor extreme observations in the lower part of the distribution.

Finally, we discuss briefly several things that can be done if the first test shows that $\pi_1(t)$ varies with t.

(a) *Test of the form of one distribution*

It may be suspected that one distribution function, say $F_2(t)$, is exponential, the other, $F_1(t)$, being of unknown form. An exact small-sample test of the hypothesis that $F_2(t)$ is exponential would be of interest; the following approximate procedure is based on a life-table analysis.

Divide the range of failure-times into intervals of width h so short that in any one interval few failures occur and both conditional failure rates are effectively constant. At the start of the i^{th} interval let n_i individuals be in service, and let N_{i0}, N_{i1}, N_{i2} be the numbers surviving the i^{th} interval, and failing from Type I and II failures respectively. Let λ_{i1}, λ_{i2} be the corresponding conditional failure rates.

Then

$$E(N_{i0}) = n_i e^{-(\lambda_{i1}+\lambda_{i2})h} \tag{16}$$

$$E(N_{i2}) = \frac{n_i\,\lambda_{i2}}{\lambda_{i1}+\lambda_{i2}}\,[1 - e^{-(\lambda_{i1}+\lambda_{i2})h}]. \tag{17}$$

Thus

$$h\lambda_{i2} = \frac{E(N_{i2})}{n_i}\left[1 - \frac{E(N_{i0})}{n_i}\right]^{-1}\left\{-\log\frac{E(N_{i0})}{n_i}\right\} \tag{18}$$

$$\simeq \frac{E(N_{i2})}{n_i} + \frac{E(N_{i2})(n_i - E(N_{i0}))}{2\,n_i^2}.$$

This suggests putting

$$h\hat{\lambda}_{i2} = \frac{N_{i2}}{n_i} + \frac{N_{i2}\,(N_{i1}+N_{i2})}{2\,n_i^2}. \tag{19}$$

For some purposes it is more useful to consider

$$\theta_{i2} = 1 - e^{-\lambda_{i2}h} \simeq \lambda_{i2}h - \tfrac{1}{2}\,\lambda_{i2}^2\,h^2,$$

the conditional probability of not surviving the i^{th} interval when only Type II failures operate; we take

$$\hat{\theta}_{i2} = \frac{N_{i2}}{n_i} + \frac{N_{i1}N_{i2}}{2\,n_i^2}. \tag{20}$$

To the order of approximation considered here

$$\hat{\theta}_{i2} = \frac{N_{i2}}{n_i - \frac{1}{2}\,N_{i1}}. \tag{21}$$

Approximately

$$V(\hat{\theta}_{i2}) = \theta_{i2}/n_i. \tag{22}$$

Formulae of this type are widely used in actuarial work; Elveback (1958) has made a thorough theoretical analysis.

We can test whether θ_{i2} is constant by a χ^2 test applied to the estimates $\hat{\theta}_{i2}$, or to the estimates $\hat{\lambda}_{i2}$, if for some reason unequal grouping intervals are used; the single degrees of freedom in χ^2 for linear, quadratic, etc. increase in θ_{i2} can be picked out in the usual way. This procedure applied to the pooled failure-times provides another solution to the problem of testing a single censored sample for agreement with the exponential form. An approximation in the procedure, additional to that in (20) and (22), is that the n_i are regarded as constant, whereas in fact they are random variables. This does not affect the approximations (20) and (22), as is easily seen by considering expectations, first conditionally on n_i, and then as n_i varies.

Example.—There is some doubt whether this analysis is completely appropriate for Mendenhall and Hader's example, because it is not clear that the two types of failure are physically distinct. Hence it may not be right to represent them by independent random variables. However, if it is correct to treat the failures as genuinely distinct, it would, in some applications, be reasonable to see whether the unconfirmed failures occur randomly; the calculations for this are shown in Table 2. The $\hat{\lambda}_{i2}$ are estimated from (19), the interval estimate for the last group being multiplied by 50/30.

Under the null hypothesis that the $\hat{\lambda}_{i2}$ are constant an estimate $\hat{\lambda}_2$ is obtained by weighting the $\hat{\lambda}_{i2}$ and then the estimated standard errors $D(\hat{\lambda}_{i2})$ of the individual $\hat{\lambda}_{i2}$'s are obtained from (22); a special calculation is needed for the last group. The pattern of differences $\hat{\lambda}_{i2} - \hat{\lambda}_2$ is quite systematic, although the total χ^2_{12} is only 9·35. Components of χ^2 for linear and quadratic regression can, if required, be isolated in the usual way by least squares fitting.

TABLE 2

Life-table Analysis to Test whether the Underlying Distribution of Unconfirmed Failure-times is Exponential

Hr.	*Number at risk* n_i	*Type I failures* N_{i1}	*Type II failures* N_{i2}	N_{i2}/n_i	$50\ \hat{\lambda}_{i2}$	$D(50\ \hat{\lambda}_{i2})$	$\hat{\theta}_{i1}$	$\hat{\theta}_{i2}$
0–	369	26	15	0·0406	0·0429	0·0121	0·0719	0·0421
50–	328	29	15	0·0457	0·0488	0·0129	0·0904	0·0478
100–	284	28	22	0·0775	0·0843	0·0138	0·1024	0·0813
150–	234	35	13	0·0556	0·0613	0·0152	0·1537	0·0597
200–	186	17	11	0·0591	0·0636	0·0171	0·0941	0·0618
250–	158	21	8	0·0506	0·0553	0·0185	0·1363	0·0540
300–	129	11	7	0·0543	0·0850	0·0205	0·0876	0·0566
350–	111	11	5	0·0450	0·0483	0·0221	0·1013	0·0473
400–	95	12	3	0·0316	0·0341	0·0239	0·1283	0·0336
450–	80	7	4	0·0500	0·0534	0·0261	0·0897	0·0522
500–	69	6	1	0·0145	0·0162	0·0281	0·0876	0·0151
550–	62	9	2	0·0323	0·0398	0·0296	0·1475	0·0346
600–629 . . .	51	6	1	0·0196	0·0378	0·0421		
				Weighted mean	0·0543			

(b) *Life-table for estimating* $F_1(t)$ *and* $F_2(t)$

We may take the most general model C, with arbitrary distributions $F_i(t)$ and estimate $F_i(t)$ from (9) by replacing $G_i(t)$ and $G(t)$ by the corresponding sample step functions, possibly after grouping. As noted in section 2, this is a life-table method in which a conditional failure rate is calculated for each type of failure and then converted into a distribution function. With ungrouped data this is very nearly the product limit method

(Kaplan and Meier, 1958); with grouped data we may estimate conditional failure probabilities $\hat{\theta}_{i1}$, $\hat{\theta}_{i2}$ using (20), then converting them into survival probabilities by finding, for example,

$$\prod_{j=1}^{i} (1 - \hat{\theta}_{j1}) \tag{23}$$

which is the probability of surviving to the end of the i^{th} interval, when only failures of Type I operate.

(c) *Parametric form for the underlying distributions*

We may assume parametric forms for $F_i(t)$ and then estimate the parameters say by maximum likelihood. Mr. P. Chandler, in some as yet unpublished work, has considered this in connection with an application to strength testing. He assumed that the distributions $F_i(t)$ are logistic and showed that a close approximation to the maximum likelihood estimates can be obtained simply.

(d) *Parametric form for the observed distributions*

A fourth possibility is to fit a simple parametric form to the observed distributions $g_i(t)$ and then to substitute the fitted functions $\hat{g}_i(t)$ into (9) in order to estimate $F_i(t)$.

Comparison of methods (b), (c) and (d) will not be attempted here.

References

Armitage, P. (1959), "The comparison of survival curves", *J. R. Statist. Soc.* A, **122**, 279–300.

Audley, R. J. (1957), "A stochastic description of the learning behaviour of an individual subject", *Q. J. Exptl. Psychol.*, **9**, 12–20.

Bartholomew, D. J. (1954), "Note on the use of Sherman's statistic as a test of randomness", *Biometrika*, **41**, 556–558.

Bartlett, M. S. (1953), "On the statistical estimation of mean life-times", *Phil. Mag.*, **44**, 249–262.

Blom, G. (1958), *Statistical Estimates and Transformed Beta-Variables*. New York: Wiley.

Cochran, W. G. (1955), "A test of a linear function of the deviations between observed and expected numbers", *J. Amer. Statist. Ass.*, **50**, 377–397.

Cox, D. R. (1958), "The regression analysis of binary sequences", *J. R. Statist. Soc.* B, **20**, 215–242.

Elveback, L. (1958), "Estimation of survivorship in chronic disease; the actuarial method", *J. Amer. Statist. Ass.*, **53**, 420–440.

Epstein, B. & Sobel, M. (1953), "Life testing", *J. Amer. Statist. Ass.*, **48**, 486–502.

Gupta, S. (1952), "Estimation of the mean and standard deviation of a normal population from a censored sample", *Biometrika*, **39**, 260–273.

Irwin, J. O. (1942), "The distribution of the logarithm of survival times when the true law is exponential", *J. Hyg., Camb.*, **42**, 328–333.

Gumbel, E. J. (1958), *Statistics of Extreme Values*. New York: Columbia University Press.

Harris, T. E., Meier, P. & Tukey, J. W. (1950), "Timing of the distribution of events between observations", *Human Biology*, **22**, 249–270.

Kaplan, E. L. & Meier, P. (1958), "Non-parametric estimation from incomplete observations", *J. Amer. Statist. Ass.*, **53**, 457–481.

Meier, P. (1955), "Note on estimation in a Markov process with constant transition rates", *Human Biology*, **27**, 121–124.

Mendenhall, W. & Hader, R. J. (1958), "Estimation of parameters of mixed exponentially distributed failure time distributions from censored life test data", *Biometrika*, **45**, 504–520.

Moran, P. A. P. (1951), "Simple estimation methods for evolutive processes", *J. R. Statist. Soc.* B, **13**, 147–150.

Sampford, M. R. (1954), "The estimation of response-time distributions, III", *Biometrics*, **10**, 531–561.

Sarhan, A. E. & Greenberg, B. G. (1956), "Estimation of location and scale parameters by order statistics from singly and doubly censored samples, I", *Ann. Math. Statist.*, **27**, 427–451.

—— —— (1958), "Estimation of location and scale parameters by order statistics from singly and doubly censored samples, II", *Ann. Math. Statist.*, **29**, 79–105.

Seal, H. L. (1954), "The estimation of mortality and other decremental probabilities", *Skand. Akt.*, **37**, 137–162.

Sherman, B. (1957), "Percentiles of the ϖ_n statistic", *Ann. Math. Statist.*, **28**, 259–261.