FIELD TRIALS FOR LIFE ESTIMATION

by

N. K. Chakravarti

Defence Research Laboratory (Stores), Kanpur.

ABSTRACT

Field trials for life estimation differ in essential respects from laboratory trials for the same purpose in that the psychological reactions of the users play an important role, specially when the users stand to gain by tampering with the data. The method of sampling and of estimation have been discussed, both when only complete lives are used and from truncated normal and exponential populations.

Introduction

The method of field-trials is used wherever the reactions of the users play a significant role. It is usually concerned with the physical conditions of a store or of a user as measured by various physiological factors or psychological reactions as elicited by a questionnaire. These trials need careful control regarding strict adherence to laid down procedure and careful measurement or recording of opinions.

The method of field trials for life estimation normally belongs to the same class. In certain special situations such as are being discussed here, it not only requires careful control and recording, but may require careful control of the behaviour, actions and habits of the trial-subjects. This contingency arises when the life of the store to be estimated affects the trial-subjects, in some way, and they stand to gain by tampering with it. The store itself is not condemned, it arouses no psychological aversion, it does not fail in its performance requirements. Still, the facts are distorted to suit the personal ends of the trial-subjects.

As an example, we may consider the estimation of average life of the common clothing items by the soldiers. These items are purchased by a soldier, but as a consideration for the cost he has to incur, he receives a fixed annual clothing allowance. Though not stated anywhere, it is but natural to suppose that the clothing allowance is based on annual clothing requirements by an average soldier. If the estimated lives of the items decrease, the requirement increases, and in all fairness, the clothing allowance should increase too. A deliberate attempt thus may be made to distort facts by willful damage, or neglect or careless use, or even deliberate substitution of new stores by older ones.

Only a very strict watch can detect such facts if they occur, and by prompt detection, reduce the possibility of further attempts at distortion.

Replacement life

Replacement life (L) is defined as follows:

When the total number of users of the store (N) is divided by the replacement life (L), it results in a correct estimate of the total replacement demand, or annual requirements (R). Thus

$$R = N/L.$$
$$\text{or, } L = N/R \text{ years.} \quad \text{.. .. ..} \quad (1)$$

If the scale of issue is 5, *i.e.*, 5 items of the store are used at a time, and if each item lasts an average of 10 years, there is a demand of 5 items in 10 years, or 1 item in 2 years. Thus the annual requirement is $\frac{1}{2}$ item on the average per person. The replacement life as defined is then

$$L = N/R = \frac{1}{1/2} = 2 \text{ years.}$$

as it should be as stated above.

Sampling

A combination of stratified and nested sampling will generally be found appropriate. Thus, in an organisation like the army, the entire organisation may be divided into five strata: (1) infantry, (2) artillery, (3) armoured corps, (4) EME Workshops, and (5) Supply and Ordnance depots. A number of units belonging to each stratum spaced all over India may be chosen, and in each unit, a number of soldiers may be chosen at random.

It is best to choose the number of soldiers from each stratum in proportion to the strength of the stratum. This number may then be distributed in proportion to the strength of the units chosen.

Estimation of life

Suppose the trial to continue for a fixed time. During this time the trial-subjects are to use only trial items. If one item has worn out, a replacement is issued, and observations on the replacement issue also recorded.

It should be possible to mark each trial item with a serial number. In that case, the exact period that an item has been in use can be obtained. Let this period be P_{ijkq} in respect of the *ith* stratum, *jth* unit, kth trial subject and the qth item used by him. Of course, the items that are not yet worn out at the termination of the trial are left out. If the scale of issue, i.e., the number of items in use by a trial subject at a time be λ_{ijk}, then the replacement life

$$L_{ijkq} = \frac{P_{ijkq}}{\lambda_{ijk}} \quad \text{.. .. ..} \quad (2)$$

Generally, however, the scale of issue is fixed. At least, it is so for the army. Hence $\lambda_{ijk} = \lambda$

Taking an average over all the items used by all the trial subjects of the *jth* unit of the *ith* stratum, we have

$$L_{ij\cdot\cdot} = \frac{1}{\lambda . m_{ij}} \sum_{k,q} P_{ijkq} \; ; \quad \text{.. .. ..} \quad (3)$$

Where the total number of trial subjects involved in the *jth* unit of the *ith* stratum is *mi j*.

The variance is

$$\sigma^2_{L_{ij}\,..} = \frac{1}{\lambda^2 . m_{ij}^2} \left\{ \sum_{k,q} p^2_{ijkq} - \frac{(\sum_{k,q} p_{ijkq})^2}{m^2_{ij}} \right\} \quad .. \quad (4)$$

If there are *r* units in a stratum, we have

$$L_{i\,...} = \frac{1}{\lambda r} \sum_{j} \frac{1}{m_{ij}} \sum_{k,q} p_{ijkq}, \text{ and} \quad .. \quad .. \quad .. \quad (5)$$

$$\sigma^2_{L_{i\,..}} = \frac{\sigma^2_{\mu}}{r} + \frac{1}{\lambda r} \sum_{j} \left(\sigma^2_{L_{ij}} \Big/ m_{ij} \right) \quad .. \quad .. \quad (6)$$

where σ^2_{μ} is the variance between the units in the *i*th stratum.

If N_i is the strength of the *i*th stratum, and N is the total strength and there are S strata, then we have

$$L\,.... = \sum_{i} \frac{N_i}{N} . L_i\,... \quad , \text{and} \quad .. \quad .. \quad .. \quad (7)$$

$$\sigma^2_{L\,....} = \frac{1}{S} \sum_{i} \frac{N_i}{N} \sigma^2_{L_{i....}} \quad .. \quad .. \quad .. \quad .. \quad (8)$$

Estimation from a truncated distribution

The fixed time that a trial continues is based generally on considerations of the life of the longest wearing store e.g., if boots are the longest wearing items with an average life of two years, the trial may last about four years. During this period, certain boots will wear out and be replaced, it is even possible that a few first replacements will also wear out by the time the trial terminates. On the other hand, it is quite likely that a few pairs of boots that are issued at the start of the trial are still in use and not worn out by the time the trial terminates. In case the number of the replacements for which the lives are available are less than the number of original issues that are still serviceable at the termination of the trial, it will be advantageous to use the method of estimation applicable to a truncated population, by ignoring the replacement issues.

On the other hand there will be stores like leather laces, socks, etc. that are also in the trial. Many replacements will usually be necessary by the time the trial terminates. Inclusion of replacement issues in the data will certainly give more information, and hence those of the replacements that are still in use at the end of the trial may be ignored. The estimates are then based on complete lives only, and the usual method of analysis applies.

In the former cases, i.e., where estimation from a truncated distribution leads to more precise estimates, we have to assume the form of the distribution. This form may be normal or exponential. Estimation from both types of populations will be considered here. These provide estimates of mean and variance for the second stage units, e.g., an infantry unit.

Truncated normal distribution

There is a considerable amount of literature in respect of estimation from a truncated normal distribution. We shall consider here a simple mode of estimation.

Let us consider a normal distribution with mean m and standard deviation σ truncated at X.

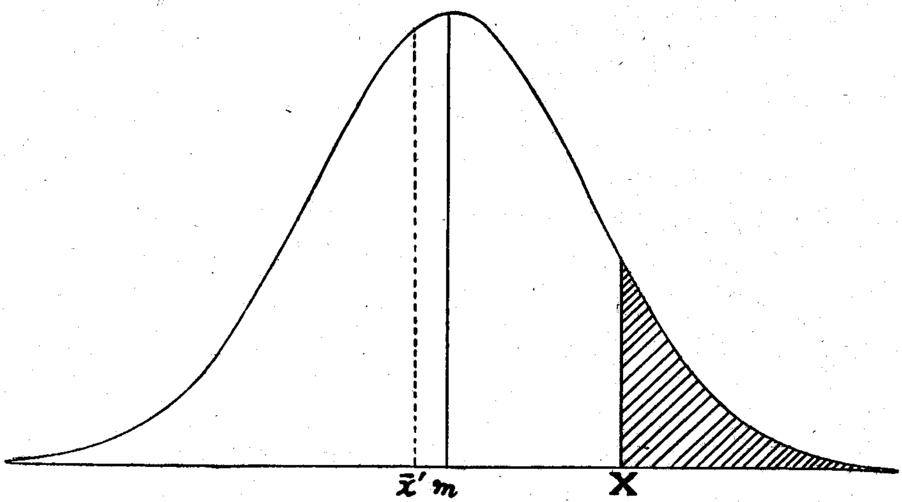

FIG. showing normal distribution.

Let $\bar{x}'$ and S' be the observed mean and standard deviation of a sample from the truncated distribution, truncated at X, and let $\bar{x}$ and S be the corresponding estimates from the sample from the untruncated distribution. Let the total number of items in the sample be N, and the observations be available on n out of N. Hence the area under the distribution from $-\infty$ to X is n/N and the area beyond X is $(N\text{-}n)/N$. Then we have

$$\bar{x}' = \int_{-\infty}^{X} \frac{x}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} dx \Bigg/ \int_{-\infty}^{X} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} dx$$

$$= \int_{-\infty}^{X} \frac{(x-m+m)}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} dx \Bigg/ \frac{n}{N}$$

$$= \int_{-\infty}^{X} \left[\frac{\sigma}{\sqrt{2\pi}} \left(\frac{x-m}{\sigma} \right) + \frac{m}{\sqrt{2\pi}} \right] e^{-\frac{(x-m)^2}{2\sigma^2}} d\left(\frac{x-m}{\sigma} \right) \bigg/ \frac{n}{N} .$$

$$= \left[m \int_{-\infty}^{X} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2} d\left(\frac{x-m}{\sigma} \right) + \right.$$

$$\left. \frac{\sigma}{2\sqrt{2\pi}} \int_{-\infty}^{X} e^{-\frac{(x-m)^2}{2\sigma^2}} d\left(\frac{x-m}{\sigma} \right)^2 \right] \bigg/ \frac{n}{N}$$

$$= \left(\frac{m\,n}{N} - \sigma\, y_T \right) \bigg/ \frac{n}{N}$$

$$= m - \frac{N}{n}\, \sigma\, y_T$$

$$\therefore m = \bar{x}' + \frac{N}{n}\, \sigma\, y_T \qquad \ldots \quad \ldots \quad \ldots \quad \ldots \qquad (9)$$

where T is the point on the standard normal curve corresponding to X on the actual curve, and y_T is the ordinate of the standard normal curve at the point T

Therefore, an estimate of the mean m for the untruncated population is given by

$$\bar{x} = \bar{x}' + \frac{N}{n}\, \sigma\, y_T \qquad \ldots \quad \ldots \quad \ldots \quad \ldots \qquad (10)$$

If σ is unknown, an estimate S for σ is obtained from the truncated sample which is then used in (10).

To obtain an estimate of σ, we proceed as follows:

$$\mu'_2 = \int_{-\infty}^{X} x^2 \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} dx \bigg/ \int_{-\infty}^{X} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} dx$$

$$= \int_{-\infty}^{X} \frac{(x-m+m)^2}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-m)^2}{\sigma}} d(x-m) \bigg/ \frac{n}{N} .$$

$$= \int_{-\infty}^{X} \left[\frac{\sigma^2}{\sqrt{2\pi}} \left(\frac{x-m}{\sigma} \right)^2 + \frac{2\,m\sigma}{\sqrt{2\pi}} \left(\frac{x-m}{\sigma} \right) + \frac{m^2}{\sqrt{2\pi}} \right]$$

$$e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2} d\left(\frac{x-m}{\sigma} \right) \bigg/ \frac{n}{N} .$$

$$= \left[\sigma^2 \left(\frac{n}{N} - T y_T \right) - 2 m\sigma y_T + m^2 \frac{n}{N} \right] \frac{N}{n}$$

$$= \sigma^2 \left(1 - \frac{N}{n} T y_T \right) - 2 m \sigma \frac{N}{n} y_T + m^2. \qquad (11)$$

If S' is the observed standard deviation for the truncated sample, and σ' is its population value, we have

$$\sigma'^2 = \mu'_2 - \mu'^2_1$$

$$= \sigma^2 \left(1 - \frac{N}{n} T y_T \right) - 2 m\sigma \frac{N}{n} T y_T + m^2 - \left(m - \frac{N}{n} \sigma y_T \right)^2$$

$$= \sigma^2 \left(1 - T y_T \frac{N}{n} - \frac{N^2}{n^2} y_T^2 \right)$$

Hence the estimate S' for σ' can be written as

$$S'^2 = S^2 \left(1 - T y_T \frac{N}{n} - \frac{N^2}{n^2} y_T^2 \right) \qquad (12)$$

$$\text{or } S^2 = S'^2 \Big| \left(1 - \frac{N}{n} T y_T - \frac{N^2}{n^2} y_T^2 \right). \qquad (13)$$

S determined from equation (13) can now be put in eqn (10) to obtain $\bar{x}$.

An example

Let $N = 100$; $n = 85$.

$\bar{x}' = 58$; $S'^2 = 148$.

The area to the right of X, the point of truncation is $N - n/N = 0{\cdot}15$.

Hence $T = 1{\cdot}036433$; $y_T = 0{\cdot}2332$.

Using these values in eqn (13), we get

$S^2 = 230{\cdot}89$.

Using this value of S in eqn (10), we have

$\bar{x} = 62{\cdot}17$.

Since it is apparent that the estimates $\bar{x}$ and S in respect of the untruncated population should be higher than $\bar{x}'$ and S' corresponding to the truncated population, the values appear reasonable.

Truncated exponential distribution

Let the exponential distribution be

$$df = \frac{1}{\lambda} e^{-\frac{x}{\lambda}} dx, \qquad x > o$$

It is easy to see that

$$E(x) = \lambda \quad \ldots \quad (14)$$

Using the same notation as in the case of the truncated normal distribution, we find the maximum likelihood estimate for λ to be[1]

$$\hat{\lambda} = \frac{x_1 + x_2 + \ldots\ldots\ldots\ldots + x_n + (N - n)X}{n}$$

If $x_n = X$, i.e., the trial is terminated as soon as the nth item fails, we have

$$\hat{\lambda} = \frac{\sum^{n-1} x_i + (N - n + 1)\, x_n}{n} \quad \ldots$$

The maximum likelihood estimate of variance is given by[2]

$$V\left(\frac{1}{\hat{\lambda}}\right) = \frac{\frac{1}{\hat{\lambda}^2}}{N\left(1 - e^{x/\lambda}\right)} \quad \ldots \quad (16)$$

$$\text{Now } d\left(\frac{1}{\lambda}\right) = -\frac{1}{\lambda^2} d\lambda$$

$$\therefore V\left(\frac{1}{\lambda}\right) = \frac{1}{\lambda^4} V(\lambda)$$

$$\text{Hence } V(\hat{\lambda}) = \frac{\hat{\lambda}^2}{N\left(1 - e^{x/\hat{\lambda}}\right)} \quad \ldots \quad (17)$$

The chance that x exceeds X is given by

$$\int_X^\infty \frac{1}{\lambda} e^{-\frac{x}{\lambda}} dx = e^{-\frac{X}{\lambda}} = \frac{N - n}{N} \quad \ldots \quad (18)$$

Using the estimate $\hat{\lambda}$ for λ, the estimated value of X can be determined and compared with the actual truncation point.

An example

As before, let

$N = 100;$ $n = 85,$

$\bar{x}' = 58;$ $S'^2 = 148.$

$$\therefore e^{-\frac{X}{\lambda}} = 0 \cdot 15$$

Or, $X = 1 \cdot 897\, \lambda$.

Substituting in eqn (15), we get

$$\hat{\lambda} = \frac{\sum^{n} x_i + (N - n) X}{n}$$

$$= 58 + \frac{15 \times 1 \cdot 897\, \hat{\lambda}}{85}$$

$$\therefore \hat{\lambda} = 89 \cdot 15$$

using the eqn (17) for estimating $V(\hat{\lambda})$ we have

$$V(\hat{\lambda}) = \frac{\hat{\lambda}^2}{N\left(1 - e^{-\frac{X}{\hat{\lambda}}}\right)} = \frac{89 \cdot 15^2}{100\,(1 - 0 \cdot 15)}$$

$$= 93 \cdot 5$$

This estimate of $\hat{\lambda}$ is very high compared to $\bar{x}'$ of 58. $V(\hat{\lambda})$ is 93 5 and much less than S'^2, whereas it is to be expected that this value should be higher than 148. It is, therefore, clear that the exponential distribution does not provide a good description of the data in this case.

Discussion

In field-testing, the user element is present, and this case differs in essential respects from that in which the user effect is absent, e.g., in the laboratory testing for estimating the life of electric glow lamps. In such cases, since a user generally uses many other stores, some of which are complementary and others competitive, the life cannot be stated unrelated to the other stores in use. Thus, the life of a pair of socks is higher if used along with a pair of new boots than with a pair with cracked insoles. It is essential, therefore, that the other stores be standardised.

It is of course necessary that the trial stores be homogeneous. In peace time, under normal conditions, there is not a great heterogeneity in the material. However, under special conditions, e.g., when cottage industry is deliberately encouraged and thereby heterogeneity is introduced, a separate trial to estimate the lives of such stores is desirable.

It may be observed that in a trial of the type envisaged here, even though the material may not be heterogeneous, the variations are large as a result of differing nature of duties of the users. This fact necessitates a large sample size. It is also to be observed that in clothing stores, another factor responsible for large variation is washing of the stores whether dhobi or self-washed. However, the practice in a unit being more or less uniform, this factor is not important.

For items that have a short life expectation, the exponential form of the distribution may be appropriate. However, in their case, generally lives of many replacements are available and hence it is not necessary to assume a truncated distribution. In long wearing items, since variations in the nature of duties of users are superimposed, it is likely that normal form of the distribution will be more appropriate.

It may be noted that the users generally keep one or two items in reserve to be used whenever the occasion demands. If these items that are kept in reserve could be identified, it would be possible to weight the lives of the items in terms of the length of use, and obtain a more precise estimate. However, this information is rarely available. Hence the assumption that the scale of use equals the scale of issue is the only possible course open, and it does not vitiate the results.

The life as determined here of course depends on the scale of issue, However, as long as the scale of issue is fixed, the estimates are valid. It is to be observed that in an organisation like the army, the scale of issue is seldom changed.

Acknowledgements

The author wishes to express his appreciation of the valuable assistance rendered by Messrs G.K. Mathur, A.S.P. Rao, B.S. Gill, Ramesh Shanker and A.K. Singh. Thanks are also due to Dr. T.S. Subramanian and Dr. Kartar Singh, for their keen interest in this work.

References

1. Epstin, B., and Sobel, M., *J. Amer. Stat. Assoc.* **48,** 486, 1953.
2. Miller, R.G. Jr., *J. Amer. Stat. Assoc.* **55,** 491, 1960.