Scand J Statist 12: 55–61, 1985

The Null Distribution for a Test of Constant versus "Bathtub" Failure Rate

MAGNE VOLLAN AARSET
University of Trondheim

ABSTRACT. The total time on test plot (TTT plot) was introduced by Barlow & Campo (1975) as a tool for analysing failure data. Bergman (1979) suggested a test statistic based on the TTT plot for testing exponentiality against "bathtub" distributions. In this paper the exact distribution of Bergman's test statistic is derived under the hypothesis of exponentiality.

Keywords: lifetime data, failure rate, "bathtub" distributions, total time on test.

1. Introduction

Let $T_1, T_2, \ldots, T_n$ be a random sample from a life distribution $F\{F(0^-)=0\}$ with finite mean, and denote the ordered sample $T_{n:1}, T_{n:2}, \ldots, T_{n:n}$.

Define

$$U_r=\frac{\sum_{j=1}^{r} T_{n:j}+(n-r)T_{n:r}}{\sum_{j=1}^{n} T_j}; \qquad r=1,2,\ldots,n,$$

and $U_0=0$. The plot of $(r/n,\ U_r)$ $(r=0, 1, 2, \ldots, n)$, where consecutive points are connected by straight lines, is known as the (empirical) TTT plot (Barlow & Campo, 1975).

The TTT transform H_F^{-1} of a life distribution F is defined by

$$H_F^{-1}(t)=\int_0^{F^{-1}(t)} \{1-F(u)\}\,du; \qquad 0\leqq t\leqq 1,$$

where $F^{-1}(t)=\inf\{u:F(u)\geqq t\}$. The scaled TTT transform is defined by

$$\phi_F(t)=\frac{H_F^{-1}(t)}{H_F^{-1}(1)}; \qquad 0\leqq t\leqq 1$$

(Barlow *et al.*, 1972, p. 235).

The scaled TTT transform and its empirical counterpart, the TTT plot, were presented by Barlow & Campo (1975) as a tool for model identification. Some properties of the (scaled) TTT transform are listed by Bergman (1979).

2. The exact null distribution of a test statistic proposed by Bergman

A class of life distributions arising naturally in reliability situations may be constructed by assuming a failure rate initially decreasing during the infant mortality phase, next constant during the so-called "useful life" phase and, finally, increasing during the so-called "wear-out" phase. In reliability literature such failure rate functions are said to have a "bathtub" shape (Barlow & Proschan, 1981, p. 55). The scaled TTT transforms of some "bathtub" distributions are displayed in Fig. 1.

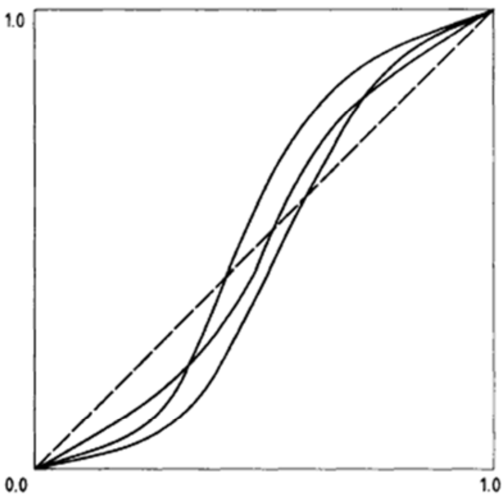

Fig. 1. Scaled TTT transform from an exponential distribution (dotted diagonal) and some "bathtub" distributions

Bergman (1979) suggests the following test procedure for testing exponentiality against the class of distributions with "bathtub"-shaped failure rate. Introduce

$$V_n = \min\left\{i \geqq 1 : U_i \geqq \frac{i}{n}\right\},$$

$$M_n = \max\left\{i \leqq n-1 : U_i \leqq \frac{i}{n}\right\},$$

$$G_n = V_n + n - M_n,$$

and reject the hypothesis of exponentiality when G_n is large. The motivation for this test is that when the distribution has a "bathtub"-shaped failure rate, then we may expect V_n as well as $(n-M_n)$ to be large. G_n obviously takes integer values in $[2, n+1]$, only.

We shall now derive expressions for $P(G_n=i)$ $(i=2, \ldots, n+1)$ under the hypothesis of exponentiality. Then $(U_1, \ldots, U_{n-1})$ are distributed as the order statistic of $(n-1)$ independent observations from a uniform distribution over $(0, 1)$ (Barlow *et al.*, 1972, p. 268). Following Bergman (1979), we may find the null distribution of G_n by considering the experiment of randomly placing $(n-1)$ balls (the U_i's) in n cells (the intervals $(\{j-1\}/n, j/n)$; $j=1, \ldots, n$) and counting the number of balls in each cell.

Let us first consider the events $\{G_n=i\}$ where $i=2, \ldots, n-1$. Observe that $\{G_n=i\}$ may be written as a union of disjoint events in the following manner:

$$\{G_n=i\} = \bigcup_{v=1}^{i-1} (\{V_n=v\} \cap \{G_n=i\}); \qquad i=2, \ldots, n-1.$$

Hence

$$P(G_n=i) = \sum_{v=1}^{i-1} P(\{V_n=v\} \cap \{G_n=i\}); \qquad i=2, \ldots, n-1.$$

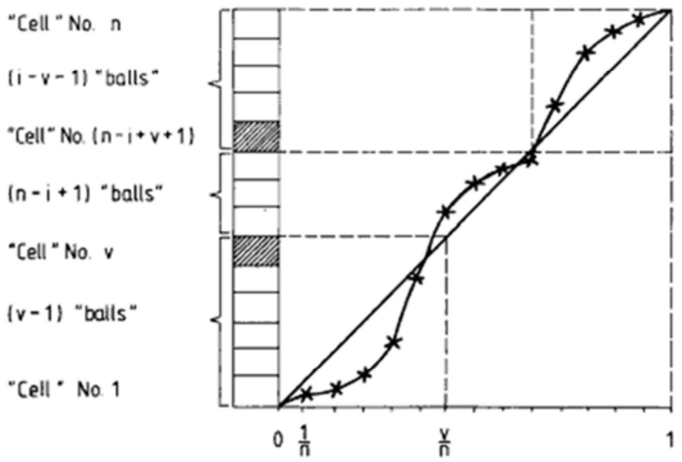

Fig. 2. TTT plot where $\{V_n=v\}\cap\{G_n=i\}$.

As illustrated in Fig. 2, the null probability of the event $\{V_n=v\}\cap\{G_n=i\}$ is that of the event

$$\begin{aligned}
&\{(v-1)\text{ "balls" in the } v \text{ first "cells"}\}\\
&\cap\{(n-i+1)\text{ "balls" in the } (n-i) \text{ next "cells"}\}\\
&\cap\{(i-v-1)\text{ "balls" in the } (i-v) \text{ last "cells"}\}\\
&\cap\{\text{the } (v-1) \text{ first plotted values under the diagonal}\}\\
&\cap\{\text{the } (i-v-1) \text{ last plotted values over the diagonal}\}.
\end{aligned} \tag{2.1}$$

Observe that "cell" no. v and "cell" no. $(n-i+v+1)$ will both be empty. We shall now derive $P(\{V_n=v\}\cap\{G_n=i\})$ using (2.1). Let us determine the probability of simultaneously having $(v-1)$ "balls" in the v first "cells", $(n-i+1)$ "balls" in the next $(n-i)$ "cells" and $(i-v-1)$ "balls" in the last $(i-v)$ "cells". This is a specific result of $(n-1)$ trinomial trials, and the probability of the mentioned result is accordingly:

$$\frac{(n-1)!}{(v-1)!(n-i+1)!(i-v-1)!}\left(\frac{v}{n}\right)^{v-1}\left(\frac{n-i}{n}\right)^{n-i+1}\left(\frac{i-v}{n}\right)^{i-v-1}.$$

Suppose we know that the "balls" are distributed in the n "cells" as described in (2.1). This corresponds to

$$U_{v-1}<v/n,\ U_v>v/n,\ U_{n-i+v}<(n-i+v)/n$$

and (2.2)

$$U_{n-i+v+1}>(n-i+v)/n.$$

When $(U_1,\ldots,U_{n-1})$ is distributed as the order statistic of $(n-1)$ independent observations from a uniform distribution over (0, 1), then $(U_1,\ldots,U_{v-1})$, given that $U_{v-1}<v/n$ and $U_v>v/n$, will be distributed as the order statistic of $(v-1)$ independent observations from a uniform distribution over $(0, v/n)$. Correspondingly, $(U_{n-i+v+1},\ldots,U_{n-1})$, given that $U_{n-i+v}<(n-i+v)/n$ and $U_{n-i+v+1}>(n-i+v)/n$, will be distributed as the order statistic of $(i-v-1)$ independent observations from a uniform distribution over $(\{n-i+v\}/n, 1)$.

Furthermore, given (2.2), $(U_1, \ldots, U_{v-1})$ and $(U_{n-i+v+1}, \ldots, U_{n-1})$ will be independently distributed.

Therefore, given that the "balls" are distributed as described in (2.1), the events {the $(v-1)$ first plotted values under the diagonal} and the {the $(i-v-1)$ last plotted values over the diagonal} will be independent. Furthermore, according to Barlow & Campo (1975), these events will occur with probability $1/v$ and $1/(i-v)$ respectively. Hence

$$P(\{V_n=v\} \cap \{G_n=i\}) = \frac{(n-1)!}{(v-1)!(n-i+1)!(i-v-1)!}\left(\frac{v}{n}\right)^{v-1}\left(\frac{n-i}{n}\right)^{n-i+1}\left(\frac{i-v}{n}\right)^{i-v-1}$$

$$\times\frac{1}{v}\frac{1}{i-v}; \qquad i=2, \ldots, n-1,\ v=1, \ldots, i-1,$$

and

$$P(G_n=i) = \sum_{v=1}^{i-1}\frac{(n-1)!}{(v-1)!(n-i+1)!(i-v-1)!}\left(\frac{v}{n}\right)^{v-1}\left(\frac{n-i}{n}\right)^{n-i+1}\left(\frac{i-v}{n}\right)^{i-v-1}$$

$$\times\frac{1}{v}\frac{1}{i-v}; \qquad i=2, \ldots, n-1. \tag{2.3}$$

Now let us consider the event $\{G_n=n\}$. This event is equivalent to the event $\{V_n=M_n\}$, which occurs if and only if the TTT plot crosses the diagonal only once and from below in one of the points $(j/n, j/n)(j=1, \ldots, n-1)$. This event has probability 0, hence

$$P(G_n=n)=0.$$

Finally let us consider the event $\{G_n=n+1\}$. In connection with the evaluation of $P(G_n=n+1)$, we shall use some definitions given by Bergman (1977). The TTT plot crosses the diagonal from below in the interval $(i/n, (i+1)/n]$ $(i=1, \ldots, n-1)$ if $U_i<i/n$ and $U_{i+1} \geqq (i+1)/n$.

Similarly a crossing from above occurs in the interval $(i/n, (i+1)/n]$ $(i=1, \ldots, n-2)$ if $U_i>i/n$ and $U_{i+1} \leqq (i+1)/n$.

A crossing is either a crossing from below occurring in $(1/n, (n-1)/n]$ or a crossing from above. Observe that, by definition, a crossing from below in the interval $(1-1/n, 1]$ is not a crossing. The definition of crossing from below, which allows a crossing from below to take place in 1, is chosen for technical reasons.

Let K_n be the number of crossings from below and let L_n be the number of crossings. Then we have the basic relation

$$L_n=(1-\delta_n)+2(K_n-\Delta_n)+(\Delta_n-1),$$

where

$$\delta_n=\begin{cases}1 & \text{if the point } (1/n,\ U_1) \text{ lies below the } 45° \text{ line } (U_1<1/n)\\ 0 & \text{otherwise}\end{cases}$$

and

$$\Delta_n=\begin{cases}1 & \text{if the point } (\{n-1\}/n,\ U_{n-1}) \text{ lies below the } 45° \text{ line } \{U_{n-1}<(n-1)/n\}\\ 0 & \text{otherwise.}\end{cases}$$

This relation is easily proved by consideration of the four possible ways for the TTT plot to start and end with regard to the diagonal.

The event $\{G_n=n+1\}$ can occur in three disjoint ways only. Either the TTT plot lies altogether above the diagonal, the TTT plot lies altogether below the diagonal or the TTT plot has only one crossing of the diagonal and this crossing is from below in the interval $(1/n, \{n-1\}/n)$. The probability of each of the first two elements are by Barlow & Campo (1975) found to be $1/n$. The third possibility may equivalently be expressed as $\{L_n=1,\ \delta_n=1\}$.

Now let us derive $P(L_n=1,\ \delta_n=1)$. We have

$$\begin{aligned}P(L_n=1,\ \delta_n=1)&=P(L_n=1)-P(K_n=1,\ \delta_n=0,\ \Delta_n=1)\\&=P(L_n=1)-P(\Delta_n=1)\ P(K_n=1,\ \delta_n=0\mid\Delta_n=1).\end{aligned}$$

Barlow & Campo (1975) found $P(\Delta_n=1)=\{(n-1)/n\}^{n-1}$.

The other probabilities are (Bergman, 1977):

$$P(L_n-1)=P(L_n\geqq 1)-P(L_n\geqq 2)=\frac{2}{n}\left(\frac{n+1}{n}\right)^{n-2}-\frac{2}{n^2}-\frac{1}{n}$$

and

$$\begin{aligned}P(K_n=1,\ \delta_n=0\mid\Delta_n=1)&=P(K_n\geqq 1,\ \delta_n=0\mid\Delta_n=1)-P(K_n\geqq 2,\ \delta_n=0\mid\Delta_n=1)\\&=\frac{(n-2)n^{n-3}}{(n-1)^{n-1}}\end{aligned}$$

Hence

$$\begin{aligned}P(L_n=1,\delta_n=1)&=\frac{2}{n}\left(\frac{n+1}{n}\right)^{n-2}-\frac{2}{n^2}-\frac{1}{n}-\left(\frac{n-1}{n}\right)^{n-1}\frac{(n-2)n^{n-3}}{(n-1)^{n-1}}\\&=\frac{2}{n}\left\{\left(\frac{n+1}{n}\right)^{n-2}-1\right\},\end{aligned}$$

and finally

$$P(G_n=n+1)=\frac{2}{n}\left\{\left(\frac{n+1}{n}\right)^{n-2}-1\right\}+\frac{2}{n}=\frac{2}{n}\left(\frac{n+1}{n}\right)^{n-2}$$

Hence the distribution of G_n is completely known under the hypothesis of exponentiality.

In Table 1 numerical values of $P(G_n\geqq n-k)$ are given for selected values of k and n.

Table 1. $P(G_n\geqq n-k)$

	n								
k	10	50	75	100	125	150	175	200	250
−1	0.42872	0.10348	0.07013	0.05303	0.04263	0.03565	0.03063	0.02685	0.02153
1	0.46698	0.11091	0.07506	0.05673	0.04559	0.03811	0.03273	0.02869	0.02300
2	0.51102	0.11843	0.08001	0.06041	0.04852	0.04055	0.03482	0.03051	0.02446
3	0.56003	0.12564	0.08470	0.06389	0.05129	0.04284	0.03678	0.03222	0.02582
4	0.61577	0.13257	0.08917	0.06718	0.05389	0.04499	0.03862	0.03383	0.02710
5	0.68139	0.13927	0.09343	0.07030	0.05636	0.04703	0.04035	0.03534	0.02830
6	0.76202	0.14579	0.09753	0.07329	0.05871	0.04897	0.04200	0.03677	0.02944
7	0.86578	0.15217	0.10149	0.07617	0.06097	0.05082	0.04358	0.03814	0.03052
8	1.00000	0.15846	0.10535	0.07895	0.06314	0.05261	0.04509	0.03945	0.03156

Table 2. $P(G_n=n-k)$, exact and approximated

	n			
	100		200	
k	Exact	Approx.	Exact	Approx.
−1	0.05303	0.05437	0.02685	0.02718
1	0.00370	0.00368	0.00184	0.00184
2	0.00368	0.00361	0.00132	0.00180
3	0.00348	0.00336	0.00171	0.00168
4	0.00329	0.00313	0.00160	0.00156
5	0.00312	0.00292	0.00151	0.00146

3. Asymptotic results

From (2.3) we get

$$\lim_{n\to\infty} P(G_n=i)=\sum_{v=1}^{i-1}\frac{v^{v-1}}{i-v}\,\frac{(i-v)^{i-v}}{v!(i-v)!}\exp(-i); \qquad i=2, 3, \ldots$$

The right-hand side of this formula seems to give good approximations of the exact distribution of G_n for small values of i. However, using G_n as a test statistic, the most interesting part of G_n's distribution is $P(G_n \geqq n-k)\,(k=-1, 1, 2, \ldots)$. From (2.3) we get

$$P(G_n=n-k)=\frac{n!}{n^n}\,\frac{k^{k+1}}{(k+1)!}\sum_{v=1}^{n-k-1}\frac{v^{v-1}}{v!}\,\frac{(n-k-v)^{n-k-v-1}}{(n-k-v)!}; \qquad k=1, 2, \ldots$$

Using Stirling's formula and the dominated convergence theorem, we get

$$\lim_{n\to\infty} nP(G_n=n-k)=2\frac{k^{k+1}}{(k+1)!}\exp(-k)\sum_{v=1}^{\infty}\frac{v^{v-1}}{v!}\exp(-v); \qquad k=1, 2, \ldots$$

Bergman (1979) shows that

$$\sum_{v=1}^{\infty}\frac{v^{v-1}}{v!}\exp(-v)=1.$$

Hence

$$\lim_{n\to\infty} nP(G_n=n-k)=2\frac{k^{k+1}}{(k+1)!}\exp(-k); \qquad k=1, 2, \ldots$$

We see that this formula is valid even for $k=-1, 0$.

In Table 2 the probabilities $P(G_n=n-k)$ are given both exact and approximated for selected values of k and n.

Acknowledgements

The author would like to thank A. Høyland and B. Lindqvist for their helpful advice and suggestions.

References

Barlow, R. E., Bartholomew, D. J., Bremner, J. & Brunk, H. D. (1972). *Statistical inference under order restrictions*. Wiley, New York.

Barlow, R. E. & Campo, R. (1975). Total time on test processes and applications to failure data analysis. In *Reliability and fault tree analysis* (ed. Barlow, Fussel & Singpurwalla). SIAM, Philadelphia.

Barlow, R. E. & Proschan, F. (1981). *Statistical theory of reliability and life testing, probability models*. Holt, Rinehart & Winston, New York.

Bergman, B. (1977). Crossings in the total time on test plot. *Scand. J. Statist.* **4,** 171–177.

Bergman, B. (1979). On age replacement and the total time on test concept. *Scand. J. Statist.* **6,** 161–168.

Received May 1983, in final form October 1984

Magne Vollan Aarset, Storebrand-Norden, H10-5, Box 1380, Vika, N-0114 Oslo 1, Norway