⊠연구논문

평균잔여수명의 경향 변화에 대한 검정에 관한 연구

나명환 · 이현우 · 김재주
서울대학교 통계학과

A Study on the Test for Trend Change in Mean Residual Life

Myoung-Hwan Na · Hyunwoo Lee · Jae-Joo Kim
Dept. of Statistics, Seoul National University

Abstract

The mean residual life function is the expected remaining life of an item at age x. The problem of trend change in the mean residual life is great interest in the reliability and survival analysis. In this paper we develop a new test statistic for testing whether or not the mean residual life changes its trend based on a complete sample. Monte Carlo simulations are conducted to compare the performance of our test statistic with that of previously known tests.

1. Introduction

Let F be a continuous life distribution(i.e., $F(x)=0$ for $x\leq 0$) with the finite first moment and let X be a nonnegative random variable with distribution F. The mean residual life(MRL) function $e(x)$ is defined as

$$e(x)=E(X-x \mid X>x), \tag{1.1}$$

The MRL is the expected remaining lifetime, $X-x$, given that the item has survived to time x. The MRL function $e(x)$ in (1.1) can also be written as

$$e(x)=\frac{\int_x^\infty \overline{F}(u)\,du}{\overline{F}(x)},$$

where $\overline{F}(x)=1-F(x)$ is the reliability function.

The MRL function plays a very important role in the area of engineering, medical science, survival studies, social sciences, and many other fields. The MRL is used by engineers in burn-in studies, setting maintenance policies, and in comparison of life distributions of different systems. Social scientists use MRL, also called as inertia, in studies of lengths of wars, duration of strikes, job mobility etc. Medical researchers use MRL in lifetime experiments under various conditions. Actuaries apply MRL to setting rates and benefits for life insurance.

Hall and Wellner(1981) derive that all MRL functions associated with distributions having a finite mean must satisfy three conditions:

$$e(x)\geq 0, \quad e'(x)\geq -1, \quad \int_0^\infty \frac{1}{e(x)}\,dx=\infty.$$

See also Bhattacharjee(1982) for another characterization of MRL. Knowledge of the MRL function completely determines the reliability function, via the relation

$$\overline{F}(x)=\frac{e(0)\exp\left\{-\int_0^x [e(u)]^{-1}du\right\}}{e(x)}, \quad x\geq 0. \tag{1.2}$$

Cox(1962) assigns as an exercise the demonstration that MRL determines the reliability. Kotz and Shanbhag(1980) derive a generalized inversion formula for distributions that are not necessarily life distributions. Hall and Wellner(1981) have an excellent discussion of (1.2) along with further references.

Guess and Proschan(1988) show that various families of life distributions defined in terms of the MRL(e. g. increasing MRL, decreasing MRL) have been used as models for lifetimes for which such prior information is available. One such family of distributions is called as "increasing initially then decreasing MRL(IDMRL)" distributions if there exists a turning point $\tau\geq 0$ such that

$$e(s) \le e(t) \quad \text{for} \ 0 \le s \le t < \tau, \quad e(s) \ge e(t) \quad \text{for} \ \tau \le s \le t. \tag{1.3}$$

The dual class of "decreasing initially, then increasing MRL(DIMRL)" distributions is obtained by reversing inequalities on the MRL function in (1.3). IDMRL distributions model life in which, in terms of residual life, aging initially is beneficial but eventually is detrimental. Such life times are exemplified. (i) Human lifetimes; High infant mortality causes the initially IMRL and deterioration with advancing age causes the subsequently DMRL. (ii) Employment time with a given company; The remaining employment time (residual life) of an employee with several years with a company is likely (due to time investment, career value, etc.) to exceed that of an employee with the company only several months. This results in increasing MRL with years of employment up to a certain point τ, after which, due to retirement, MRL decreases. Also see Guess and Proschan(1988) and the references therein for further applications of the IDMRL family.

It is well known that $e(x)$ is constant for all $x \ge 0$ if and only if F is an exponential distribution (i.e., $F(x) = 1 - \exp(-x/\mu)$ for $x \ge 0$, $\mu > 0$). Due to this "no-aging" property of the exponential distribution, it is of practical interest to know whether a given life distribution F is constant MRL or IDMRL. Therefore, we consider the problem of testing

H_0 : F is constant MRL,

against

H_1 : F is IDMRL (and not constant MRL),

based on random samples. When the dual model is proposed, we test H_0 against

H_1' : F is DIMRL (and not constant MRL).

This problem is noted by Guess, et al.(henceforth GHP, 1986), who obtain tests when the change point is known or when the proportion before the change point takes place is known. Aly(1990) also discusses this problem. Both of GHP(1986) and Aly(1990) tests are based an estimates of functional which distinguish that F is constant MRL against that F is IDMRL(DIMRL).

In Section 2 we develop a test statistic for testing exponentiality against IDMRL (DIMRL) alternative. We assume that the turning point is known. GHP(1986) provide excellent explaining that this assumption is very realistic in many interesting situations. To establish the asymptotic distribution of our test statistic, we used the differential statistical function approach. In Section 3 Monte Carlo simulations are conducted to compare the performance of our test statistic with the GHP's(1986) test and Aly's(1990) test for various values of sample size n.

2. Test for Trend Change in MRL

In this section we propose a test statistic for testing exponentiality against IDMRL(DIMRL) alternative. We assume that the turning point τ is known or has been specified by the user. Our test statistic is motivated by a simple observation of Ahmad(1992). If $e(x)$ is differentiable and $e(x)$ is decreasing(increasing), then

$$\frac{de(x)}{dx} = \frac{f(x)v(x) - \overline{F}^2(x)}{\overline{F}^2(x)} \le (\ge) 0,$$

where $v(x) = \int_x^\infty \overline{F}(u)du$ and $f(x)$ denotes the probability density function corresponding to F. Thus $e(x)$ is decreasing(increasing) if and only if $f(x)v(x) \le (\ge) \overline{F}^2(x)$. Hence, as a measure of the deviation from the null hypothesis H_0 in favor of H_1, we propose the parameter

$$T(F) = \int_0^\tau (f(x)v(x) - \overline{F}^2(x))dx + \int_\tau^\infty (\overline{F}^2(x) - f(x)v(x))dx$$

Note that $T(F)$ is zero for the exponential distribution F and strictly positive for the IDMRL F. Using integration by parts, we can rewrite $T(F)$ as

$$T(F) = \int_0^\infty \overline{F}(x)dx - 2\int_0^\tau \overline{F}^2(x)dx + 2\int_\tau^\infty \overline{F}^2(x)dx - 2\overline{F}(\tau)\int_\tau^\infty \overline{F}(x)dx.$$

Let $F_n(x)$ be the empirical distribution formed by a random sample $X_1, \cdots, X_n$ from F and let $X_{(1)} < \cdots < X_{(n)}$ be the order statistics of the sample. Then we can estimate $T(F)$ by

$$T(F_n) = \sum_{i=1}^{i^*} B_1\left(\frac{n-i+1}{n}\right)(X_{(i)} - X_{(i-1)}) + B_1\left(\frac{n-i^*}{n}\right)(\tau - X_{(i^*)})$$

$$+ B_2\left(\frac{n-i^*}{n}\right)(X_{(i^*+1)} - \tau) + \sum_{i=i^*+2}^{n} B_2\left(\frac{n-i+1}{n}\right)(X_{(i)} - X_{(i-1)}),$$

where $0 = X_{(0)} < X_{(1)} < \cdots < X_{(i^*)} \le \tau < X_{(i^*+1)} < \cdots < X_{(n)}$, $B_1(u) = u - 2u^2$ and $B_2(u) = (1 - 2\overline{F}_n(\tau))u + 2u^2$.

To establish asymptotic distribution of $T(F_n)$, we use the differentiable statistical function approach of von Mises(1947) (cf. Boos and Serfling(1980) and Serfling(1980)). The differential statistical approximation of $T(F_n)$ is defined as

$$T(F_n) = T(F) + d_1T(F)(F_n - F) + R_n(F)$$

where $d_1T(F)(G-F)$ is the first-order Gateaux differential of functional T at the point F in the direction G, and F and G are life distributions in the domain of $T(\cdot)$. For the IDMRL functional T, the Gateaux differential

$$\begin{aligned} d_1T(F)(F_n - F) &= \frac{1}{n}\sum_{i=1}^{n} d_1T(F)(\delta_{X_i} - F) \\ &= \int_0^\infty D_n(x)\,dx + 4\int_0^\tau \overline{F}(x)D_n(x)dx - 4\int_\tau^\infty \overline{F}(x)D_n(x)dx \\ &\quad + 2\overline{F}(\tau)\int_\tau^\infty D_n(x)dx + 2D_n(\tau)\int_\tau^\infty \overline{F}(x)\,dx \end{aligned}$$

where $D_n(x) = \overline{F}(x) - \overline{F}_n(x)$ and $\delta_{X_i}(x) = 0$ if $x < X_i$ and $=1$ if $x \ge X_i$. Our proof of asymptotic normality approximates $T(F_n) - T(F)$ by $d_1T(F)(F_n - F)$ and shows that the term $\sqrt{n}R_n(F)$ converges in probability to 0. Let $\mu(T, F) = E_F[d_1T(F)(\delta_{X_1} - F)]$ and $\sigma^2(T, F) = \mathrm{Var}_F[d_1T(F)(\delta_{X_1} - F)]$. Then we can obtain the following result.

THEOREM 2.1 Let F be the life distribution such that $0 < F(\tau) < 1$ and $\sigma^2(T, F) < \infty$. Then

$$\sqrt{n}(T(F_n) - T(F)) \xrightarrow{d} N(0, \sigma^2(T, F)).$$

PROOF. Applying the classical Lindberg-Levy central limit theorem, we have

$$\sqrt{n}d_1T(F)(F_n-F) \xrightarrow{d} N(0,\sigma^2(T,F)).$$

Next we show that $\sqrt{n}R_n(F)$ converges in probability to 0. By straightforward calculation, for the life distribution F we have

$$\begin{aligned}R_n(F) &= T(F_n)-T(F)-d_1T(F)(F_n-F)\\ &= -2\int_0^\tau(\overline{F}_n(x)-\overline{F}(x))^2dx+2\int_\tau^\infty(\overline{F}_n(x)-\overline{F}(x))^2dx\\ &\quad -2(\overline{F}_n(\tau)-\overline{F}(\tau))\int_\tau^\infty(\overline{F}_n(x)-\overline{F}(x))dx.\end{aligned}$$

Thus for any $\tau>0$ and the life distribution F,

$$\begin{aligned}\sqrt{n}|R_n(F)| &\le 4\sqrt{n}\int_0^\infty(\overline{F}_n(x)-\overline{F}(x))^2dx+2\sqrt{n}|\overline{F}_n(\tau)-\overline{F}(\tau)|\int_0^\infty|\overline{F}_n(x)-\overline{F}(x)|dx\\ &\le 6\sqrt{n}\sup|\overline{F}_n(\tau)-\overline{F}(\tau)|\int_0^\infty|\overline{F}_n(x)-\overline{F}(x)|dx.\end{aligned}$$

By the classical weak convergence of the empirical process, $\sqrt{n}R_n(F)$ converges in probability to 0. This completes the proof. □

Under H_0, (i.e. F is exponential with mean μ), we have that $\sqrt{n}T(F_n)$ is asymptotically normal distributed with mean 0 and variance $\mu^2/3$. The distribution of $T(F_n)$ is not scale invariant. In order to make our test scale invariant we use the test statistic

$$T_n^* = \frac{\sqrt{n}T(F_n)}{\overline{X}}$$

where $\overline{X}$ denote the sample mean. By Slusky's theorem, T_n^* is asymptotically normal distributed with mean 0 and variance $1/3$, under H_0.

The IDMRL test procedures rejects H_0 in favor of H_1 at the approximation level α if $\sqrt{3}T_n^*\ge z_\alpha$. Analogously, the DIMRL test rejects H_0 in favor of H_1' at the approximation level α if $\sqrt{3}T^*\le -z_\alpha$.

3. Power Comparison

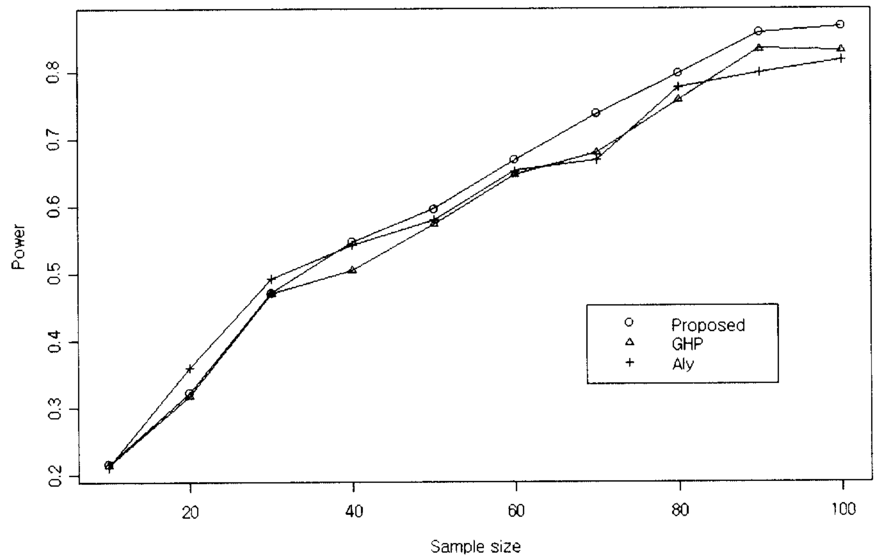
In this section we perform a Monte Carlo simulation to compare the performance of our test with that of GHP's(1986) test and Aly's(1990) test by simulating the power of test. For Monte Carlo study we used the subroutine IMSL of the package FORTRAN on the IBM super computer SP2 at Seoul National University.

To compare the power of our test based on T_n^* with that of GHP's(1986) test based on U_n^* and Aly's(1990) test based on V_n^*, the random numbers are generated from

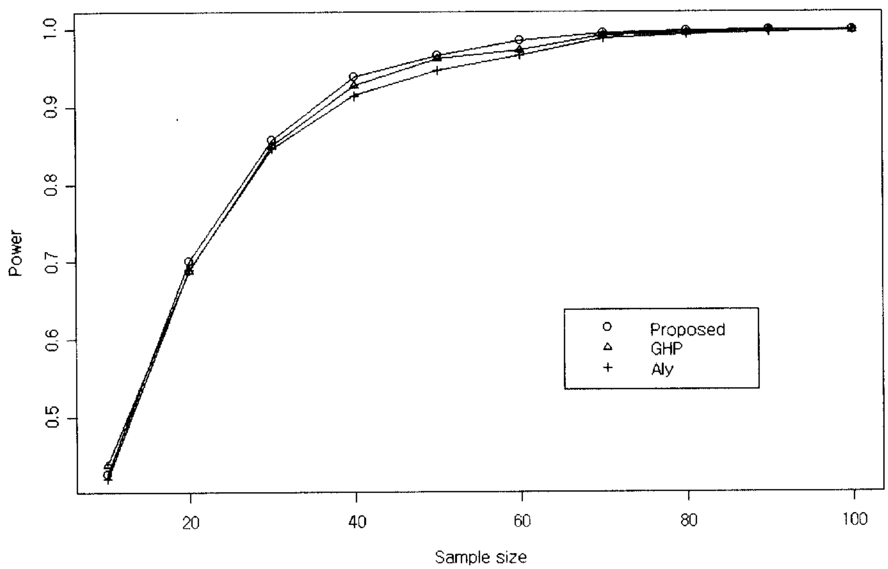
$$\overline{F}_{\alpha,\beta,\gamma}(x)=\left\{\frac{\beta}{\beta+\gamma\exp(-\alpha x)(1-\exp(-\alpha x))}\right\}\left\{\frac{[1+d]^2-c^2}{[\exp(\alpha x)+d]^2-c^2}\right\}^{1/2\alpha\beta}$$
$$\times\left\{\frac{\exp(\alpha x)+d-c}{\exp(\alpha x)+d+c}\,\frac{1+d+c}{1+d-c}\right\}^{\gamma/4\alpha\beta^2 c},\quad x\geq 0$$

where $d=\gamma/2\beta$, $c^2=[4(\beta/\gamma)+1]/[4(\beta/\gamma)^2]$. This distribution has MRL function $e_{\alpha,\beta,\gamma}(x)=\beta+\gamma\exp(-\alpha x)(1-\exp(-\alpha x))$, $x\geq 0$. The motivation for choosing $\overline{F}_{\alpha,\beta,\gamma}$ is that $\overline{F}_{\alpha,\beta,\gamma}$ has IDMRL structure with the turning point $\tau=\ln 2/\alpha$ for any choice of (α,β,γ) and $\overline{F}_{\alpha,\beta,\gamma}$ is exponential distribution if $\gamma=0$.

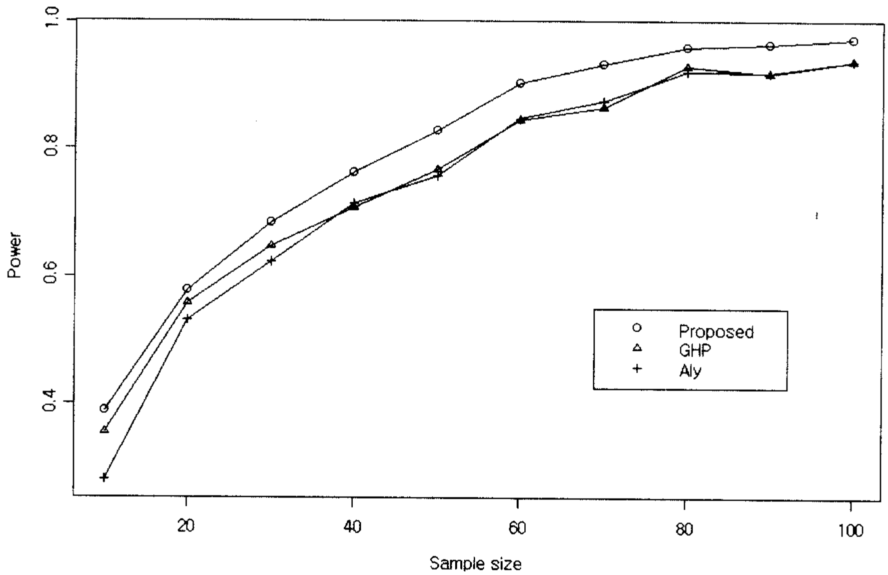
Figures 1~4 contain Monte Carlo estimated powers based on 1000 replications of sample size $n=10,20,\cdots,100$ from $\overline{F}_{\alpha,\beta,\gamma}$ for $\beta=1$ and a selection of (α,γ) when the level of significance is 0.05. From figures, we notice that the powers of 3 tests increase rapidly as γ increases when α is fixed and also as α increases (i.e., the turning point τ decreases) when γ is fixed. It is further better to increase γ than α. This is generally to be expected since the width of $e(x)$ increases as γ increases. Figures show that our test generally dominates the other tests except small α and small γ. But the power of our test increase more rapidly than those of the other tests as n increases for any α and γ.



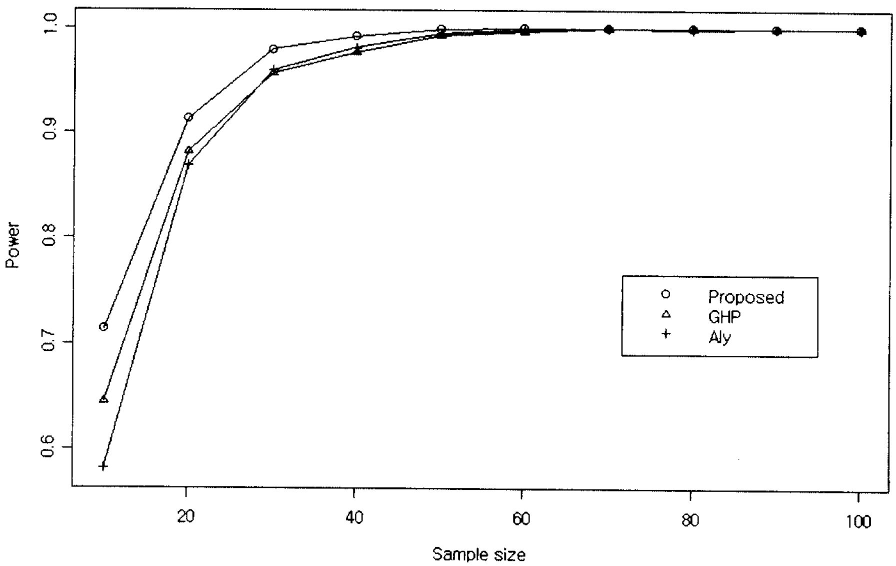
< Figure 1 > Monte Carlo power comparison from 1000 replications with $\alpha=1$, $\beta=1$ and $\gamma=1$



< Figure 2 > Monte Carlo power comparison from 1000 replications with $\alpha=1$, $\beta=1$ and $\gamma=2$



< Figure 3 > Monte Carlo power comparison from 1000 replications with $\alpha=2$, $\beta=1$ and $\gamma=1$



< Figure 4 > Monte Carlo power comparison from 1000 replications with $\alpha=2$, $\beta=1$ and $\gamma=2$

References

[1] Ahmad, I.A.(1992): "A new test for mean residual life times," *Biometrika*, Vol. 79, No. 2, pp. 416-419.

[2] Aly, E.E.A.A.(1990): "Tests for Monotonicity Properties of the Mean Residual Life Function," *Scandinavian Journal of Statistics*, Vol. 17, No. 3, pp. 189-200.

[3] Bhattacharjee, M.C.(1982), "The Class of Mean Residual Lives and Some Consequences," *SIAM Journal of Algebraic Discrete Methods*, Vol. 3, pp. 56-65.

[4] Boos, D.D. and Serfling, R.J.(1980), "A Note on Differentials and the CLT and LIL for Statistical Functions, with Application ti M-Estimates," *Annals of Statistics*, Vol. 8, pp. 618-624.

[5] Cox, D.R.(1962), *Renewal Theory*, Methuem, London.

[6] Guess, F., Hollander, M. and Proschan, F.(1986), "Testing Exponentiality versus a Trend Change in Mean Residual Life," *Annals of Statistics*, Vol. 14, pp. 1388-1398.

[7] Guess, F., and Proschan, F.(1988), "Mean Residual Life: Theory and Applic-ation," *Handbook of Statistics: Quality Control and Reliability 7*, pp. 215-224, North-Holland, Amsterdam.

[8] Hall, W. and Wellner, J.A.(1981), "Mean Residual Life," *Statistics and Related Topics* pp. 169-184, North-Holland Amsterdam.

[9] Kotz, S. and Shanbhag, D.N.(1980), "Some new Approaches to Probability Distributions," *Advance in Applied Probability*, Vol. 12, pp. 903-921.

[10] Serfling, R.J.(1980), *Approximation Theorems of Mathematical Statistics*, Wiely, New York.

[11] von Mises, R.(1947), "On the Asymptotic Distribution of Differentiable Statistical Functions," *Annals of Mathematical Statistics*, Vol. 18, pp. 309-348.