ON A CORRELATION INEQUALITY AND ITS APPLICATIONS

By Mark Brown[1]

The City College, CUNY

Consider a continuous distribution on $[0, \infty)$ with cdf F, survival function $\bar{F} = 1 - F$ and cumulative hazard function $H = -Ln\bar{F}$. For F NBUE it is shown that the correlation coefficient between $X \sim F$ and $H(X)$ is bounded below by σ/μ, the coefficient of variation of F, while for F NWUE the correlation coefficient is bounded below by μ/σ. Several applications of this inequality and its generalizations are discussed, including Monte Carlo simulation of the renewal function, exponential approximation of DMRL distributions, moment inequalities for record values, and a variance inequality for random event epochs in a homogeneous Poisson process.

1. Introduction and Summary. Consider a continuous distribution on $[0, \infty)$, with cdf F, survival function $\bar{F} = 1 - F$ and cumulative hazard function $H = -Ln\bar{F}$. If $X \sim F$ then $H(X)$ is exponentially distributed with mean 1. The random variable $H(X)$ measures lifetime by total hazard overcome until death, while X measures lifetime in ordinary time units. Since H is an increasing function we know that $H(X)$ and X are positively correlated. The question of how positively correlated arose naturally in Brown, Solomon, and Stephens (1981) and Brown (1987) in different contexts. In the former paper the asymptotic relative savings in risk between two Monte Carlo estimators of the renewal function was given by the square of the correlation coefficient between X and $H(X)$. In Brown (1987), a quantity closely related to the correlation coefficient was needed to bound the distance between a DMRL (decreasing mean residual life) distribution and its stationary renewal distribution.

In this paper we show that for X NBUE (new better than used in expectation):

$$\rho(X, H(X)) \geq \frac{\sigma}{\mu} \tag{1}$$

while for X NWUE (new worse than used in expectation):

[1]Research supported by the Air Force Office of Scientific Research under Grant No. 84–0095.

AMS 1980 subject classifications. Primary 60E15; secondary 60G55.

Key words and phrases. Inequalities, correlation coefficient, record values, exponential approximation, Poisson processes, NBUE and NWUE distributions.

I wish to thank the editors and the referee for their helpful suggestions.

$$\rho(X, H(X)) \geq \frac{\mu}{\sigma}. \tag{2}$$

The lower bound for the correlation coefficient is also applicable to record value processes. Record value processes are of interest in reliability theory due to their connection with minimal repair (Crow, 1974). Consider an i.i.d. sequence $\{X_i, i \geq 1\}$ with $X_i \sim F$. Define $S_1 = X_1$, $N_2 = \min\{i : X_i > X_1\}$, $S_2 = X_{N_2}$, $N_k = \min\{i : X_i > X_{N_{k-1}}\}$, $S_k = X_{N_k}$, $k = 3, 4, \ldots$; S_k is referred to as the k^{th} record value. Using (1) and (2) we show that for F NBUE:

$$\frac{\sigma^2}{\mu} \leq E(S_2 - S_1) \leq \sigma \tag{3}$$

and for F NWUE:

$$\mu \leq E(S_2 - S_1) \leq \sigma. \tag{4}$$

Using similar methodology, inequalities are derived for the moments of higher record values. For example, it is shown (Section 3.4) that for F IFRA:

$$\frac{\mu_{k+r-1}}{\mu_{r-1}} \leq ES_r^k \leq \binom{k+r-1}{k} \mu_k \tag{5}$$

where $\mu_j = \int x^j dF(x)$.

In Section 3.5 we show that if $\{T_i, i \geq 1\}$ are the arrival epochs for a homogeneous Poisson process with parameter λ, and N is a stopping time, then $\text{Var } T_N \geq \lambda^{-2}$. We further demonstrate among distributions with failure rate uniformly bounded above by λ, the exponential distribution with parameter λ has minimum variance for the k^{th} record value, for $k \geq 1$. Equivalently, among non-homogeneous Poisson processes with intensity function uniformly bounded above by λ, the homogeneous Poisson process with parameter λ has minimum variance for S_k, the time of the k^{th} event, for $k \geq 1$.

2. A Correlation Inequality. Consider a non-negative random variable X with continuous cdf F. Denote by X^* a random variable with cdf $G(x) = \mu^{-1} \int_0^x \bar{F}(t)dt$, the stationary renewal distribution corresponding to F. Let T denote a random variable with distribution $dF_T(t) = t\mu^{-1}dF(t)$. T is distributed as the length of the interval covering an arbitrary fixed point in a stationary renewal process with interarrival time distribution F (Feller, 1971, p. 371). For fixed $x \geq 0$ define $g(t) = \bar{F}(x \vee t)/\bar{F}(t)$, where $x \vee t = \max(x, t)$. Note that:

$$\begin{aligned} \Pr(T > x) &= \int_x^\infty t\mu^{-1}dF(t) = x\mu^{-1}\bar{F}(x) + \bar{G}(x) \\ &= \int \left(\frac{\bar{F}(x \vee t)}{\bar{F}(t)}\right)\left(\frac{\bar{F}(t)}{\mu}\right) dt = Eg(X^*). \end{aligned}$$

Thus:

$$\Pr(T > x) = Eg(X^*). \tag{6}$$

Next, consider the record value process corresponding to F, described in Section 1. The sequence of record values $\{S_i, i \geq 1\}$ generates a nonhomogeneous Poisson process with $EN(t) = -Ln\bar{F}(t) = H(t)$ (Shorrock, 1972). Now:

$$\begin{aligned} \Pr(S_2 > x) = \Pr(N(x) \leq 1) &= \bar{F}(x)[1 + H(x)] = \\ &\int \left(\frac{\bar{F}(x \vee t)}{\bar{F}(t)}\right) dF(t) = Eg(X). \end{aligned} \tag{7}$$

LEMMA 2.1. *If F is NBUE (NWUE) then S_2 is stochastically larger (smaller) than T.*

PROOF. F NBUE is equivalent to $X \overset{\text{st}}{\geq} X^*$. Since g is increasing in t, it follows from (6) and (7) that:

$$\Pr(T > x) = Eg(X^*) \leq Eg(X) = \Pr(S_2 > x).$$

The NWUE case similarly follows.

LEMMA 2.2. *If F is NBUE then $\rho(X, H(X)) \geq \sigma/\mu$. If F is NWUE then $\rho(X, H(X)) \geq \mu/\sigma$.*

PROOF. Note that $dF_{S_2}(t) = H(t)dF(t)$ while $dF_T(t) = t\mu^{-1}dF(t)$. By Lemma 2.1, if F is NBUE then:

$$ES_2 = E(XH(X)) \geq ET = \mu_2/\mu. \tag{8}$$

Now subtract μ and divide by σ on both sides of (8) and the NBUE result follows.

Next, assume that X is NWUE. It follows from Lemma 2.1 that for any increasing function ℓ (with the expectations existing):

$$E\ell(S_2) = \int \ell(x)H(x)dF(x) \leq \mu^{-1} \int x\ell(x)dF(x) = E\ell(T) \tag{9}$$

choose $\ell(x) = H(x)$, then:

$$EH^2(X) = 2 \leq \mu^{-1}E(XH(X)). \tag{10}$$

From (10) the NWUE result easily follows.

3. Applications.

3.1. Monte Carlo Estimation of the Renewal Function. Suppose we wish to estimate $M(t)$, the expected number of renewals in $[0, t]$ for a renewal process with interarrival distribution F, by Monte Carlo simulation. An obvious approach is to simulate $N(t)$, the number of renewals in $[0,t]$, K times $(N_1(t),\ldots,N_K(t))$, and to estimate $M(t)$ by the sample mean. In Brown, Solomon, and Stephens (1981) an unbiased estimator $M^*(t)$ was proposed and it was shown that as $t \to \infty$ the asymptotic relative savings in risk between $M^*(t)$ and the estimator based on $N(t)$ was given by $\rho^2(X, H(X))$. Lemma (2.2) gives a lower bound on ρ and thus a lower bound on the asymptotic relative savings in risk.

3.2. Exponential Approximation of DMRL Distributions. Consider a continuous DMRL (decreasing mean residual life) distribution F on $[0,\infty)$ with stationary renewal distribution G. In Brown (1987) it is shown that:

$$\mathcal{D}^*(F,G) = \sup \mid F(B) - G(B) \mid \leq 1 - EH(X^*) \tag{11}$$

where $H = -Ln\bar{F}$, the cumulative hazard function, $X^* \sim G$, and the sup is taken over all Borel sets. Now:

$$ES_2 = \int (H(t)+1)\bar{F}(t)dt = \mu[1 + EH(X^*)]. \tag{12}$$

But F DMRL implies F NBUE, thus (12) and Lemma 2.1 give:

$$ES_2 = \mu[1 + EH(X^*)] \geq ET = \mu_2/\mu \tag{13}$$

thus:

$$EH(X^*) \geq \sigma^2/\mu^2. \tag{14}$$

From (11) and (14) we obtain:

$$\mathcal{D}^*(F,G) \leq 1 - (\sigma^2/\mu^2). \tag{15}$$

The inequality (15) thus extends the result of Brown (1987) from F IFR to F DMRL. Moreover it follows from (15), employing the methodology of Brown (1987), that for F DMRL:

$$\sup \mid \bar{F}(t) - e^{-t/\mu} \mid \leq 1 - (\sigma^2/\mu^2). \tag{16}$$

Thus if F is DMRL with coefficient of variation close to 1, then F is approximately exponential.

3.3. The Second Record Value. Consider $S_2 - S_1$ the interarrival time between the first and second record values in a record value process corresponding to F (equivalently the interarrival time between the first and second events in a non-homogeneous Poisson process with $EN(t) = H(t) = -Ln\bar{F}(t)$). It follows from Lemma 2.1 that F NBUE implies:

$$E(S_2 - S_1) \geq ET - \mu = \frac{\mu_2}{\mu} - \mu = \frac{\sigma^2}{\mu} \tag{17}$$

while F NWUE leads to $E(S_2 - S_1) \leq (\sigma^2/\mu)$.

The quantity $E(S_2 - S_1)$ is the expected residual life for an item which is minimally repaired at its first failure. It is of interest in the evaluation and planning of maintenance policies.

Lemma 3.3.1, below, presents an upper bound of σ for $E(S_2 - S_1)$, derived without aging assumptions of F. As is done throughout this paper we assume that F is a continuous distribution on $[0, \infty)$.

LEMMA 3.3.1. *Let $X \sim F$ and g a function on $[0, \infty)$ with $Eg^2(X) < \infty$. Then:*

$$| E(g(S_2) - g(S_1)) | \leq \sigma_g$$

where σ_g is the standard deviation of $g(X)$. In particular the choice $g(x) = x$ gives:

$$E(S_2 - S_1) \leq \sigma$$

where σ is the standard deviation of X.

PROOF. $Eg(S_2) = E(g(X)H(X)) = Eg(X)EH(X) + \sigma_g \sigma_{H(X)} \rho(g(X), H(X)) \leq Eg(X) + \sigma_g$. Thus $E(g(S_2) - g(S_1)) \leq \sigma_g$. Substituting $-g$ for g yields $E(g(S_1) - g(S_2)) \leq \sigma_g$ from which the result follows. ||

COROLLARY 3.3.1. *For F NBUE, $\sigma^2/\mu \leq E(S_2 - S_1) \leq \sigma$. For F NWUE, $\mu \leq E(S_2 - S_1) \leq \sigma$.*

PROOF. The NBUE case follows from expression (17) and Lemma 3.3.1. The NWUE case follows from Lemma 3.3.1 and the obvious NWUE inequality $E(S_2 - S_1) \geq \mu$. ||

A function $g(x)$ on $[0, \infty)$ is defined to be starshaped if $\frac{g(x)}{x}$ is increasing (meaning non-decreasing). If g is non-negative and starshaped then g is increasing.

Consider, now, a function g which is non-negative and starshaped on $[0, \infty)$, with $\mu_g = Eg(X) < \infty$. Define:

$$dF_g(t) = g(t) dF(t)/\mu_g.$$

Then: $dF_g(t)/dF_T(t) = \mu_g^{-1} \mu (g(t)/t)$ which is increasing. Thus F_g is larger than F_T under the partial ordering of monotone likelihood ratio (Lehmann [1959] p.73) and is thus stochastically larger. It follows that:

$$E[Xg(X)] \geq \mu_g \mu_2/\mu. \tag{18}$$

Now assume that F is NBUE. By Lemma 2.1 and (18):

$$Eg(S_2) \geq Eg(T) = \mu^{-1} E(Xg(X)) \geq \mu_g \mu_2/\mu^2. \tag{19}$$

Thus for F NBUE and g non-negative and starshaped it follows from Lemma 3.3.1 and (19) that:

$$\frac{\sigma^2}{\mu^2}\mu_g \leq E(g(S_2) - g(S_1)) \leq \sigma_g. \tag{20}$$

The choice $g(x) = x$ leads to the NBUE inequality of Corollary 3.3.1.

3.4. Higher Record Values. Let S_k denote the k^{th} record value in a record value process corresponding to F continuous. Since S_k is the k^{th} event epoch in a non-homogeneous Poisson process with $EN(t) = H(t)$ it follows that:

$$dF_{S_k}(t) = [(H(t))^{k-1}/(k-1)!]dF(t) \tag{21}$$

and also that:

$$dF_{S_k}(t) = [(H(t)/k-1)]dF_{S_{k-1}}(t), \;\; k \geq 2. \tag{22}$$

Consequently (from 22):

$$Eg(S_k) = (k-1)^{-1}E[g(S_{k-1})H(S_{k-1})]. \tag{23}$$

Now $H(S_{k-1})$ is gamma distributed with parameters $k-1$ and 1 (the sum of $k-1$ i.i.d. exponentials with parameter 1) thus $ES_{k-1} = \text{Var}S_{k-1} = k-1$.

Using the mean and variance of $H(S_{k-1})$, (23) and the upper bound for the product moment, EUV $\leq$ EUEV $+\sigma_U\sigma_V$ with $U = S_{k-1}, V = H(S_{k-1})$ we obtain:

$$Eg(S_k) \leq Eg(S_{k-1}) + (\sigma(g(S_{k-1}))/\sqrt{k-1}). \tag{24}$$

From (24) we obtain the following generalizations of Lemma (3.3.1):

$$\mid E[(g(S_k) - g(S_{k-1})] \mid \leq \sigma(g(S_{k-1}))/\sqrt{k-1}. \tag{25}$$

The case $k = 2$ corresponds to Lemma 3.3.1. However the more general inequality appears to be computationally useful only when $k = 2$. For general k, $\sigma(g(S_{k-1}))$ is no easier to compute than $E(g(S_k) - g(S_{k-1}))$.

We have no analogue of Lemma 2.1 for F NBUE or NWUE. However if we strengthen the restriction on F from NBUE (NWUE) to IFRA (DFRA) then we obtain the following:

LEMMA 3.4.2. *Let F be a continuous IFRA distribution, and T_r be a random variable with distribution $dF_{T_r}(t) = x^{r-1}dF(x)/\mu_{r-1}$, where μ_m is the m^{th} moment of F. Then S_r is stochastically larger than T_r and:*

$$\frac{\mu_{k+r-1}}{\mu_{r-1}} \leq ES_r^k \leq \binom{k+r-1}{k}\mu_k.$$

If F is a continuous DFRA distribution with finite $(r-1)^{st}$ moment then S_r is stochastically smaller than T_r. If in addition F has finite $(k+r-1)^{st}$ moment then the above inequality reverses.

For $r=2$ the above inequalities hold under the weaker condition that F is NBUE or NWUE.

PROOF. Note that $dF_{S_r}(t)/dF_{T_r}(t) = \frac{\mu_{r-1}}{(r-1)!}\left(\frac{H(t)}{t}\right)^{r-1}$ which is increasing, as F is IFRA. Thus S_r is larger than T_r under monotone likelihood ratio and is thus stochastically larger. Thus:

$$ES_r^k \geq ET_r^k = \mu_{k+r-1}/\mu_{r-1}. \tag{26}$$

Next:

$$\frac{H^k}{k!}dF \overset{\text{st}}{\geq} \frac{x^k}{\mu_k}dF. \tag{27}$$

Multiply both sides of (27) by $H^{r-1}/(r-1)!$ and integrate obtaining:

$$\binom{k+r-1}{k} \geq \frac{1}{\mu_k}ES_r^k. \tag{28}$$

Thus $ES_r^k \leq \mu_k\binom{k+r-1}{k}$ and this inequality and (26) yield the IFRA result. The DFRA case similarly follows. By Lemma 2.1, for F NBUE and $r=2$, $S_2 \overset{\text{st}}{\geq} T_2 = T$ which is sufficient, by our above derivation, for (26) and (28) to follow (with $r=2$). ||

Note that the various inequalities derived above for record value processes hold for non-homogeneous Poisson processes.

3.5. A Variance Inequality. Consider an absolutely continuous distribution F with failure rate function $h(t)$ bounded above by $\lambda(h(t) \leq \lambda$ for all $t \geq 0)$. Let S_1 and S_2 denote the first two record values in a record value process corresponding to F. The failure rate function of $S_2 - S_1$ evaluated at t is a mixture of the values $\{h(s), s \geq t\}$ and is thus bounded above by λ for all t. Consequently:

$$E(S_2 - S_1) \geq \lambda^{-1}. \tag{29}$$

By Lemma 3.3.1:

$$E(S_2 - S_1) \leq \sigma \tag{30}$$

where σ is the standard deviation corresponding to F. From (29) and (30) we obtain:

$$\sigma^2 \geq \lambda^{-2}. \tag{31}$$

Thus among distributions on $[0,\infty)$ with failure rate bounded above by λ, the exponential distribution with parameter λ has smallest variance.

Next, consider a homogeneous Poisson process on $[0,\infty)$ with intensity λ and event epochs $\{T_i, i \geq 1\}$. Let N be a stopping time and consider the random variable T_N, letting h^* denote its failure rate function. Since T_N can only occur at an event epoch for the Poisson process, and since the conditional intensity of an event at t (given the past) for the Poisson process is always λ, it follows that:

$$h^*(t) = \lambda \Pr(T_N = t \mid T_N \geq t, T_i = t \text{ for some } i) \leq \lambda. \tag{32}$$

Thus (31) and (32) imply:

$$\mathrm{Var}(T_N) \geq \lambda^{-2}. \tag{33}$$

Note that λ^{-2} is the variance of T_1 as well as the variance of $T_{N(t)+1}$, the time of the first event after time t. These event epochs have smallest variance among all random event epochs for the Poisson process.

The inequality (33) holds for a large variety of random variables arising in secondary processes generated by a Poisson process. These include counter models, queues with Poisson input and uniformizable Markov chains.

Also note that if the failure rate of F is uniformly bounded above by λ, then by (25) and the argument used to derive (29):

$$\lambda^{-1} \leq E(S_{k+1} - S_k) \leq k^{-1/2}\sigma(S_k). \tag{34}$$

Thus $\sigma(S_k) \geq k^{1/2}\lambda^{-1}$, the lower bound being achieved in the exponential case. Thus for $k \geq 1$, the exponential distribution with parameter λ minimizes the variance of the k^{th} record value, among distributions with failure rate function uniformly bounded above by λ. Equivalently, consider a non-homogeneous Poisson process with intensity function $\lambda(t)$ bounded above by λ. Then Var $S_k \geq k\lambda^{-2}$ where S_k is the k^{th} arrival epoch. Thus among all non-homogeneous Poisson processes with intensity functions uniformly bounded above by λ, the homogeneous Poisson process with intensity λ minimizes the variance of S_k, for all $k \geq 1$.

REFERENCES

BROWN, M. (1987). Inequalities for distributions with increasing failure rate. *Contributions to the Theory and Application of Statistics, A Volume in Honor of Herbert Solomon*, pp. 3–17. Academic Press, New York.

BROWN, M., SOLOMON, H., and STEPHENS, M.A. (1981). Monte Carlo simulation of the renewal function. *J. Appl. Prob.* **18** 426–434.

CROW, L.H. (1974). Reliabilitly analysis for complex repairable systems. *Reliability and Biometry, Statistical Analysis of Lifelength*, (F. Proschan and R.J. Serfling, eds.), SIAM, Philadelphia, 379–414.

FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications, II, 2nd Edition.* John Wiley and Sons, New York.

LEHMANN, E.L. (1959). *Testing Statistical Hypotheses.* John Wiley and Sons, New York.

SHORROCK, R.W. (1972). A limit theorem for inter-record times. *J. Appl. Prob.* **9** 219–223.

DEPARTMENT OF MATHEMATICS
CITY COLLEGE, CUNY
138 ST. & CONVENT AVENUE
NEW YORK, NY 10031