Approximating IMRL Distributions by
Exponential Distributions, with Applications
 to First Passage Times

by

Mark Brown
Florida State University and City College, CUNY

June, 1981

AFOSR Technical Report No. 81-125

The Florida State University
Department of Statistics
Tallahassee, Florida 32306

and

City College, CUNY
Department of Mathematics
New York, New York 10031

Research supported by Air Force Office of Scientific Research, AFSC, USAF, under Grant No. AFSOR F49620-79-C-0157 and Grant No. AFSOR 81-0038.

AMS 1970 subject classifications. Primary 62E10; Secondary 60K10.

Key words and phrases: IMRL distributions, approximate exponentiality, first passage times, Markov processes, reliability theory, inequalities.
Approximating IMRL Distributions by Exponential Distributions, with Applications to First Passage Times.

by

Mark Brown

City College, CUNY and Florida State University

Summary It is shown that if $F$ is an IMRL (increasing mean residual life) distribution on $[0, \infty)$ then:

$$\max(\exp |G(t)|, \sup_t |F(t) - e^{-t/\mu}|, \sup_t |G'(t) - e^{-t/\mu}|, \sup_t G(t) - e^{-t/\mu} |) = 1 - \frac{\mu}{\mu_G}$$

Where $F(t) = 1 - F(t)$, $\mu = E_X$, $\mu_2 = E(X^2)$, $G(t) = \mu^{-1} \int_0^t F(x) dx$.

$\mu_G = E_X^2/2\mu^2$, and $\rho = \frac{\mu_2}{2\mu^2} - 1 \leq \frac{\mu_G}{\mu} - 1$.

Thus if $F$ is IMRL and $\rho$ is small then $F$ and $G$ are approximately equal and exponentially distributed. IMRL distributions with small $\rho$ arise naturally in a class of first passage time distributions for Markov processes, as first illuminated by Accett. The current results thus provide error bounds for exponential approximations of these distributions.
The exponential distribution satisfies
\[ \rho = \frac{\nu_2}{2\nu_1} - 1 = 0 \]
where \( \mu = \text{EX} \) and \( \nu_2 = \text{EX}^2 \). In general a small value of \( |\rho| \) does not in itself imply approximate exponentiality.

For example the binomial distribution with \( n=1 \) and \( p=\frac{1}{2} \) has \( \rho=0 \). It is reasonable to conjecture however that within certain classes of distributions a small value of \( |\rho| \) does imply approximate exponentiality. In these cases, given the class and the first two moments of the distribution, it would be desirable to obtain bounds on the distance from exponentiality.

The above problem for the class of completely monotone distributions (mixtures of exponential distributions) has received some recent attention. The motivation for this interest is that completely monotone distributions with small \( \rho \) arise naturally in first passage time distributions for Markov processes, as first illuminated by Keilson ([11]-[14]). Keilson and Steutel [15] suggested \( \rho \) as a measure of departure from exponentiality within this class. Keilson [11] using results of Heyde [8] derived
\[ (1 \cdot 1) \ d(F, F_{\mu}) = \sup |F(x) - e^{-x/\mu}| \leq k \rho^{1/4} \]
where \( k \) is bounded above by 4*41. Heyde and Leslie [9] improved the right hand side of \( (1 \cdot 1) \) to 3*74\( \rho \) and Hall [7] obtained a further improvement to 2*77\( \rho \).
This paper considers INRL (increasing mean residual life) distributions, a class considerably larger than that of completely monotone distributions. (INRL distributions are defined and discussed in section 2). Defining \( G(x) = s^{-1} \int_0^x f(s) \) (the stationary renewal distribution), \( \mathbb{E}_D \) as an exponential distribution with failure rate \( \lambda \) (and mean \( \frac{1}{\lambda} \)), \( \mu_G = \mathbb{E}_G X = \frac{\mu}{2\mu} \), \( \sigma^2 = \text{Var}_F X \), \( d(F_1, F_2) = \sup_{t \geq 0} |F_1(t) - F_2(t)| \) and \( d^*(F_1, F_2) = \sup_{\beta} |F_1(\beta) - F_2(\beta)| \) where \( \beta \) is the collection of \( \sigma \)-dense subsets of \( \{0, \infty\} \), we obtain:

\[
(1.2) \quad d(F, E_{\mu - 1}) \leq \frac{\rho}{\rho + 1} = 1 - \frac{1}{\mu_G} = 1 - \frac{2\mu^2}{\sigma^2 + 2\mu^2} = \frac{\sigma^2 - \mu^2}{\sigma^2 + 2\mu^2} \\
(1.3) \quad d^*(F, G) \leq \frac{\rho}{\rho + 1} \\
(1.4) \quad d^*(G, E_{\mu^*}) \leq \frac{\rho}{\rho + 1} \\
(1.5) \quad d(G, E_{\mu - 1}) \leq \frac{\rho}{\rho + 1} \\

Since \( F = G \) if and only if \( F \) is exponential one would expect that for small \( \rho \), \( F \) and \( G \) are approximately equal and that \( G \) is approximately exponential. Expressions (1.2) (1.5) give error bounds for these approximations.

The inequalities (1.2) and (1.3) are shown to be sharp even within the subclass of completely monotone distributions.
Thus $\frac{d}{\rho+1}$ is the best upper bound for both $d(F, F_\mu)$ and $d(F, G)$ (for $F$ IMRL, $d(F, G) = d^*(F, G)$).

The current results thus extend the domain of applicability from completely monotone to IMRL distributions, improve the upper bound for $d(F, F_\mu)$ from $2.77\rho$ to its best value $\frac{d}{\rho+1}$, apply to distributions of the form $G(x) = \int_0^x u^{-1} F(s) ds$, with $F$ IMRL, without requiring knowledge of $E_G X^2$ (which may be infinite), and bound the distance between $F$ and $G$.

The inequalities (1.2)-(1.5) are derived along with related results in section 4. The sharpness of (1.2) and (1.3) is discussed in section 5. Further bounds for the completely monotone class are derived in section 6. It is remarked in section 7 that the results immediately apply to the mixtures of distributions from proportional hazard families. In section 8 the above inequalities are applied to the class of first passage time distributions considered by Keilson ([11],[14]), which, as mentioned above, motivated much of the interest in this problem.

2 - IMRL AND RELATED DISTRIBUTIONS. A distribution $F$ is defined to be IMRL on $[0, \infty)$ if $\mu = \int x dF(x) < \infty, F'(0-) = 3, F'(1) < 1$ and $E(\lambda-t|X>t) = \left( \int_t^\infty F(x) dx \right) / F(t)$ is increasing in $t \geq 0$.
A distribution $F$ is defined to be DFR on $[0,\infty)$ if $F(0^-)=0$, $F(0)<1$ and $P(x>s+t|x>t)=\frac{F(s+t)}{F(t)}$ is increasing in $t\geq0$ for each $s>0$. In our work below all INRL and DFR distributions will be on $[0,\infty)$ as defined above and we will leave out the phrase "on $[0,\infty)$" in referring to them.

Since stochastic ordering implies ordering of means, $F$ DFR and $\mu<\infty$ implies $F$ INRL. However it is easy to provide examples in which $F$ is INRL but not DFR.

Define $G(t)=\mu^{-1}\int_0^t \frac{F(x)}{x}dx$, the stationary renewal distribution corresponding to $F$. Note that $G$ is absolutely continuous with failure rate function $\lambda_G(t)=\frac{F(t)}{\mu G(t)}=E(x-t|x>t)^{-1}$. Therefore $F$ INRL $\iff$ $G$ DFR. It follows that if $F$ is DFR with $F(0)=0$ and $f(0^+)\in(0,\infty)$ (F is necessarily absolutely continuous with its pdf $f$ possessing a version which is decreasing) then $F$ is the stationary renewal distribution corresponding to the INRL distribution with survival function $L(x)=f(x)/f(0^+)$. If $F$ is INRL then $\lambda_G$ is decreasing. Defining $q=F(0)$ we thus have:

$$(2\cdot1)\lambda_G(t)\leq\lambda_G(0)\leq q/\mu \leq 1/\mu$$
It immediately follows that:

\[(2.2) \quad \tilde{G}(t) = e^{-tq/\mu} - e^{-t/\mu} \]

A consequence of (2.2) is that for all decreasing functions \( \lambda \):

\[(2.3) \quad \int \lambda(t) dG(t) = \mu^{-1} \int \lambda(t) \tilde{F}(t) dt \leq \mu^{-1} \int \lambda(t) e^{-t/\mu} dt \]

Since \( \tilde{G}(t)/\tilde{F}(t) = \mu^{-1} E(X-t \mid X \geq t) \) it follows that for \( F \)

\[(2.4) \quad \tilde{G}(t)/\tilde{F}(t) \text{ is increasing} \]

Furthermore since \( \tilde{G}(0)/\tilde{F}(0) = q^{-1} \), it follows from (2.4) that:

\[(2.5) \quad \tilde{G}(t) \geq q^{-1} \tilde{F}(t) \leq \tilde{F}(t) \]

3. Distance between distributions. Suppose that \( X \sim F_1 \),

and \( Y \sim F_2 \). Then:

\[(3.1) \quad \sup_t |F_1(t) - F_2(t)| \leq \sup_t |F_1(t) - F_2(t)| \leq \Pr(X \neq Y) \]

Thus given two distributions \( F_1, F_2 \) we try to construct

random variables \( X, Y \) with \( X \sim F_1, Y \sim F_2 \) and \( \Pr(X \neq Y) \) small.

Then (3.1) is invoked to show that \( F_1 \) and \( F_2 \) are close.

This is the approach of Hodges and LeCam [10] in their study

of Poisson approximation to sums of independent Bernoulli

variables. The contruction for \( X, Y \) is found in Lemma 3.3.

The author employed a similar construction in [6].
If $F_1, F_2$ are absolutely continuous with respect to

$\mu$, with Radon-Nikodym derivatives $f_1, f_2$ then:

$$
(3.2) \quad \int \frac{|f_1 - f_2|}{f_1 f_2} \, d\mu = 2 \int f_1 - f_2 \, d\mu = 2 \sup_{\beta} |F_1(\beta) - F_2(\beta)|.
$$

Thus if $Pr(X \neq Y)$ is small then $f_1, f_2$ are close in $L_1(\mu)$ norm.

(3.3) **Lemma** Suppose that $F_1, F_2$ are distributions with

$\text{support } (0, \infty)$, $H_i(t) = -\ln(F_i(t)/P_i(0))$ $i = 1, 2$ has no atoms

and $\overline{F}_2(t)/\overline{F}_1(t)$ is increasing in $t$.

Then:

$$
(3.4) \quad \sup_{\beta} \int F_1^{-1}(\beta) - F_2^{-1}(\beta) \, d\overline{F}_1(t) \overline{H}_2(t)
$$

**Proof.** Construct two independent non-homogeneous Poisson

processes, $(H_i(t), t \geq 0)$ $i = 1, 2$, with $\overline{H}_i(t) = \overline{H}_2(t)$ and

$E[H_2(t)] = H_2(t)$ Note that $\overline{F}_2/\overline{F}_1$ increasing implies

$H_1 - H_2$ is increasing. Define $Y_i$ to be the first event

epoch from process $i, i = 1, 2$. Construct $\rho_i$ independent

of the two Poisson processes with $Pr(\rho_i = 0) = P = F_1(0)$,

$Pr(\rho_i = 1) = F_i(0)$ Define $X = \rho_{\min}(Y_1, Y_2)$ and $Y = Y_1$. Now

$Pr(Y \geq t) = Pr(H_1(t) = 0) \stackrel{a}{=} \overline{F}_1(t)$, thus $X \sim F_2$. 


Similarly, \( \Pr(X>t) = \Pr(\rho = 1) \Pr(N_1(t) + N_2(t) = j) \)

\[= qe^{-(\bar{h}_1(t) + \bar{h}_2(t))} \int_0^t e^{-(\bar{h}_2(t))} \, d\bar{h}_2(t) \]

Finally:

\[= \int_0^t F_1(t) \, d\bar{h}_2(t) \]

The result now follows from (3.1) and (3.3).

Given two distribution functions \( F_1, F_2 \) define \( d(F_1, F_2) = \sup_t |F_1(t) - F_2(t)| = \sup_t |F_1(t) - \bar{F}_2(t)|. \)

(3.6) **Lemma** Suppose that for each \( t \) either \( F_3(t) = \max(F_1(t), F_2(t)) \) or \( F_3(t) = \min(F_1(t), F_2(t)) \). Then \( d(F_1, F_2) \leq \max\{d(F_1, F_3), d(F_2, F_3)\}. \)

**Proof** The above condition immediately implies that \( |F_1(t) - F_2(t)| \leq \max(|F_1(t) - F_3(t)|, |F_2(t) - F_3(t)|) \) for all \( t \).

Taking the \( \sup \) of both sides gives the result.

4. **Inequalities.** Recall that \( d(F_1, F_2) \) is defined to equal \( \sup_t |F_1(t) - F_2(t)|. \) Similarly define

\[d^*(F_1, F_2) = \sup_{\beta} |F_1(\beta) - F_2(\beta)| \]

where \( \beta \) is the collection of Borel subsets of \([0, \infty)\). Let \( F_\beta \) denote an exponentially distributed random variable with failure rate \( \beta \) (and mean \( \beta^{-1} \).
Theorem Assume that $F$ is LRM with $\mu^\infty$, $G(x) = \mu^{-1} \int_0^\infty \bar{F}(s) ds = F(0)$ and $q = 1 - p$.

Then:

(i) $d^*(F,G) \leq \frac{\mu}{\rho + 1} = 1 - \frac{\mu}{\mu_G} = 1 - \frac{2\mu^2}{\mu_2^2} = \frac{q^2 - \mu^2}{\sigma^2 + \mu^2}$

(ii) $\bar{F}(t) \leq \bar{G}(t) \leq \bar{F}(t) + \frac{\rho}{\rho + 1}$ for all $t$.

(iii) $d^*(G,E_{\mu^{-1}}) \leq \frac{\rho}{\rho + 1}$

(iv) $e^{-t/\mu} \leq \bar{G}(t) \leq e^{-t/\mu} + \frac{\rho}{\rho + 1}$ for all $t$.

(v) $d(G,E_{\mu^{-1}}) \leq \frac{\rho}{\rho + 1}$

(vi) $d(G,E_{\mu}) \leq \frac{\rho}{\rho + 1}$ for $\mu_G^{-1} \leq bs < \mu_G^{-1}$

In particular $d(G,E_{\mu^{-1}}) \leq \frac{\rho}{\rho + 1}$.

(vii) $\sup_t \left| \bar{F}(t) - qe^{-qt/\mu} \right| \leq \frac{q^2 - \mu^2}{\rho + 1}$

(viii) $d^*(G,E_{q\mu^{-1}}) \leq \frac{\rho}{\rho + 1} + \frac{\rho}{q}$

(ix) $e^{-qt/\mu} \leq \bar{G}(t) \leq e^{-qt/\mu} + \frac{\rho}{\rho + 1}$ for all $t \geq 0$.

(x) $d(G,E_{\mu}) \leq \begin{cases} \frac{\rho}{\rho + 1} + \frac{\rho}{q} & \text{for } \mu_G^{-1} \leq bs < \mu_G^{-1} \\ 1 - \mu_G^{-1} & \text{for } bs < \mu_G^{-1} \end{cases}$

(xi) Assume that $F$ is DFR with $F(0) = 0$ and $f(0^+) = \mu^\infty$. Then $d^*(F,E) = 1 - (\alpha\mu)^{-1}$.
(xii) Under the conditions of (xi), \(e^{-at} \leq \overline{G}(t) \leq e^{-at} + (\mu)^{-1}\) for all \(t \geq 0\).

(xiii) Under the conditions of (xi):

\[
d(F, G) = \begin{cases} 
(1 - (\mu)^{-1}) & \text{for } \mu^{-1}s > s \\
1 - b & \text{for } b > s^{-1}
\end{cases}
\]

Proof (1) By (2.4), \(\overline{G}(t)/\overline{F}(t)\) is increasing. Therefore Lemma (3.3) is applicable with \(F_1 = F, F_2 = G\). Using the Cauchy-Schwartz inequality and Lemma (3.3) we obtain,

\[
\frac{\mu}{\mu} \int_0^\infty \left(\frac{\overline{F}(t)}{\overline{G}(t)}\right)^2 \overline{G}(t) \text{d}t \leq 1 - \frac{\mu}{\mu} \int_0^\infty \frac{\overline{F}(t)}{\overline{G}(t)} \overline{G}(t)^{1/2} \text{d}t = 1 - \frac{\mu}{\mu} \leq \frac{\mu}{\mu + 1}.
\]

(ii) Follows from (1) and (2.5).

(iii) By (2.1) and Lemma (3.5).

\[
(4.2) \quad d^*(G, \overline{F}) \leq 1 - \int_0^\infty e^{-t/\mu} h_G(t)dt
\]

Since \(h_G\) is decreasing it follows from (2.3) that:

\[
(4.3) \quad \int_0^\infty e^{-t/\mu} h_G(t)dt \geq \int_0^\infty \overline{F}(t)h_G(t)d\tau
\]

The result now follows from (4.2), (4.3) and the inequality.
\[ J(t)h_G(t)dt \geq \frac{1}{\rho + 1}, \] derived in the proof of part (i) of this theorem.

(iv) Follows from (iii) and (2.2).

(v) From (2.2) and (2.5), \( \bar{S}(t) \geq \max(F(t), e^{-t/\mu}) \). The result now follows from Lemma (3.6) and parts (i) and (iii) of this theorem.

(vi) For \( b \leq \mu^{-1} \), \( e^{-tb} \geq e^{-t/\mu} \), thus by (2.2), \( e^{-t/\mu} \leq \min(\bar{S}(t), e^{-bt}) \) for all \( t \geq 0 \). By Lemma (3.3), \( d(E^{-1}, F) \leq 1 - b\mu \), thus by Lemma (3.6) and part (iii) of this theorem:

\[ d(G, E_B) \leq \max\left(\frac{-\rho}{\rho + 1}, 1 - b\mu\right). \]

For \( b \geq \nu^{-1}_G \), \( \frac{\rho}{\rho + 1} = 1 - \nu^{-1}_G \geq 1 - b\mu \), and for \( b \geq \nu^{-1}_G \) the inequality reverses.

(vii) \( \bar{K}(t) = q^{-1}F(t) \) is the survival function of an IMRL distribution with \( \nu_K = q^{-1}\nu_F, \nu_{2K} = q^{-1}\nu_{2F} \) and \( \rho_K + 1 = q(\rho_F + 1) \). Thus from (v):

\[ \sup_t |\bar{K}(t) - e^{-qt/\nu_F}| \leq \frac{\rho_K}{\rho_K + 1} \leq \frac{\rho_F - \eta}{\rho_F + 1} \]

Now multiply both sides of (4.6) by \( q \) to obtain the desired result.

(viii) \( \bar{K}(t) = q^{-1}F(t) \) is the survival function of an IMRL distribution with the same stationary renewal distribution as \( F \) and the same failure rate function as \( F \). Simply apply (iii) using \( \nu_K = q^{-1}\nu_F \) and \( \nu_{2K} = q^{-1}\nu_{2F} \).

(ix) Follows from (viii) and (2.2).

(x) This result follows from (vi) using the distribution \( K \) as in (viii).
(xi) Take a version of \( f \) for which \( f(0) = f(0+) = a \). Set \( L(x) = \frac{f(x)}{f(0)}, \quad x \geq 0 \). Then \( L \) is the survival function of an INRL distribution on \((0, +\infty)\). Moreover \( F \) is the stationary distribution corresponding to \( L \), \( \mu_L = \frac{1}{f(0)} = \frac{1}{a} \), and \( \mu_F = \frac{\mu_L}{2} \). The result now follows from (iii).

(xii) Since \( h_F(t) \simeq h_F(0) = a, \ F(t) \simeq e^{-at} \). The result thus follows from (xi).

(xiii) Follows from (vi) using a similar argument as in (xi).

(4.7) Remark The inequality cited in the summary follows from (i), (iii), (v) and (vi) of Theorem (4.1).

(4.8) Remark If, in addition to \( \mu_1 \) and \( \mu_2 \), \( D = F(0) \) is known then the approximation \( F(t) \simeq ae^{-qt/\mu} \) is suggested. Part (vi) of Theorem (4.1) gives an error bound of \( \frac{qD - D}{D + 1} \) for this approximation.

(4.9) Remark If \( F \) is DFR then its failure rate function, \( h_F(t) \), has a limit as \( t \to +\infty \). Call this limit \( \gamma \). Consider the case \( \gamma > 0 \). Noting that \( h_F \simeq \gamma \) and applying Lemma (3.3) we obtain:

\[
(4.10) \quad a^*(F, E, \gamma) \leq 1 - \gamma u
\]

It is easily shown that \( h_G \) has the same limit as \( h_F \), and thus applying Lemma (3.3) to \( G \) we obtain:

\[
(4.11) \quad c^*(G, E, \gamma) \leq 1 - \gamma u_G
\]

Note that \( 1 - \gamma u_G \leq 1 - \gamma u \). Since \( \ln(e^{\gamma tF(t)}) = \ln(\gamma) + \int_0^t (\gamma - h_F(s))ds \), it follows that \( e^{\gamma tF(t)} \) is decreasing. Call its limit \( c; \ c \geq 0 \) with equality if and only if \( \int_0^\infty (h_F(s) - \gamma)ds = -\infty \).
When $c > 0$, $\gamma > 0$, the approximation $F(t) \approx ce^{-\gamma t}$ will be preferable to $e^{-t/\mu}$ for $t$ sufficiently large.

Note that since $F(t) \approx c_Fe^{-\gamma t}$, $F(t) = c_Fe^{-\gamma t}$, and

$$\lim_{t \to \infty} \frac{F(t)}{G(t)} = \mu \lim_{t \to \infty} h_G(t) = \gamma \mu,$$

it follows that $c_F = \gamma \mu c_G$.

(4.12) **Remark** Let $F$ be a distribution on $[0, \infty)$ with failure rate function $h$. Define $\gamma$ to be the essential infimum of $h$, and assume that $\gamma > 0$. Then by Lemma (3.3):

$$d^*(F, E_\gamma) \leq 1 - \gamma \mu = 1 - \frac{\text{E}(\frac{1}{h(X)})}{\text{sup}(\frac{1}{h(x)})}$$

where $\text{sup}(\frac{1}{h(x)})$ is the essential supremum.

For $b \leq \gamma$ the same argument gives:

$$d^*(F, E_b) \leq 1 - \gamma b \mu \text{ for } b \leq \gamma.$$

Furthermore for $b \geq \gamma$, $e^{-\gamma t} \geq \text{max}(e^{-bt}, F(t))$, thus by Lemma (3.6):

$$d(F, E_b) \leq \begin{cases} 
1 - \gamma b \mu \text{ for } \gamma \leq b \leq \mu^{-1} \\
1 - \gamma b^{-1} \text{ for } b \geq \mu^{-1}
\end{cases}$$

(4.16) **Remark** $F$ is defined to be NHUE (new better than used in expectation) if $E(X - t|X > t) = E[X] < \infty$ for all $t \geq 0$. This is equivalent to $F$ stochastically larger than $G$, where $G$ is the stationary renewal distribution corresponding to $F$. Since $h_G(t) = \text{E}(X - t|X > t)^{-1}$, NHUE implies that $h_G(t) \geq \mu^{-1} = h_G(0)$ for all $t$. Thus $\mu^{-1}$ plays the role of $\gamma$ in remark (4.12). It thus follows from (4.14) and (4.15) that:
(4.17) \[ d^*(G, E_{\mu} - 1) \leq 1 - \frac{\mu_C}{\mu} = \rho \]

(4.18) \[ d^*(G, E_b) \leq 1 - b\mu_C \quad \text{for} \quad b \leq \mu^{-1} \]

(4.19) \[ d(G, E_{b_1}) \leq \begin{cases} 
|\rho| & \text{for} \quad \mu^{-1} \leq b \leq \mu^{-1} \\
1 - (b\mu)^{-1} & \text{for} \quad b \geq \mu^{-1} 
\end{cases} \]

(4.20) **Remark** In Brown [6], Theorem 2, it was shown that if \( Z(t) \) is the forward recurrence time at \( t \) for a renewal process with IMRL interarrival time distribution \( F \), then \( F_{Z(t)} \), the distribution of \( Z(t) \) is stochastically larger than \( F \) and stochastically smaller than \( G \). Thus:

(4.21) \[ \max(d(F, F_{Z(t)}), d(F_{Z(t)}, G)) \leq d(F, G) \leq \frac{\rho}{\rho + 1} \]

Since \( \bar{F}(x) \geq \max(\bar{F}(x), e^{-x/j}) \) it follows from Lemma (3.6), Theorem (4.1) and (4.21) that:

(4.22) \[ d(F_{Z(t)}, E_{\mu} - 1) \leq \max(d(F_{Z(t)}, G), d(E_{\mu} - 1, G)) \]

\[ \leq \max(d(F, G), d(E_{\mu} - 1, G)) \leq \frac{\rho}{\rho + 1} \]

(4.23) **Remark** If \( a < b \) then \( d(E_a, E_b) = (1 - \frac{a}{b}) \left( \frac{a}{b} \right)^{a/b-a} \).

In various bounds derived above (Theorem (4.1) parts (vi), (x) and (xiii), (4.15) and (4.19)) the upper bound, \( 1 - \frac{a}{b} \), rather than the exact distance was used. Thus the results can be somewhat strengthened at the cost of some extra computation. For example part (vi) of theorem (4.1) can be improved as follows:
\[ c(G, F_B) \leq \max \left\{ \frac{\mu - \lambda}{\mu + 1}, (1 - \lambda)(\lambda \mu)/(\lambda \mu + 1) \right\} \text{ for } \lambda \leq \mu^{-1}. \]

(4.24) **Remark** For \( F \) DFR with known first two moments Barlow and Marshall ([2, p. 1266]) derive sharp two sided bounds for each \( t \). Calling their lower and upper bounds \( L(t) \) and \( U(t) \) respectively, the sharpness of their bounds and part (v) of Theorem (4.6) imply:

\[ e^{-t/\mu} - \frac{\mu}{\mu + 1} \leq L(t) \leq F(t) \leq U(t) \leq e^{-t/\mu} + \frac{\mu}{\mu + 1} \]

Therefore:

\[ \sup_t (U(t) - L(t)) \leq \frac{2\mu}{\mu + 1} \]

Remark.

(4.27) Suppose that \( F \) is DFR. The following inequality will be derived:

\[ F(t) \geq e^{-(t+\mu)} \text{ for } t \geq 0 \]

It is known (Barlow and Marshall [27, p. 1267]) that

\[ F(t) \geq e^{-t/\mu} \text{ for } 0 \leq t \leq \mu, \text{ so (4.28) leads to} \]

\[ \frac{\mu}{\mu + 1} \leq F(t) \leq e^{-t/\mu} \text{ for } 0 \leq t \leq \mu. \]

To prove (4.29), let \( m \) denote the renewal density for a renewal process with interarrival time distribution \( F \), \( A(t) \) the renewal age at time \( t \), \( h \) the failure rate function of \( F \), \( p = F'(0) \) and \( q = F(0) \). Since \( M(s) = q^{-1} \mathbb{E}(A(s)) \) with \( h^+ \) and \( A(s) \leq s \), it follows that:

\[ M(s) \geq q^{-1} h(s) \]

Integrate (4.29) from 0 to \( t \) to obtain:
Thus from (4.30):

\[ \int_0^t h(s) \, ds \]

\[ \text{(4.31)} \quad \bar{F}(t) = qe^{-\int_0^t h(s) \, ds} \geq qe^{-\left(\frac{\mu t}{\mu} - 1\right)}. \]

From Brown [6] Lemma 2:

\[ \text{(4.32)} \quad \bar{M}(t) \leq \frac{t}{\mu} + \frac{\mu^2}{2\mu}. \]

Thus (4.31) and (4.32) yield:

\[ \text{(4.33)} \quad qe^{-\left(\frac{\mu t}{\mu} - 1\right)} \geq e^{-\left(\frac{\mu t}{\mu} + \frac{\mu^2}{2\mu}\right)} \]

Finally for DFR with \( F(0) = p, \bar{F}(t) = q\bar{K}(t) \) where \( \bar{K}(t) = q^{-1}\bar{F}(t) \) is DFR. But then \( \rho_{\bar{K}} \geq 0 \) and

\[ \rho_F = q^{-1}\rho_{\bar{K}} + p/q \geq p/q. \]

Thus \( qe^{\rho_{\bar{K}}(t + 1)} \geq qe^{p/q} \geq q(t + p/q) = 1. \)

Therefore the quantity in brackets in (4.33) is at least 1, and (4.26) follows.

5. Sharpness of bounds. Given \( \rho \) and \( \mu \), a convenient IHRL distribution with these parameters is the one with survival function:

\[ \text{(5.1)} \quad \bar{F}(t) = \frac{1}{\rho + 1} e^{-\frac{t}{\rho + 1}}. \]

The stationary distribution \( G \) corresponding to \( F \) is exponential with parameter \( 1/(\rho + 1) \). It follows that:

\[ \text{(5.2)} \quad d^*(F, G) = |\bar{F}(0) - \bar{G}(0)| = \frac{\rho}{\rho + 1} \]

\[ \text{(5.3)} \quad d(F, \bar{F}) = |\bar{F}(0) - 1| = \frac{\rho}{\rho + 1} \]
Since $F$ is INRL, (5.2) and (5.3) demonstrate that the inequalities in parts (i) and (v) of Theorem (4.1) are sharp.

The distribution $F$ is a mixture of two exponentials, one with mean zero (failure rate $\infty$) and the other with failure rate $(\rho + 1)\mu$. The class of completely monotone distributions consists of all mixtures of exponential distributions. It is not clear whether a degenerate distribution at $0$ is allowable as an exponential distribution (with failure rate $\infty$). Thus depending on which definition of completely monotone is employed, either $F$ (defined in (5.1) above) is completely monotone or else is the limit in distribution of a sequence of completely monotone distributions. Either way it follows that $\rho/\rho + 1$ is the best upper bound for $d(F, E^\mu)$ and $d^*(F, G)$ as $F$ ranges through the class of completely monotone distributions.

6. Further results for completely monotone distributions.

It is instructive to look at the bounds of Theorem (4.1) for the class of completely monotone distributions.

Represent a completely monotone random variable $Y$ by $Y = UE = A^{-1}E$ where $U$ and $E$ are independent, $A = U^{-1}$, and $E$ is exponential with parameter $1$. A failure rate of $\infty$ (a mean of 0) corresponds to a degenerate distribution at $0$ while a failure rate of 0 (a mean of $\infty$) corresponds to a degenerate distribution at $\infty$.

Note that:

$$\rho = \frac{\sigma_U^2}{(E_U)^2}.$$  

(6.1)

Thus $\rho$ is the coefficient of variation of the random mean $U$.

Therefore (v) of theorem (4.6) can be expressed as:
\[
(6.2) \quad d(F, \mathcal{E}(\mathcal{E}U)^{-1}) \leq \frac{\sigma_u^2}{\sigma_u^2 + (EU)^2}.
\]

Suppose that \( F(0) = 0 \) and \( f(0+) = EA = EU^{-1} < \infty \). Then (xi) and (xiii) of Theorem (4.1) give:

\[
(6.3) \quad d^n(F, \mathcal{E}(\mathcal{E}A)^{-1}) \leq 1 - (EAEU)^{-1}
\]

\[
(6.4) \quad d(F, \mathcal{E}_b) \leq \begin{cases} 
1 - (EAEU)^{-1} & \text{for } EA \leq b \geq (EU)^{-1} \\
1 - b(EA)^{-1} & \text{for } b \leq (EU)^{-1}
\end{cases}
\]

Finally other bounds for \( d(F, \mathcal{E}_b) \) can be obtained from the following approach:

\[
(6.5) \quad \sup_{\lambda} \left| e^{-\lambda x} dH(\lambda) - e^{-bx} \right| \leq \int d(\mathcal{E}_x, \mathcal{E}_b) dH(\lambda)
\]

\[
\leq E\left(\frac{|A - b|}{\max(A, b)}\right).
\]

Since \( E\left(\frac{|A - b|}{\max(A, b)}\right) \leq \frac{1}{\max(A, b)} - 1 \) we have an alternative deviation for \( d(F, \mathcal{E}_b) \leq E|AB^{-1} - 1| \) derived by Shaked [16]. Furthermore \( d(F, \mathcal{E}_b) \leq E\left(\frac{|A - b|}{\max(A, b)}\right) \leq E|1 - bA^{-1}| = E|1 - bu| \), gives still another upper bound.

7. Proportional hazard families. Consider a family of distributions on \( 0, \infty \) with survival functions:

\[
(7.1) \quad \overline{F}_\lambda(t) = (\overline{F}(t))^\lambda
\]
where \( \lambda > 0 \). If \( F \) is continuous with \( R(t) = -\ln F(t) \), and \( E \) is exponential with parameter 1, then \( Y_\lambda = R^{-1}(E\lambda^{-1}) \) is easily seen to have the distribution \( F_\lambda \) of (7.1).

Suppose now that the parameter \( \lambda \) is random with distribution \( H \). We are interested in bounding the distance between the mixture \( \int F_\lambda \omega(\lambda) \) and \( F_B \) a distribution with fixed parameter. Note that:

\[
\begin{align*}
(7.2) \quad & \sup |Pr(R^{-1}(E_1^{-1}) > t) - Pr(R^{-1}(E_0^{-1}) > t)| \\
& = \sup |Pr(E_\lambda^{-1} > R(t)) - Pr(E_0^{-1} > R(t))| = \sup |F_{E_\lambda^{-1}}(x) - e^{-bx}| \\
& = d(E_\lambda^{-1}, E_0).
\end{align*}
\]

Consequently the desired distance is equivalent to the distance between the completely monotone distribution of \( E_\lambda^{-1} \) and the exponential distribution with parameter \( b \). Therefore the bounds (6.2) - (6.5) apply.

9. **Application to First Passage Times.** Neilsen [11] p. 133 writes "If a system is modeled by a finite Markov chain which is ergodic, the passage time from some specified initial distribution over the state space to a subset \( \mathcal{B} \) of the state space visited frequently is often exponentially distributed to good approximation ..."

For engineering purposes, it is essential to quantify departure from exponentiality via error bounds. When one is dealing with time reversible chains, e.g., systems with independent Markov components, the complete monotonicity present permits quantification and the error bounds needed.

Neilsen's interesting approach defines two special (meaning specific distributions governing the initial state) first passage times, \( T_V \) (the ergodic sojourn time) and \( T_E \) (the ergodic exit time).
For finite state ergodic Markov processes in continuous time
Neilson ([11], Theorem 6.7A) derives:

\[(8.1) \quad \overline{F}_E(t) = \frac{1}{\mu_V} \int_0^t \overline{F}_V(s) \, ds\]

where $\overline{F}_E$, $\overline{F}_V$ are the cdf's of $T_E$, $T_V$ and $\mu_V = ET_V$. Thus $\overline{F}_E$ is
the stationary renewal distribution corresponding to $\overline{F}_V$.

If in addition the process is time reversible $\overline{F}_E$ and $\overline{F}_V$ are
shown to be completely monotone (Neilson [11] Theorem 6.9A) and
thus IMRL.

In view of (8.1) and the fact that $\overline{F}_E$ and $\overline{F}_V$ are IMRL,
Theorem (8.1) is applicable with $F = \overline{F}_V$ and $G = \overline{F}_E$. We thus obtain:

\[(8.2) \quad \max(d^*(\overline{F}_V, \overline{F}_E), d(\overline{F}_V, E^{-1}_\mu, d^*(\overline{F}_E, E^{-1}_\mu), d(\overline{F}_E, F^{-1}_\mu)) \leq \frac{\rho}{\rho + 1} = 1 - \frac{\mu_V}{\mu_E}.\]

Note that the exponential approximation to $\overline{F}_E$ is obtained
without requiring knowledge or finiteness of $ET_E^2$. This latter
quantity is needed to apply the bounds of Neilson [11], Huyse and
Leslie [9], and Hall [7].

\[(8.3) \quad \text{Example. Consider a system with three i.i.d. components.}\]
\[\text{A component alternates between exponential visits to state 1, with} \]
\[\text{parameter } \lambda = .01, \text{ and exponential visits to state 0 with parameter} \]
\[\text{\mu = 1. Of interest is the first passage time to } D = \{(0, 0, 0)\}.\]
\[\text{In the language of reliability theory, we have a repairable three} \]
\[\text{component parallel system with component failure rate .01 and} \]
\[\text{component repair rate 1. The time to first system failure is the} \]
first passage time to \( B \). Since \( B \) is a rarely visited set (the stationary probability of \( B \) is \( (101)^{-3} \)) we anticipate approximate exponentiality for \( T_B, T_V \). Now (Brown [5]) \( E_{T_B} = \)

\[
\frac{1}{1.01} \left( \frac{(101)^3}{(101)^3 - 1} \right) \sum_{r=1}^{3} \frac{(100)^r}{r!} = 345,181.85, \text{ furthermore} \]

\[E_{T_V} = \frac{(101)^3}{3} - 1 = 343,433.33, \text{ thus } \rho = .005001 \text{ and } \frac{\rho}{\rho + 1} = .005066.\]

Therefore by (8.3) \( F_V(t), F_B(t) \) and \( e^{-t/\mu_V} \) are all within a distance of .005066 for all \( t \).
REFERENCES


