CHARACTERIZATION OF BATHTUB DISTRIBUTIONS VIA PERCENTILE RESIDUAL LIFE FUNCTIONS

Alba M. Franco-Pereira*, Rosa E. Lillo, and Juan Romo

Abstract

In reliability theory and survival analysis, many set of data are generated by distributions with bathtub shaped hazard rate functions. Launer (1993) established several relations between the behaviour of the hazard rate function and the percentile residual life function. In particular, necessary conditions were given for a special type of bathtub distributions in terms of percentile residual life functions. The purpose of this paper is to complete the study initiated by Launer (1993) and to characterize (necessary and sufficient conditions) all types of bathtub distributions.

Keywords: Percentile residual life, bathtub hazard rate, aging notions.

* Universidad Carlos III de Madrid, Departament de Statistics, Facultad de Ciencias Sociales y Jurídicas, C/ Madrid 126, 28903 Getafe (Madrid), e-mail: alba.franco@uc3m.es (Alba M. Franco-Pereira).
1 Introduction

Attempts in modeling or summarizing survival data have largely been confined to three major types of distributions: exponential distribution (the constant hazard rate model), increasing hazard rate (IHR) and decreasing hazard rate (DHR) distributions. However, along the years there have been a growing interest in non-monotone hazard rate distributions, particularly bathtub distributions. These distributions offer possibly most natural models for survival times of biological organisms as well as many industrial units or components.

The hazard rate life distributions with bathtub shape, often known simply as bathtub distributions, have a hazard rate curve that resembles to a bathtub shape. There are several variants of the definition of a bathtub shaped hazard rate (see Lai and Xie (2006)). Here we consider the following. Let $X$ be a random variable with a continuous hazard rate function $r_X$. Then $X$ has a bathtub distribution (BT distribution) if there exist $t_1 \leq t_2$ such that,

(i) $r_X(t)$ is strictly decreasing for $t < t_1$,

(ii) $r_X(t)$ is a constant for $t_1 \leq t \leq t_2$, and

(iii) $r_X(t)$ is strictly increasing for $t > t_2$.

Bathtub distributions may offer reasonable models in survival analysis. The initial phase or the infant mortality period shows a very high death rate, due for example to hereditary defects, genetic disorders, infant diseases, or hostile environment. In the context of a manufactured unit, the initial decreasing phase could be motivated by production errors, inferior raw material, errors that escape quality assurance departments, etc. The middle-age group shows deaths mainly due to sudden jolts or accidents. Deaths are mostly met by chance in this period and the hazard rate remains approximately constant during this phase. The final period shows deaths due to actual aging or wearing out, resulting in an increasing hazard rate.

From the customer satisfaction viewpoint, the initial phase of a bathtub hazard rate is unacceptable. It causes ‘death-of-arrival’ products and undetermines customer confidence. It is caused by defects designed into or built into a product. Therefore, to avoid this fact, the product manufacturer must estimate the point at which the hazard rate of a BT distribution attains its minimum, which is the end of ‘burn-in-period’. The determination of the time at which the hazard rate attains the minimum can be also important in fixing product warranty. For example, product burn-in could be used to eliminate the units which fail early, and thus, maximize the reliability of the remaining product.

Launer (1993) establishes several relations between the behavior of the hazard rate function and the percentile residual life function. In particular, he proves that the maximum of the $\alpha$-percentile residual life function precedes in time the minimum of the hazard rate (providing a minimum exists) and that the minimum of the $\alpha$-percentile residual life precedes the maximum of the hazard rate. He also establishes necessary conditions for a special type of BT distributions in terms of percentile residual life functions.

The $\alpha$-percentile residual life function was introduced in Haines and Singpurwalla (1974) as an alternative to the mean residual life function. The mean residual life function is a...
useful tool for analyzing important properties of $X$ when it exists because it characterizes the distribution, however, it may not exist. Even when it exists it may have some practical shortcomings, especially in situations where the data are censored, or when the underlying distribution is skewed or heavy-tailed. In such cases, either the empirical mean residual life function cannot be calculated, or a single long-term survivor can have a marked effect upon it which will tend to be unstable due to its strong dependence on very long durations. The $\alpha$-percentile residual life functions were studied in some detail by Arnold and Brockett (1983), Gupta and Langford (1984), Joe and Proschan (1984), and Joe (1985), as well as by Haines and Singpurwalla (1974). Families of distributions for which simple expressions for the $\alpha$-percentile residual life functions can be obtained, are identified in Raja Rao, Alhumoud, and Damaraju (2006). Besides, Haines and Singpurwalla (1974), Joe and Proschan (1984), and Franco-Pereira, Lillo, and Shaked (2010) studied some aspects of the classes of distribution functions with decreasing $\alpha$-percentile residual life (DPRL($\alpha$)), $\alpha \in (0, 1)$.

Let $X$ be a random variable and let $u_X$ be the right endpoint of its support. If $F_X$ denotes the distribution function of $X$ and $\bar{F}_X = 1 - F_X$ denotes the corresponding survival function, for any $0 < \alpha < 1$, the $\alpha$-percentile residual life function of $X$, denoted by $q_{X,\alpha}$, is defined by

$$q_{X,\alpha}(t) = \bar{F}_X^{-1}(\alpha F_X(t)) - t, \quad t < u_X,$$

where $\overline{\alpha} = 1 - \alpha$. Or, alternatively,

$$q_{X,\alpha}(t) = F_X^{-1}(\alpha + \overline{\alpha} F_X(t)) - t, \quad t < u_X.$$

In this paper we complete the study carried out by Launer (1993) characterizing all types of BT distributions. In Section 2 we introduce new definitions of aging notions based on percentile residual life functions. In Section 3 we prove some results that characterize BT distributions. We will assume that all the variables considered along the paper are nonnegative.

Some conventions that we use in this paper are the following. By “increasing” and “decreasing” we mean “nondecreasing” and “nonincreasing”, respectively. For any distribution function $F$ we let function $F^{-1}$ be the left continuous version of the inverse of $F$, that is

$$F^{-1}(p) = \inf\{x : F(x) \geq p\}, \quad p \in (0, 1).$$

2 New aging notions

Recently, Franco-Pereira, Lillo, and Shaked (2010) derived new properties of the DPRL($\alpha$) aging notion, $\alpha \in (0, 1)$. This notion is based on the monotone behaviour of the $\alpha$-percentile residual life function in the whole support. In this section, we introduce four new definitions of aging notions which are also based on the monotone behaviour of the $\alpha$-percentile residual life function but for all $\alpha \in (0, 1)$ and not necessarily in the whole support. These notions will be used in Section 3 to characterize BT distributions. We also give some characterizations of these concepts in terms of the hazard rate function.
Definition 1. Let $t_0 > 0$. A random variable $X$ is said to be decreasing percentile residual life up to time $t_0$, denoted $t_0$-DPRL, if its $\alpha$-percentile residual life function is decreasing for every $\alpha \in (0, 1)$ and for every $t \leq t_0$. That is,

$$q_{X,\alpha}(t) \geq q_{X,\alpha}(t'), \quad \text{for all } t < t' \leq t_0, \text{ and for all } \alpha \in (0, 1).$$

Analogously, we can define the $t_0$-IPRL aging notion.

Definition 2. Let $t_0 > 0$. A random variable $X$ is said to be increasing percentile residual life from time $t_0$ on, denoted IPRL-$t_0$, if its $\alpha$-percentile residual life function is decreasing for every $\alpha \in (0, 1)$ and for every $t \geq t_0$. That is,

$$q_{X,\alpha}(t) \leq q_{X,\alpha}(t'), \quad \text{for all } t \geq t' \text{, and for all } \alpha \in (0, 1).$$

Analogously, we can define the DPRL-$t_0$ aging notion. Note that, if $X$ is IPRL-$t_0$, it is necessary that $u_X = \infty$.

Some useful equivalent conditions for the $t_0$-DPRL, the $t_0$-IPRL, the IPRL-$t_0$ and the DPRL-$t_0$ notions are given in the following propositions for absolutely continuous random variables with interval support (which may be finite or infinite). For such random variable $X$ we denote by $f_X$ its density function and by $r_X \equiv f_X/F_X$ its hazard rate function.

Proposition 2.1. Let $X$ be an absolutely continuous random variable with interval support $(l_X, u_X)$. Let $t_0 > 0$. The following conditions are equivalent:

(i) $X$ is $t_0$-DPRL($t_0$-IPRL);

(ii) $\bar{\alpha}f_X(t) \leq (\geq)f_X(\bar{F}_X^{-1}(\bar{\alpha}F_X(t)))$, for all $t \in (l_X, t_0]$ and all $\alpha \in (0, 1)$;

(iii) $\bar{\alpha}f_X(\bar{F}_X^{-1}(p)) \leq (\geq)f_X(\bar{F}_X^{-1}(\bar{\alpha}p))$, for all $p \in [\bar{F}_X(t_0), 1)$ and all $\alpha \in (0, 1)$;

(iv) $r_X(t) \leq (\geq)r_X(t + q_{X,\alpha}(t))$, for all $t \in (l_X, t_0]$ and all $\alpha \in (0, 1)$.

Proof. Assume (i). Then $q_{X,\alpha}(t)$ is decreasing in $t \in (l_X, t_0]$ for every $\alpha \in (0, 1)$. Therefore, by differentiating $q_{X,\alpha}$ we see that

$$0 \geq \frac{d}{dt}q_{X,\alpha}(t) = \frac{\bar{\alpha}f_X(t)}{f_X(\bar{F}_X^{-1}(\bar{\alpha}F_X(t)))} - 1,$$

for all $\alpha \in (0, 1)$ and $t \leq t_0$ and (ii) follows. In fact, the proof shows that $(i) \iff (ii)$.

Next assume (ii). Putting there $t = \bar{F}^{-1}(p)$ we obtain (iii). In fact, the proof shows that $(ii) \iff (iii)$.

Finally, assume (ii) again. For $t \in (l_X, t_0]$ divide the left hand side by $\bar{\alpha}F_X(t)$ and the right hand side by $\bar{F}_X(t + q_{X,\alpha}(t))$, which are equal by the definition of percentile residual life function. We obtain

$$r_X(t) \leq f_X(\bar{F}_X^{-1}(\bar{\alpha}F_X(t))) = \frac{f_X(t + q_{X,\alpha}(t))}{\bar{F}_X(t + q_{X,\alpha}(t))},$$

where the last equality follows from the definition of hazard rate function. This gives (iv). In fact, the proof shows that $(ii) \iff (iv)$.

\hfill $\Box$
The equivalence (i)$\iff$(iv) can be found already in Haines and Singpurwalla (1974) and in Joe and Proschan (1984). The difference here is that we consider all $\alpha \in (0, 1)$ and $t \leq t_0$.

From (iv) it is seen that if $r_X$ is increasing (that is, if $X$ has an increasing hazard rate (IHR)) then $X$ is $t_0$-DPRL for any $\alpha \in (0, 1)$ and every $t_0 > 0$.

Analogously, the following proposition holds.

**Proposition 2.2.** Let $X$ be an absolutely continuous random variable with interval support $(l_X, u_X)$. Let $t_0 > 0$. The following conditions are equivalent:

(i) $X$ is IPRL-$t_0$ (DPRL-$t_0$);

(ii) $\overline{\sigma f_X(t)} (t) \geq (\leq) f_X(\overline{F_X}^{-1}(\alpha F_X(t)))$, for all $t \geq t_0$ and all $\alpha \in (0, 1)$;

(iii) $\overline{\sigma f_X(\overline{F_X}^{-1}(p))} \geq (\leq) f_X(\overline{F_X}^{-1}(\alpha p))$, for all $p \in (0, \overline{F_X}(t_0)]$ and all $\alpha \in (0, 1)$;

(iv) $r_X(t) \geq (\leq) r_X(t + q_{X, \alpha}(t))$, for all $t \geq t_0$ and all $\alpha \in (0, 1)$.

3 Characterization of BT distributions

In Launer (1993) some results relating the behavior of the hazard rate function and the percentile residual life function are given. He states and illustrates how those relationships can be useful for studying the behavior of the empirical hazard rate function. In particular, he shows that the maximum of the $\alpha$-percentile residual life function precedes in time the minimum of the hazard rate (providing a minimum exists) and that the minimum of the $\alpha$-percentile residual life function precedes the maximum of the hazard rate. The determination of the time at which the $\alpha$-percentile residual life function is a maximum can be important in fixing product warranty. For example, product burn-in could be used to eliminate the units which fail early, and thus, maximize the reliability of the remaining product.

In this section we complete the study carried out in Launer (1993), providing some new results in terms of the aging notions defined in Section 2. First of all, we introduce the following propositions which show how the conditions $t_0$-DPRL and IPRL-$t_0$ have implications on the behavior of the hazard rate function.

**Proposition 3.1.** Let $t_0 > 0$ and $X$ be an absolutely continuous random variable with hazard rate $r_X$. Then, $X$ is $t_0$-DPRL if, and only if, $r_X(t) \leq r_X(t')$, for all $t < t'$, $t \leq t_0$.

**Proof.** In order to prove the conclusion of the theorem, consider $0 < t \leq t_0 < u_X$ and $t' > t$. Since $X$ is a nonnegative random variable, we can write $t' = t + q_{X, \alpha}(t)$ where $\alpha = 1 - \frac{f_X(t)}{F_X(t)} \in (0, 1)$. That is,

$$t' = t + q_{X, \alpha}(t) = t + \overline{F_X}^{-1}(\alpha \overline{F_X}(t)) - t = \overline{F_X}^{-1}(\alpha \overline{F_X}(t)) = t'. \quad (3.1)$$

Therefore,

$$r_X(t) \leq r_X(t') \iff r_X(t) \leq r_X(t + q_{X, \alpha}(t)) \quad (3.2)$$

and, by Proposition 2.1(iv), the right side of equation (3.2) is equivalent to $X$ being $t_0$-DPRL.

□
Proposition 3.2. Let $t_0 > 0$ and $X$ be an absolutely continuous random variable with hazard rate $r_X$. Then,

(i) $X$ is $t_0$-IPRL if, and only if, $r_X(t) \geq r_X(t')$, for all $t < t'$, $t \leq t_0$.

(ii) $X$ is IPRL-$t_0$ if, and only if, $r_X(t) \geq r_X(t')$, for all $t < t'$, $t \geq t_0$.

(iii) $X$ is DPRL-$t_0$ if, and only if, $r_X(t) \leq r_X(t')$, for all $t < t'$, $t \geq t_0$.

Launer (1993) stated the following result that gives necessary conditions for a special kind of BT distributions.

Theorem 3.3. Let $X$ be an absolutely continuous random variable with hazard rate function $r_X$. Let $t_0 > 0$ be such that $r_X(0) = r_X(t_0)$. If $r_X$ has a bathtub shape, there is a minimum $\alpha = \alpha_0$ for which $q_{X,\alpha}(t)$ is a decreasing function of $t$, for $\alpha > \alpha_0$. For $\alpha \leq \alpha_0$, however, $q_{X,\alpha}(t)$ attains a maximum for some $t > 0$.

We give some results in terms of the aging notions defined on Section 2 and complete the study carried out in Launer (1993). First, let us to introduce the following notation:

$$t_1^* = \max \{ t : X \text{ is } t\text{-DPRL} \},$$
$$t_2^* = \min \{ t : X \text{ is } \text{IPRL}-t \},$$
$$t_3^* = \max \{ t : X \text{ is } t\text{-IPRL} \},$$
$$t_4^* = \min \{ t : X \text{ is } \text{DPRL}-t \}.$$

Then, Theorem 3.3 is a particular case of the following result.

Theorem 3.4. Let $X$ be an absolutely continuous random variable with hazard rate function $r_X$. If $r_X$ has a bathtub shape, then $X$ is DPRL-$t_4^*$ with $t_4^* \in (l_X, u_X)$. Besides, $t_4^*$ is the point where $r_X$ starts increasing.

Proof. Since $X$ is BT, there exits $t_1$ such that $r_X$ is increasing for $t > t_1$. Then, by Proposition 3.2(iii), $X$ is DPRL-$t_4^*$ and $t_4^* = \min \{ t : r_X(t) \text{ is increasing} \}$.

Theorem 3.5. Let $X$ be an absolutely continuous random variable with hazard rate function $r_X$. Let $t_+ > 0$ be such that $r_X(u_X) = r_X(t_+)$. If $r_X$ has a bathtub shape, then $X$ is DPRL-$t_4^*$ and $t_3^*$-IPRL. Besides, $t_4^*$ is the point where $r_X$ starts increasing and $t_3^* = t_+$.

Proof. Since $X$ is BT, there exits $t_2$ such that $r_X$ is strictly increasing for $t > t_2$. Then, by Proposition 3.2(iii), $X$ is DPRL-$t_4^*$ and $t_4^* = \min \{ t : r_X(t) \text{ is increasing} \}$. Besides, there exits $t_1$ such that $r_X$ is strictly decreasing for $t < t_1$. Then, by Proposition 3.2(i), $X$ is $t_3^*$-IPRL.

Obviously, $t_3^* \leq t_4^*$. Besides, since $X$ is $t_3^*$-IPRL, by Proposition ??,

$$r_X(t) \geq r_X(t + q_{X,\alpha}(t)) \text{ for all } t \in (l_X, t_3^*) \text{ and all } \alpha \in (0, 1).$$

In particular, this inequality holds for $t = t_3^*$ and $\alpha = 1$, that is,

$$r_X(t_3^*) \geq r_X(t_3^* + q_{X,1}(t_3^*)) = r_X(u_X) = r_X(t_+). \tag{3.3}$$

And, since $t_3^*$, by definition, is the maximum value that verifies (3.3), this proofs that $t_+ = t_3^*$.

\[ \square \]
Remark 3.6. Notice that $X$ is IHR if, and only if, $t^*_1 = u_X$ and $t^*_4 = l_X$. Analogously, $X$ is DHR if, and only if, $t^*_2 = u_X$ and $t^*_3 = l_X$.

The following results give necessary conditions for upside-down bathtub distributions.

**Theorem 3.7.** Let $X$ be an absolutely continuous random variable with hazard rate function $r_X$. If $r_X$ has a upside-down bathtub shape, then $X$ is IPRL-$t^*_2$ with $t^*_2 \in (l_X, u_X)$. 

6
Theorem 3.8. Let $X$ be an absolutely continuous random variable with hazard rate function $r_X$. Let $t_*>0$ be such that $r_X(u_X) = r_X(t_*)$. If $r_X$ has a upside-down bathtub shape, then $X$ is $t_1^*$-DPRL and $IPRL-t_2^*$.

![Illustration of Theorem 3.7 and Theorem 3.8](image)

Figure 3: Illustration of Theorem 3.7 and Theorem 3.8, respectively

Next result allows us to characterize the BT distributions in terms of the percentile residual life functions.

Theorem 3.9. Let $X$ be an absolutely continuous random variable with hazard rate function $r_X$. Then, $r_X$ is BT if, and only if, $X$ is $DPRL-t_4^*$ and one of the two following conditions holds.

(i) $X$ is not $t_5^*$-IRPL for any $l_X < t < u_X$, and for any $t < t_4^*$, $q'_{X,\alpha}(t) = \frac{\partial}{\partial t}q_{X,\alpha}(t)$ verifies

$$
\begin{cases}
q'_{X,\alpha}(t) < 0 & \text{if } \alpha > 1 - \frac{\bar{F}_X(t_4^*)}{\bar{F}_X(t)}, \\
q'_{X,\alpha}(t) = 0 & \text{if } \alpha = 1 - \frac{\bar{F}_X(t_4^*)}{\bar{F}_X(t)}, \\
q'_{X,\alpha}(t) > 0 & \text{if } \alpha < 1 - \frac{\bar{F}_X(t_4^*)}{\bar{F}_X(t)}. 
\end{cases}
$$

(3.4)

In this case, there exists $t_* > 0$ such that $r_X(0) = r_X(t_*)$.

(ii) $X$ is $t_3^*$-IRPL, and if given $t_3^* < t < t_4^*$, $q'_{X,\alpha}(t) = \frac{\partial}{\partial t}q_{X,\alpha}(t)$ verifies

$$
\begin{cases}
q'_{X,\alpha}(t) < 0 & \text{if } \alpha > 1 - \frac{\bar{F}_X(t_3^*)}{\bar{F}_X(t)}, \\
q'_{X,\alpha}(t) = 0 & \text{if } \alpha = 1 - \frac{\bar{F}_X(t_3^*)}{\bar{F}_X(t)}, \\
q'_{X,\alpha}(t) > 0 & \text{otherwise.}
\end{cases}
$$
Proof. (i) First, since $X$ is DPRL-$t^*_4$, by Proposition 3.2(iii), $r_X(t)$ is increasing from $t^*_4$ on. Now, if we show that $r_X(t)$ is decreasing up to $t^*_4$ we get that $X$ is BT and $r_X(t)$ attains a minimum in $t^*_4$.

Let us to define the function $\alpha(t, \tilde{t}) = 1 - \frac{\bar{F}_X(\tilde{t})}{\bar{F}_X(t)}$. Consider $t < t' < t^*_4$. We can write $t' = t + q_{X,\alpha(t,t')}(t)$, see equation (3.1). Now, since $\bar{F}_X$ is decreasing,

$$\bar{F}_X(t^*_4) < \bar{F}_X(t') \iff 1 - \frac{\bar{F}_X(t')}{\bar{F}_X(t)} < 1 - \frac{\bar{F}_X(t^*_4)}{\bar{F}_X(t)} \iff \alpha(t, t') < \alpha(t, t^*_4).$$

Therefore, $\alpha(t, t') < 1 - \frac{\bar{F}_X(t^*_4)}{\bar{F}_X(t)}$ and, by (3.4), $\frac{\partial}{\partial t} q_{X,\alpha(t,t')}(t) > 0$. That is,

$$\frac{\partial}{\partial t} q_{X,\alpha(t,t')}(t) = \frac{h(t)}{h(t')} - 1 > 0 \iff h(t) > h(t'). \quad (3.5)$$

The last equality can be found, for example, in Launer (1993) (equation (2c)). Therefore, by (3.5), $r_X(t)$ is decreasing up to $t^*_4$.

(ii) First, since $X$ is DPRL-$t^*_4$, by Proposition 3.2(iii), $r_X(t)$ is increasing from $t^*_4$ on. Second, since $X$ is $t^*_3$-IPRL, by Proposition 3.2(i), $r_X(t)$ is decreasing up to $t^*_3$. Now, analogously to part (i), it is straightforward to proof that $r_X(t)$ is decreasing from $t^*_3$ up to $t^*_4$. That is, $X$ is BT.

\[ \square \]

Remark 3.10. The value of $\alpha_0$ in Theorem 3.3 and in Launer (1993), that is the minimum $\alpha$ for which $q_{X,\alpha}(t)$ is a decreasing function of $t$, verifies

$$\alpha_0 = \alpha(0, t^*_4).$$

Additionally, the two following results provide an intuition of the shape of the hazard rate function given some knowledge of the percentile residual life function.

**Theorem 3.11.** Let $X$ be an absolutely continuous random variable with hazard rate function $r_X$. If $X$ is $t^*_1$-DPRL and DPRL-$t^*_4$, then $r_X$ has, at least, one maximum value.

**Theorem 3.12.** Let $X$ be an absolutely continuous random variable with hazard rate function $r_X$. If $X$ is $t^*_3$-IPRL and DPRL-$t^*_4$, then $r_X$ has, at least, one minimum value.

## 4 Conclusions

Launer (1993) established several relations between the behavior of the hazard rate function and the percentile residual life function. In particular, necessary conditions were given for a special type of bathtub distributions in terms of percentile residual life functions.

We have defined several notions of aging based on the monotone behaviour of the percentile residual life functions and that were employed to characterize characterize all types of bathtub distributions. The main contribution of our paper is to complete the study initiated by Launer (1993). First, giving not only necessary conditions for BT distributions but characterization results and, second, characterizing the point at which the failure rate of the bathtub distribution attains its minimum.
Figure 4: Illustration of Theorem 3.11 and Theorem 3.12, respectively

References


