CHAPTER THREE

LOG ODDS RATE

3.1 Introduction

The failure rate/hazard rate is one of the fundamental elements of reliability theory and therefore in many practical situations it has been considered as a useful measure in modeling statistical data to derive appropriate model. Based on the physical properties of the component, the monotone behavior of the failure pattern is also an effective method to identify the underlying model.

Recently, with the need of high reliability of the components, non-monotone hazard or failure rates has also been played an important role in the study of engineering reliability and biological survival analysis. The important distributions such as lognormal, Burr, Inverse Gaussian and truncated normal are appropriate in such situations. The use of odds ratio and proportional odds is becoming more common in the field of reliability or survival analysis when the data exhibits non-proportional hazards (see Kirmani and Gupta (2001)). However, there are certain other situations in which the survival data indicate a

---

2 Some of the results in this Chapter have been published entitled “Characterizations of distributions using log odds rate”, *Statistics*, 41(5), 443-451, see Sunoj, Sankaran and Maya (2007).
non-monotone failure rate, and then the modeling by either proportional hazard or proportional odds may be inappropriate for the description of the situation of failure.

Accordingly, it has been identified recently that log odds rate (LOR) is a useful measure to model statistical data that shows a non-monotone failure rate (see Wang et al. (2003)). A formal definition of LOR is as follows. Let \( X \) be a random variable representing the lifetime of a component/system, \( F(t) \) is the cumulative distribution function (cdf) and \( R(t) = 1 - F(t) \) is the reliability function, then the log odds function is

\[
LO(t) = \ln \frac{F(t)}{R(t)} = \ln F(t) - \ln R(t).
\]

Wang et al. (2003) have shown that the distributions that are non-monotone in terms of failure rate are monotone in terms of LOR in \( \log t \) or \( \log(\log t) \). They established some bounds on reliability based on increasing LOR and characterized logistic distribution in terms of constant LOR.

In view of the usefulness of LOR for modeling statistical data that exhibits non-monotone failure rate, the present chapter focuses attention to examine the relationships between LOR and various reliability measures such as hazard rate and reversed hazard rate in the context of repairable systems. Some families of distributions are characterized and discuss the properties and applications of log odds ratio in weighted models. Further we extend this concept to the bivariate set up and study its properties.

### 3.2 Properties and characterizations

In this section, we discuss some properties of LOR and characterize some families of distributions viz. general family of distributions, Burr, Pearson and log exponential models.

From the definition of log odds function (3.1),

\[
\frac{F(t)}{R(t)} = \exp(LO(t))
\]

or equivalently,
\[ F(t) = \frac{\exp(LO(t))}{1 + \exp(LO(t))}. \]  \( (3.2) \)

Thus the log odds function determines the distribution uniquely through the relation (3.2).

Then log odds rate

\[ \psi(t) = LO'(t) = \frac{f(t)}{F(t)R(t)}. \]  \( (3.3) \)

As mentioned in the previous chapter, reliability and maintainability are important measures to study the effectiveness of systems/components. The major difference between these two measures is that reliability is the probability that a component has survived (or does not fail) in a particular time, whereas maintainability is the probability that required maintenance will be successfully completed in a given time period. Let \( Y \) denotes the repair time of a component and \( \lambda_r(t) \) be the corresponding reversed repair rate. When \( X \) and \( Y \) are independent and identically distributed (i.i.d.) random variables, using the definitions of hazard and reversed repair rate, the LOR (3.3) becomes

\[ \psi(t) = \lambda(t) + h(t). \]  \( (3.4) \)

Therefore \( \psi(t) \) reduces to the sum of reversed repair rate and failure rate. One important property (3.4) posses is that even if the survival data shows a non-monotone failure rate, the log odds rate might be monotone. For various properties of \( \psi(t) \), one could refer to Wang et al. (2003).

Consider a random variable \( X \) with the support of \((a,b)\) with an absolutely continuous cdf \( F(t) \), the system of distributions, introduced by Burr (1942), is given by

\[ f(t) = F(t)(1 - F(t))k(t) \]  \( (3.5) \)

where \( k(t) \) is some convenient function, which must be non-negative in \( 0 \leq F(t) \leq 1 \) and the range of \( X \). The solution to this differential equation, for given \( k(t) \) is obtained as

\[ F(t) = (1 + \exp(-K(t)))^{-1} \]
where \(K(t) = \int_a^t k(u) du\) with \(\lim_{t \to a} K(t) = -\infty\) and \(\lim_{t \to b} K(t) = \infty\). Therefore \(k(t)\) uniquely determine the df. From (3.5), we have

\[
dF(t) \left( \frac{1}{F(t)} + \frac{1}{R(t)} \right) = k(t)dt
\]

i.e.

\[
\lambda(t) + h(t) = k(t).
\]

(3.6)

Equations (3.4) and (3.6) together implies that

\[
\psi(t) = k(t).
\]

Hence for Burr family of distributions, \(k(t)\) directly gives the log odds rate and vice versa.

We now prove a characterization theorem for Pearson family of distributions using the relation connecting IOR and the conditional expectations.

**Theorem 3.1:** Let \(X\) be a rv having an absolutely continuous df \(F(t)\) with the support of \((a, b)\), a subset of the real line. Assume that \(E(X) < \infty\), \(m(t) = E(X \mid X > t)\) and \(\bar{m}(t) = E(X \mid X \leq t)\) denotes the conditional expectations of \(X\). Then the relationship

\[
m(t) = \bar{m}(t) + (c_0 + c_1 t + c_2 t^2)\psi(t)
\]

holds for all \(t \in (a, b)\) if and only if the pdf of \(X\) satisfies the equation (1.30).

**Proof:** The family of distributions (1.30) is characterized by the identity

\[
m(t) = \mu + (c_0 + c_1 t + c_2 t^2)h(t)
\]

where \(\mu = E(X)\) (see Nair and Sankaran (1991)). One can also establish that for the family (1.30),

\[
\bar{m}(t) = \mu - (c_0 + c_1 t + c_2 t^2)\lambda(t)
\]

(see Navarro and Ruiz (2004) and Nair et al. (2005)).
From (3.8) and (3.9), we get

\[ m(t) = \bar{m}(t) + (c_0 + c_1 t + c_2 t^2) \left( h(t) + \lambda(t) \right) \]

which yields (3.7).

Conversely, assume that (3.7) holds, multiplying (3.7) by $F(t)R(t)$ and on simplification we get,

\[ F(t) \int_a^b x f(x) dx = R(t) \int_a^b x f(x) dx + (c_0 + c_1 t + c_2 t^2) f(t) \].

(3.10)

Differentiating (3.10) with respect to $t$, and simplifying we obtain the result (1.30). This completes the proof.

**Examples:** Here we consider some of the important members of the Pearson family and their respective forms (3.7).

1. Normal: $f(t) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left\{ -\frac{1}{2} \left( \frac{t-\mu}{\sigma} \right)^2 \right\}$; $-\infty < t < \infty$, $-\infty < \mu < \infty$, $\sigma > 0$.

Comparing with equation (1.30), we have $c_0 = \sigma^2$, $c_1 = 0$ and $c_2 = 0$. Then equation (3.7) becomes

\[ m(t) = \bar{m}(t) + \sigma^2 \psi(t) \].

2. Beta: $f(t) = \frac{1}{B(a,b)} t^{a-1} (1-t)^{b-1}$; $0 < t < 1$, $a,b > 0$.

Here $c_0 = 0$, $c_1 = \frac{1}{(a+b)}$ and $c_2 = \frac{-1}{(a+b)}$, equation (3.7) yields

\[ m(t) = \bar{m}(t) + \frac{t(1-t)}{(a+b)} \psi(t) \].

3. Gamma: $f(t) = \frac{m^p}{\Gamma(p)} t^{p-1} \exp(-mt)$; $0 < t < \infty$, $m, p > 0$. 

49
In this case, \( c_0 = 0, c_i = \frac{1}{m} \) and \( c_2 = 0 \). Then (3.7) becomes
\[
m(t) = \tilde{m}(t) + \frac{t}{m} \psi(t).
\]

**Theorem 3.2:** The df \( F(t) \) of a rv \( X \) belong to the general family of distributions (1.31) if and only if it satisfies the relationship
\[
m(t) = \tilde{m}(t) + g(t) \psi(t).
\]

**Proof:** For the general family of distributions (1.31), we have
\[
m(t) = \mu + g(t) h(t)
\]
(see Ruiz and Navarro (1994)).

Similarly from (2.25), the right truncated moment function of the family (1.31),
\[
\tilde{m}(t) = \mu - g(t) \lambda(t).
\]
Now (3.12) and (3.13) together implies (3.11). The proof of the converse part is similar to that of the Theorem 3.1.

Next we prove a characterization theorem using \( \psi(t) \) for the one parameter log exponential family. Let \( m_c(t) = E \left( \frac{XC'(X)}{C(X)} \mid X > t \right) \) and \( \bar{m}_c(t) = E \left( \frac{XC'(X)}{C(X)} \mid X < t \right) \)
and \( E \left( \frac{XC'(X)}{C(X)} \right) < \infty \).

**Theorem 3.3:** Assume that \( \lim_{t \to \infty} C(t)t^{\delta+1} = 0 \). Then the distribution of \( X \) belongs to one parameter log exponential family (1.29) if and only if
\[
m_c(t) = \bar{m}_c(t) - t \psi(t).
\]
Proof: For the family (1.29), we have

\[
R(t) = -\frac{C(t)t^{\theta+1}}{A(\theta)(\theta+1)} - \frac{1}{A(\theta)(\theta+1)} \int_t^b C'(x)x^{\theta+1}dx
\]

or

\[
m_c(t) = -th(t) - (\theta + 1). \tag{3.15}
\]

Similarly, one can obtain the df of the log exponential family (1.29) as

\[
F(t) = -\frac{C(t)t^{\theta+1}}{A(\theta)(\theta+1)} - \frac{1}{A(\theta)(\theta+1)} \int_t^b C'(x)x^{\theta+1}dx
\]

or

\[
m_c(t) = t\lambda(t) - (\theta + 1). \tag{3.16}
\]

Combining (3.15) and (3.16), we obtain the required form (3.14). The converse part is straightforward.

3.3 Weighted models

In this section we examine the application of LOR in the context of weighted models. Denoting \( R^w(t) = P(X^w > t) \), the survival function of the weighted rv \( X^w \), then the log odds function denoted by \( LO^*(t) \) is given by

\[
LO^*(t) = \ln \left( \frac{F^w(t)}{R^w(t)} \right) = \ln F^w(t) - \ln R^w(t). \tag{3.17}
\]

But it can be obtained directly from the relations (1.1) and (2.8), as

\[
R^w(t) = \frac{m_c(t)}{\mu_w} R(t) \tag{3.18}
\]

where \( m_c(t) = E\{w(X) | X > t\} \), is the conditional mean of \( w(X) \). From (2.8), (3.1), (3.17) and (3.18), the log odds function becomes
\[ LO^w(t) = \ln \left( \frac{\bar{m}_w(t)}{m_w(t)} \right). \]

The corresponding weighted log odds rate is given by

\[ \psi^w(t) = \frac{d}{dt} \ln \left( \frac{\bar{m}_w(t)}{m_w(t)} \right) = \dot{\lambda}^w(t) + h^w(t) \]

where \( \lambda^w(t) \) and \( h^w(t) \) are the reversed hazard rate and hazard rate of the rv \( X_w \) respectively. Using (1.1), (2.8), (3.3) and (3.18), we obtain

\[ \psi^w(t) = \frac{f^w(t)}{F^w(t)R^w(t)} = \frac{w(t)\mu_w}{\bar{m}_w(t)m_w(t)} \psi(t). \quad (3.19) \]

In view of the form-invariance property for families (1.29) and (1.30), the analogous statements for Theorems 3.1 and 3.3 in the context of weighted models are immediate, which are stated as follows.

**Theorem 3.4:** Let \( X^w \) be a size-biased rv associated to \( X \) with \( w(t) = t^\alpha, \alpha > 0 \). Then the pdf of \( X \) is a member of the Pearson system of distributions (2.48) with \( c_0 = 0 \) and

\[ \lim_{t \to a} (c_2 t^2 + c_1 t + m(t)) f(t) = 0 \quad \text{if and only if} \]

\[ m^S(t) = \bar{m}^S(t) + (v_1 t + v_2 t^2) \psi^S(t) \quad (3.20) \]

where \( m^S(x) = E(X_s | X_s > t), \bar{m}^S(t) = E(X_s | X_s \leq t), \) and \( \psi^S(t) = \frac{t^\alpha \mu_a}{m^S(t)m^w(t)} \).

**Proof:** Under the weight function \( w(t) = t^\alpha \) and \( c_0 = 0 \), the Pearson system of distributions is characterized by the relationship,

\[ m^S(t) = \mu_a + \frac{(k_1 + k_2 t^2)}{1 - 2k_2} \lambda^S(t) \quad (3.21) \]

and

\[ \bar{m}^S(t) = \mu_a - \frac{(k_1 + k_2 t^2)}{1 - 2k_2} \lambda^S(t) \quad (3.22) \]
where $\mu_i = \frac{k_1 - d_i}{1 - 2k_2}$, $1 - 2k_2 \neq 0$. Using (3.21) and (3.22), we obtain the relationship (3.20).

Conversely, assume that (3.20) holds. Then multiplying (3.20) by $R^S(t)F^S(t)$, we get

$$F^S(t) \int_a^b xf^S(x)dx = R^S(t) \int_a^t xf^S(x)dx - \int_a^t xf^S(x)dx + (v_1t + v_2t^2)f^S(t)$$

Differentiating (3.23) with respect to $t$ and on simplification, we obtain

$$[(\mu_i - v_1) - (1 + 2v_2)t]f^S(t) = (v_1t + v_2t^2)f^S(t)$$

which on further simplification, yields (2.48) with $d_i = \left(\frac{v_1 - \mu_i}{1 + 2v_2}\right)$ and $k_i = \frac{v_i}{(1 + 2v_2)}$; $i = 1, 2$ provided $(1 + 2v_2) \neq 0$.

**Theorem 3.5:** Assume that $\lim_{t \to b} c(t)^{\theta + 1} = 0$, with $c(t) = t^\alpha$, $\alpha > 0$, the relationship $m^S(t) = m^S(t) - t\psi^S(t)$

if and only if the pdf of $X_s$ belongs to the one parameter log exponential family (1.29), where

$$m^S(t) = E \left( \frac{X_sC'(X_s)}{C(X_s)} \middle| X_s > t \right), \quad m^S(t) = E \left( \frac{X_sC'(X_s)}{C(X_s)} \middle| X_s < t \right)$$

and

$$E \left( \frac{X_sC'(X_s)}{C(X_s)} \right) < \infty.$$ 

**Proof:** When $c(t) = t^\alpha$, (1.29) becomes $f^S(t) = \frac{t^{\theta + \alpha}C(t)}{A(\theta + \alpha)}$. Since $\mu_i = \frac{A(\theta + \alpha)}{A(\theta)}$, the rest of the proof is similar to the proof of the Theorem 3.3.

### 3.4 Bivariate case

In this section, we extend the concept of log odds function and log odds rate to higher dimensions. We confine our study to the bivariate setup. The extensions to higher
dimensions are direct. Let \( X = (X_1, X_2) \) be a bivariate random vector in the support of \( R^*_i = \{(t_i, t_j) | 0 < t_i < \infty \}; \ i = 1, 2 \) with an absolutely continuous distribution function \( F(t_i, t_j) \) and survival function \( R(t_i, t_j) \) and pdf \( f(t_i, t_j) \). Let \( F_i(t_i) \) and \( R_i(t_i); \ i = 1, 2 \) denote the marginal distribution function and survival function of \( X_i \). Let \( f_i(t_i) \) be the density function of \( X_i \). Then we propose the bivariate log-odds function by

\[
L = LO(t_1, t_2) = \ln \frac{F(t_1, t_2)}{R(t_1, t_2)} = \ln F(t_1, t_2) - \ln R(t_1, t_2)
\]

which gives

\[
\frac{F(t_1, t_2)}{R(t_1, t_2)} = \exp(LO(t_1, t_2)).
\]

The corresponding LOR is defined as a vector

\[
\psi(t_1, t_2) = (\psi_1(t_1, t_2), \psi_2(t_1, t_2))
\]

where

\[
\psi_i(t_1, t_2) = \frac{\partial LO(t_1, t_2)}{\partial t_i}; \ i = 1, 2.
\]

Using the bivariate vector failure rate due to Johnson and Kotz (1975) and bivariate reversed hazard rate due to Roy (2002), (3.27) becomes

\[
\psi_i(t_1, t_2) = \lambda_i(t_1, t_2) + h_i(t_1, t_2)
\]

where \( \lambda_i(t_1, t_2) = \frac{\partial}{\partial t_i} \ln F(t_i, t_j) \) and \( h_i(t_1, t_2) = -\frac{\partial}{\partial t_i} \ln R(t_i, t_j); \ i = 1, 2 \), are the \( i^{th} \) components of the reversed hazard rates and failure rates respectively.

**Examples:** Here we consider some bivariate densities having simple vector valued log odds rate.

1. Bivariate normal:

\[
f(t_1, t_2) = \frac{1}{2\pi \sigma_1 \sigma_2 \sqrt{1 - \rho^2}} \exp \left\{ -\frac{1}{2(1 - \rho^2)} \left[ \frac{t_1^2}{\sigma_1^2} - \frac{2\rho t_1 t_2}{\sigma_1 \sigma_2} + \frac{t_2^2}{\sigma_2^2} \right] \right\}; \quad -\infty < t_1, t_2 < \infty, \sigma_1, \sigma_2 > 0, |\rho| < 1.
\]
Taking logarithm on both sides of (3.29) and differentiating with respect to \( t_j \), we obtain

\[
(1-\rho^2)\sigma_i^2 \frac{\partial f}{\partial t_j} = (\rho \sigma_i t_j - \sigma_j t_j) f, \quad i \neq j = 1, 2 \quad (3.30)
\]

Now integrating (3.30) twice between the limits \( t_i \) to \( b_i \) and \( t_j \) to \( b_j \), \( i \neq j = 1, 2 \), we get

\[
(1-\rho^2)\sigma_i^2 \sigma_j \bar{h}(t_t, t_j) = \sigma_i m_i(t_t, t_j) - \rho \sigma_i m_j(t_t, t_j), \quad i \neq j = 1, 2 \quad (3.31)
\]

where \( m_i(t_t, t_j) = E(X_i|X_i > t_t, X_j > t_j) \), \( i \neq j = 1, 2 \). Similarly integrating (3.30) twice between the limits \( a_i \) to \( t_i \) and \( a_j \) to \( t_j \), \( i \neq j = 1, 2 \), we obtain

\[
(1-\rho^2)\sigma_i^2 \sigma_j \bar{h}(t_i, t_j) = \rho \sigma_i n_i(t_t, t_j) - \sigma_j \bar{m}_i(t_t, t_j), \quad i \neq j = 1, 2 \quad (3.32)
\]

where \( \bar{m}_i(t_t, t_j) = E(X_i|X_i < t_t, X_j < t_j) \), \( i \neq j = 1, 2 \). Now adding (3.31) and (3.32), we get

\[
(1-\rho^2)\sigma_i^2 \sigma_j \psi_i(t_i, t_j) = \sigma_j (m_i(t_t, t_j) - \bar{m}_i(t_t, t_j)) - \rho \sigma_i (\bar{m}_j(t_t, t_j) - m_j(t_t, t_j)), \quad i \neq j = 1, 2 \quad (3.33)
\]

Example 2: Bivariate exponential

The joint density function of the exponential conditional due to Arnold and Strauss (1988) is

\[
f(t_1, t_2) = C \exp\left(-\alpha_1 t_1 - \alpha_2 t_2 - \beta t_1 t_2\right), \quad t_1, t_2, \alpha_1, \alpha_2 > 0, \beta \geq 0 \quad (3.34)
\]

where \( C = -\beta \exp\left[\frac{-\alpha_1 \alpha_2}{\beta} \left(\frac{1}{E_i \left(\frac{-\alpha_1 \alpha_2}{\beta}\right)}\right)^{-1}\right]. \)

Now proceeding in the similar manner as above, the identity connecting the vector valued log odds rate and the conditional moments for (3.33) becomes

\[
\psi_i(t_i, t_j) = \beta (m_j(t_i, t_j) - \bar{m}_j(t_i, t_j)), \quad i = 1, 2 \text{ and } j = 3 - i.
\]

Theorem 3.6: The relationship

\[
LO_{X_i, X_j}(t_i, t_j) = LO_{X_i}(t_i) + LO_{X_j}(t_j) \quad (3.35)
\]

holds for all \( t_i, t_j \), if and only if \( X_1 \) and \( X_2 \) are independent.
**Proof:** Suppose (3.34) holds, then

\[
\frac{F(t_1,t_2)}{R(t_1,t_2)} = \frac{F_1(t_1) F_2(t_2)}{R_1(t_1) R_2(t_2)}
\]

which is equivalent to

\[
F(t_1,t_2) (1 - F_1(t_1))(1 - F_2(t_2)) = F_1(t_1) F_2(t_2) (1 - F_1(t_1) - F_2(t_2) + F(t_1,t_2)).
\]

On simplification, we obtain

\[
F(t_1,t_2) (1 - F_2(t_2) - F_1(t_1)) = F_1(t_1) F_2(t_2) (1 - F_1(t_1) - F_2(t_2))
\]

or

\[
F(t_1,t_2) = F_1(t_1) F_2(t_2).
\]

which proves the result. The converse part is straightforward.

**Remark:** Theorem 3.6 can be useful to test the independence among the variables. This might be helpful in reliability analysis to study the dependence structure between the components of a system.