

## EM Algorithm and Positive Quadrant Dependence in the context of Multi-State Stress-Strength Model

Sanjeev Sabnis\*

*Department of Mathematics, IIT Bombay, India  
sabnissanjeev@gmail.com*

Mayank Jain

*Department of Mathematics, IIT Bombay, India*

Kartik Verma

*Department of Mathematics, IIT Bombay, India*

Received 7 August 2014

Accepted 12 November 2015

This research article is related to a multi-state stress-strength model proposed by Eryilmaz and İşçiöğlü (2011). It deals with (i) demonstration of the use of EM algorithm for the estimation of parameters of distributions of random variables representing strength and two stress levels. The numerical results are illustrated using exponential distribution for each of the random variables under consideration, and (ii) comparison of results under the assumption of the random variables representing two stress levels being independent vis-à-vis they being positively quadrant dependent (PQD). The corresponding numerical results are based on Farlie-Gumbel-Morgenstern PQD distribution having non-identical exponential distributions as marginal distributions, and the distribution of strength variable is also an exponential distribution with a different parameter. As far as the EM algorithm related exercise is concerned, the numerical results, by and large, show that the EM algorithm, the one which makes clever use of data by pretending presence of missing observations, provides efficient estimators of the parameters than those provided by direct use of maximum likelihood (ML) estimators. For the latter exercise related to the PQD property, the numerical results highlight the fact that not only it is incorrect to assume two random variables representing two stress levels to be independent when in fact they are dependent, but the degree of dependence also can not be ignored.

*Keywords:* Multi-State system, EM algorithm, PQD property, Stress-strength model, k-out-of-n system

### 1. Introduction

If two random variables  $X$  and  $Y$  represent stress and strength of a component or a system then stress-strength models involve studying different probabilistic and statistical, estimation and inferential, aspects of  $P(Y > X)$  under various distributional assumptions for  $X$  and  $Y$ . Over the years, these models in the binary set up (state space  $S = \{0, 1\}$ ) have been studied extensively in the literature and a very good reference for the same is a book by Kotz et al. (2003). Eryilmaz and İşçiöğlü(2011) deviated from binary level stress-strength models and proposed the following multi-state level stress-strength model given in (1.1) below. See also Eryilmaz (2008a, 2008b).

---

\*Corresponding author.

Consider a coherent system consisting of  $n$  components. Let  $Y_i$  represent the strength of the  $i^{th}$  component,  $i = 1, 2, \dots, n$ . Assume that  $Y_1, Y_2, \dots, Y_n$  are independent and identically distributed random variables with common cumulative distribution function  $F(x) = P(Y_i \leq x)$ ,  $i = 1, 2, \dots, n$ . Assume that each component and, hence, the system which is made up of these components are subjected to two stress levels. Let these two stress levels be denoted by random variables  $Z_1$  and  $Z_2$  with  $G_1$  and  $G_2$  being their respective cumulative distribution functions. Let

$$X_i = \begin{cases} 0 & \text{if } Y_i < Z_{1:2} \\ 1 & \text{if } Z_{1:2} < Y_i < Z_{2:2} \\ 2 & \text{if } Y_i > Z_{2:2} \end{cases} \quad (1.1)$$

where  $Z_{1:2} = \min(Z_1, Z_2)$  and  $Z_{2:2} = \max(Z_1, Z_2)$ . It may be noted that the random variables  $X_1, X_2, \dots, X_n$  are not independent due to common random stresses.

Using various distributional assumptions and statistical independence between  $Z_1$  and  $Z_2$ , maximum likelihood approach can be employed to estimate various system reliabilities of interest using the invariance property of maximum likelihood principle. However, a closer look at this model proposed by Eryilmaz and İşçioğlu(2011) reveals that it has one limitation, namely, that the state of the  $i^{th}$  component  $X_i$  assumes value 1 when  $Z_{1:2} < Y_i < Z_{2:2}$ , for  $i = 1, 2, \dots, n$ , and thus it is not exactly known as to which stress level is responsible for the  $i^{th}$  component to be in state 1. This can be regarded as a case of missing data and, in fact, one can invoke EM algorithm for estimating the unknown parameters of the distributions of random variables corresponding to strength and two stress levels. Simulation results carried out by nonidentical exponential distributions show that the estimators based on EM algorithm are, by and large, more efficient than the ones based on the classical maximum likelihood estimation method.

Also, in this research article a deviation from the assumption of statistical independence between random variables representing two stress levels has been made to gauge the effect of a particular type of dependence on system reliabilities. Here it is assumed that these two random variables are positively quadrant dependent as it is more realistic to obtain results for this very bivariate notion of dependence. A PQD bivariate Farlie-Gumbel-Morgenstern distribution with exponential marginals is used to illustrate results for the model proposed by Eryilmaz and İşçioğlu(2011). The relevant numerical results do bring out an important point that stringent assumption of dependence such as PQD is more realistic and should not be replaced with that of statistical independence just for the sake of algebraic simplification.

The organization of this paper is as follows: Section 2 contains the preliminaries. Section 3 provides details of the EM algorithm and the corresponding numerical results for the model in (1.1), whereas Section 4 presents the relevant theory for PQD concepts along with numerical results for the same model.

## 2. EM Algorithm application

The model proposed by Eryilmaz and İşçioğlu(2011) has one limitation, namely, that it is not known exactly as to which of the two stress levels is responsible for the  $i^{th}$ ,  $i = 1, 2, \dots, n$ , component to be in state 1. All the model definition given in (1.1) says that the random variable representing strength exceeds only one of the stress levels. In view of this, the information about as to which stress level causes  $i^{th}$  component to be in state 1 can be regarded as missing for each of the  $n$  components. In EM algorithm parlance, observations  $\underline{y} = \{(Y_i, Z_{(1:2)i}, Z_{(2:2)i}), i = 1, 2, \dots, n\}$  would constitute an

incomplete and observable data, whereas  $\underline{\mathbf{x}} = \{(Y_i, Z_{(1:2)i}, Z_{(2:2)i}), \delta_i, i = 1, 2, \dots, n\}$  would constitute a complete and unobservable data. Here  $Z_{(1:2)i} = \min(Z_{1i}, Z_{2i})$  and  $Z_{(2:2)i} = \max(Z_{1i}, Z_{2i})$  and  $\delta_i$  is defined to be equal to 1 if  $Z_{1i} < Y_i < Z_{2i}$  and it is zero otherwise for  $i = 1, 2, \dots, n$ . Under the assumption that (i)  $Y_i \sim \exp(\lambda)$ ,  $Z_{1i} \sim \exp(\theta_1)$  and  $Z_{2i} \sim \exp(\theta_2)$  for  $i = 1, 2, \dots, n$ , and (ii)  $Y_i$ 's are independent, (iii)  $Z_{1i}$ 's and  $Z_{2i}$ 's are statistically independent, the E-step of the EM algorithm entails computing

$$\phi(\psi; \psi^{(k)}) = E_{\psi^{(k)}}[\log L_c(\psi)/\underline{\mathbf{y}}].$$

Here

$$L_c(\lambda, \theta_1, \theta_2/\underline{\mathbf{y}}) = (\lambda^n e^{-\lambda \sum_{i=1}^n y_i}) (\theta_1^n e^{-\theta_1 \sum_{i=1}^n [(z_{1i})^{\delta_i} (z_{2i})^{1-\delta_i}]} \\ (\theta_2^n e^{-\theta_2 \sum_{i=1}^n [(z_{1i})^{1-\delta_i} (z_{2i})^{\delta_i}]})$$

and, thus

$$\ln L_c(\lambda, \theta_1, \theta_2/\underline{\mathbf{y}}) = n \ln \lambda - \lambda \sum_{i=1}^n y_i + n \ln \theta_1 - \theta_1 \sum_{i=1}^n [(z_{1i})^{\delta_i} (z_{2i})^{1-\delta_i}] \\ + n \ln \theta_2 - \theta_2 \sum_{i=1}^n [(z_{1i})^{1-\delta_i} (z_{2i})^{\delta_i}]$$

This involves unobservable data values, namely,  $\delta_i$ 's. Using E-step, it follows using statistical independence of  $Z_1$  and  $Z_2$  that

$$E[\delta_i/\underline{\mathbf{y}}] = P[\delta_i = 1/\underline{\mathbf{y}}] \\ = P[Z_{(1:2)i} = Z_{1i}, Z_{(2:2)i} = Z_{2i}/\underline{\mathbf{y}}] \\ = P[Z_{1i} < y_i < Z_{2i}/\underline{\mathbf{y}}] \\ = P[Z_{1i} < y_i] P[y_i < Z_{2i}] \\ = [1 - e^{-\theta_1 y_i}] [e^{-\theta_2 y_i}]$$

The M-step of the EM algorithm on the  $(k + 1)^{th}$  iteration chooses the values of the parameters that maximize  $\phi(\psi; \psi^{(k)})$ . An excellent reference for EM algorithm is a book by McLachlan and Krishnan (2008).

### 2.1. Simulations

Under the multi-state stress-strength model given in (1.1), a k-out-of-n system is said to be in state  $i$  if at least k out of n components are in state  $i$  and  $R_i(k, n)$  is used to denote the probability that a k-out-of-n system is in state  $i$  or above for  $i = 1, 2$ . It may be noted that the expressions (2.1) and (2.2) given below are appearing on page number 554 of Eryilmaz and İşçioğlu(2011).

$$R_1^{k,n} = \sum_{m=k}^n \binom{n}{m} E[(F(Z_{1:2}))^{n-m} (\bar{F}(Z_{1:2}))^m] \\ = \sum_{m=k}^n \binom{n}{m} \int_0^\infty (1 - e^{-\lambda z})^{n-m} (e^{-\lambda z})^m (\theta_2 e^{-\theta_2 z} e^{-\theta_1 z} + \theta_1 e^{-\theta_1 z} e^{-\theta_2 z}) dz \\ = \sum_{m=k}^n \sum_{i=0}^{n-m} \binom{n}{m} \binom{n-m}{i} (-1)^i \frac{\theta_1 + \theta_2}{(m+i)\lambda + \theta_1 + \theta_2} \tag{2.1}$$

$$\begin{aligned}
 R_2^{k,n} &= \sum_{m=k}^n \binom{n}{m} E[(F(Z_{2:2}))^{n-m}(\bar{F}(Z_{2:2}))^m] \\
 &= \sum_{m=k}^n \binom{n}{m} \int_0^\infty (1 - e^{-\lambda z})^{n-m} (e^{-\lambda z})^m (\theta_1 e^{-\theta_1 z} (1 - e^{-\theta_2 z}) + \theta_2 (1 - e^{-\theta_1 z}) e^{-\theta_2 z}) dz \quad (2.2) \\
 &= \sum_{m=k}^n \sum_{i=0}^{n-m} \binom{n}{m} \binom{n-m}{i} (-1)^i \left[ \frac{\theta_1}{(m+i)\lambda + \theta_1} + \frac{\theta_2}{(m+i)\lambda + \theta_2} - \frac{\theta_1 + \theta_2}{(m+i)\lambda + \theta_1 + \theta_2} \right]
 \end{aligned}$$

$$eR_i(k, n) = \frac{MSE(\widehat{R_i^{EM}(k, n)})}{MSE(\widehat{R_i^{ML}(k, n)})} \quad (2.3)$$

where  $eR_i(k, n)$  denotes the efficiency of EM based reliability of a k-out-of-n system being in state i or above as compared to the one based on ML estimator. Here  $\widehat{R_i^{EM}(k, n)}$  ( $\widehat{R_i^{ML}(k, n)}$ ) is based on parameter estimates obtained using EM algorithm (ML principle) for  $i = 1, 2$ ;  $k$  are 2 and 3 and the corresponding values of  $n$  are 3 and 5.

Under the assumptions that for the  $i^{th}$  component,  $i = 1, 2, \dots, n$ , (i) the strength variable  $Y_i \sim \exp(\lambda)$ , (ii) 1st stress level variable  $Z_{1i} \sim \exp(\theta_1)$ , (iii) 2nd stress level variable  $Z_{2i} \sim \exp(\theta_2)$  and (iv)  $Z_{1i}$  and  $Z_{2i}$  are statistically independent, 2-out-of-3 and 3-out-of-5 systems were simulated 65000 and 40000 times, respectively, for various values of  $\lambda$ ,  $\theta_1$ , and  $\theta_2$  and the values of efficiencies such as  $eR_1(2, 3)$ ,  $eR_2(2, 3)$ ,  $eR_1(3, 5)$  and  $eR_2(3, 5)$  were obtained using the estimates of the expressions indicated in (2.1) and (2.2). The corresponding numerical results are given in Table 2.1.

**Remark 2.1.** It may be noted that this particular simulation exercise is repeated for fewer number of times, namely, 7000 and 4000, respectively, for 2-out-of-3 and 3-out-of-5 systems and related numerical results are reported in Table 2.2.

**Remark 2.2.** This idea of using EM algorithm can be employed for stress-strength models having more than two stress levels as well.

**Remark 2.3.** Table 2.1 and Table 2.2, given on the next page, provide comparison between two estimation procedures, namely, EM algorithm and maximum likelihood estimation. It may be noted that the former one takes into account the missing data in terms of the stress level which is causing the component to be in state 1, while the latter one does not use this information at all. This exercise involves comparing the performance of the cumulative survival probabilities  $R_1(2, 3)$ ,  $R_2(2, 3)$  corresponding to 2-out-of-3 system and  $R_1(3, 5)$ ,  $R_2(3, 5)$  of 3-out-of-5 system using the concept of respective efficiencies. The numerical results show that, by and large, the estimator obtained using EM algorithm tend to outperform those based on ML estimation procedure. This in a way reinforces the fact that clever use of the data at hand does lead to better estimation results.

Table 2.1: Efficiency of EM Algorithm based estimators vs ML Principle based estimators

$\lambda$	$\theta_1$	$\theta_2$	$eR_1(2,3)$	$eR_2(2,3)$	$eR_1(3,5)$	$eR_2(3,5)$
0.1	0.2	0.3	1.5291	1.2038	1.7657	1.2397
0.5	0.2	0.3	0.8571	0.8043	0.8370	0.7633
0.5	0.8	0.3	1.1107	1.0154	1.1303	0.9938
1	3	5	2.1993	1.1433	2.1692	1.1677
1	5	3	1.7699	1.1400	2.2038	1.1635
1	3	3	1.6823	1.1950	2.0259	1.2369
4	5	6	1.2517	1.1502	1.3224	1.1464
3	8	4	1.3889	1.1777	1.5250	1.1945
0.2	1	2	1.9185	1.0237	2.6003	0.9972
2	2	0.2	0.9701	0.8434	0.9653	0.8357

Table 2.2: Efficiency of EM Algorithm based estimators vs ML Principle based estimators

$\lambda$	$\theta_1$	$\theta_2$	$eR_1(2,3)$	$eR_2(2,3)$	$eR_1(3,5)$	$eR_2(3,5)$
0.1	0.2	0.3	1.4738	1.1736	1.6423	1.2004
0.5	0.2	0.3	0.8988	0.8487	0.8823	0.8145
0.5	0.8	0.3	1.0923	1.0119	1.1102	0.9960
1	3	5	1.6546	1.1222	1.9534	1.1420
1	5	3	1.6400	1.1304	1.9375	1.1512
1	3	3	1.5773	1.1676	1.8135	1.1990
4	5	6	1.2308	1.1344	1.2857	1.1341
3	8	4	1.3219	1.1471	1.4219	1.1600
0.2	1	2	1.8214	1.0212	2.3005	1.0041
2	2	0.2	0.9723	0.8421	0.9682	0.8334

### 3. Positive Quadrant Dependent Stress Levels

We give below the definition of two random variables being positively quadrant dependent, its consequence and an example.

**Definition 3.1.** Two random variables  $Z_1$  and  $Z_2$ , with the joint cumulative distribution function  $G(z_1, z_2)$  and the corresponding joint probability density function  $g(z_1, z_2)$ , are said to be positively quadrant dependent (PQD) if and only if

$$P(Z_1 > z_1, Z_2 > z_2) \geq P(Z_1 > z_1)P(Z_2 > z_2)$$

for all  $z_1$  and  $z_2$ .

**Remark 3.1.** The above definition is equivalent to

$$P(Z_1 \leq z_1, Z_2 \leq z_2) \geq P(Z_1 \leq z_1)P(Z_2 \leq z_2)$$

for all  $z_1$  and  $z_2$ .

**Remark 3.2.** Let the marginal cumulative distribution functions of  $Z_1$  and  $Z_2$  be respectively denoted by  $G_{Z_1}$  and  $G_{Z_2}$ . Let the respective marginal probability density functions be given by  $g_{Z_1} = G'_{Z_1}$  and  $g_{Z_2} = G'_{Z_2}$ . Then, for a PQD bivariate distribution, the joint distribution function may be written as

$$G(z_1, z_2) = G_{Z_1}(z_1)G_{Z_2}(z_2) + w(z_1, z_2)$$

where  $w(z_1, z_2)$  satisfies the following conditions: (i)  $w(z_1, z_2) \geq 0$

(ii)  $w(z_1, \infty) \rightarrow 0, w(\infty, z_2) \rightarrow 0, w(z_1, -\infty) = 0, w(-\infty, z_2) = 0$

(iii)  $\frac{\partial^2 w(z_1, z_2)}{\partial z_1 \partial z_2} + g_{Z_1}(z_1)g_{Z_2}(z_2) \geq 0$ .

Note that if both  $Z_1 \geq 0$  and  $Z_2 \geq 0$  then the condition in (ii) may be replaced by  $w(z_1, \infty) \rightarrow 0, w(\infty, z_2) \rightarrow 0, w(z_1, 0) = 0, w(-\infty, 0) = 0$ .

Lai and Xie [2000, 2006] have used these conditions to construct a family of PQD distributions with uniform marginals.

**Example 3.1.** Farlie-Gumbel-Morgenstern bivariate distribution is given by

$$G(z_1, z_2) = G_{Z_1}(z_1)G_{Z_2}(z_2)[1 + \alpha(1 - G_{Z_1}(z_1))(1 - G_{Z_2}(z_2))] \quad (3.1)$$

where  $0 < \alpha \leq 1$ . It is easy to verify that  $Z_1$  and  $Z_2$  are positively quadrant dependent if  $\alpha > 0$ .

When both the marginals  $G_{Z_1}$  and  $G_{Z_2}$  are exponential the joint distribution given in Equation 3.1 acquires the following form:

$$G(z_1, z_2) = (1 - e^{-\theta_1 z_1})(1 - e^{-\theta_2 z_2})[1 + \alpha e^{-\theta_1 z_1 - \theta_2 z_2}], \quad (3.2)$$

for  $0 < \alpha \leq 1$ . Here

$$\begin{aligned} w(z_1, z_2) &= G(z_1, z_2) - G_{Z_1}(z_1)G_{Z_2}(z_2) \\ &= \alpha e^{-\theta_1 z_1 - \theta_2 z_2} (1 - e^{-\theta_1 z_1})(1 - e^{-\theta_2 z_2}) \end{aligned} \quad (3.3)$$

for  $0 < \alpha \leq 1$ . This  $w(z_1, z_2)$  satisfies the conditions (i)-(iii) mentioned above in Remark 3.2 and hence  $Z_1$  and  $Z_2$  are PQD. This distribution is denoted by  $FGM(\theta_1, \theta_2, \alpha)$ .

In the context of Stress-Strength model proposed by Eryilmaz and İşçioğlu(2011), assume that  $Y \sim \exp(\lambda)$  and  $(Z_1, Z_2) \sim FGM(\theta_1, \theta_2, \alpha)$ . With these assumptions, it follows that

$$P[Z_{1:2} \leq z] = 1 - \bar{G}(z, z) = G_1(z) + G_2(z) - G(z, z)$$

Here

$$G(z, z) = (1 - e^{-\theta_1 z})(1 - e^{-\theta_2 z})(1 + \alpha e^{-\theta_1 z - \theta_2 z})$$

$$G_1(z) = 1 - e^{-\theta_1 z}$$

$$G_2(z) = 1 - e^{-\theta_2 z}$$

$$\begin{aligned} P[Z_{1:2} \leq z] &= 1 - e^{-\theta_1 z} + 1 - e^{-\theta_2 z} + (1 - e^{-\theta_1 z})(1 - e^{-\theta_2 z})(1 + \alpha e^{-\theta_1 z - \theta_2 z}) \\ &= 1 - e^{-(\theta_1 + \theta_2)z} - \alpha e^{-(\theta_1 + \theta_2)z} (1 - e^{-\theta_1 z})(1 - e^{-\theta_2 z}) \end{aligned}$$

$$\begin{aligned}
 p_{i,0} &= P[X_i = 0] \\
 &= P[Y_i < Z_{1:2}] \\
 &= \int F(z) dP[Z_{1:2} \leq z] \\
 &= \left[ 1 + \alpha(\theta_1 + \theta_2) \left\{ \frac{1}{\theta_1 + \theta_2} - \frac{1}{2\theta_1 + \theta_2} - \frac{1}{\theta_1 + 2\theta_2} + \frac{1}{2(\theta_1 + \theta_2)} \right\} - \frac{\alpha\theta_1}{2\theta_1 + \theta_2} + \frac{\alpha(\theta_1 + \theta_2)}{2(\theta_1 + \theta_2)} \right. \\
 &\quad \left. - \frac{\alpha\theta_2}{\theta_1 + \theta_2} \right] - \left[ \frac{\theta_1 + \theta_2}{\lambda + \theta_1 + \theta_2} + \alpha(\theta_1 + \theta_2) \left\{ \frac{1}{\lambda + \theta_1 + \theta_2} - \frac{1}{\lambda + 2\theta_1 + \theta_2} - \frac{1}{\lambda + \theta_1 + 2\theta_2} \right. \right. \\
 &\quad \left. \left. + \frac{1}{\lambda + 2(\theta_1 + \theta_2)} \right\} - \frac{\alpha\theta_1}{\lambda + 2\theta_1 + \theta_2} + \frac{\alpha(\theta_1 + \theta_2)}{\lambda + 2(\theta_1 + \theta_2)} - \frac{\alpha\theta_2}{\lambda + \theta_1 + \theta_2} \right]
 \end{aligned}$$

Next,

$$\begin{aligned}
 P[Z_{1:2} \leq z_1, Z_{2:2} \leq z_2] &= G(z_1, z_2) + G(z_2, z_1) - G(z_1, z_1) \\
 &= (1 - e^{-\theta_1 z_1})(1 - e^{-\theta_2 z_2})(1 + \alpha e^{-\theta_1 z_1 - \theta_2 z_2}) \\
 &\quad + (1 - e^{-\theta_1 z_2})(1 - e^{-\theta_2 z_1})(1 + \alpha e^{-\theta_1 z_2 - \theta_2 z_1}) \\
 &\quad - (1 - e^{-\theta_1 z_1})(1 - e^{-\theta_2 z_1})(1 + \alpha e^{-\theta_1 z_1 - \theta_2 z_1})
 \end{aligned}$$

$$\begin{aligned}
 p_{i,1} &= P[X_i = 1] \\
 &= P[Z_{1:2} < Y_i < Z_{2:2}] \\
 &= \int \int_{z_1 < z_2} (F(z_2) - F(z_1)) dP\{Z_{1:2} \leq z_1, Z_{2:2} \leq z_2\} \\
 &= \int_0^\infty \int_0^{z_2} (e^{-\lambda z_1} - e^{-\lambda z_2}) dP\{Z_{1:2} \leq z_1, Z_{2:2} \leq z_2\} \\
 &= \frac{(1 + \alpha)(\theta_1 + \theta_2)}{\lambda + \theta_1 + \theta_2} + \frac{2\alpha(\theta_1 + \theta_2)}{\lambda + 2\theta_1 + 2\theta_2} - \frac{\alpha(2\theta_1 + \theta_2)}{\lambda + 2\theta_1 + \theta_2} - \frac{\alpha(\theta_1 + 2\theta_2)}{\lambda + (\theta_1 + 2\theta_2)} \\
 &\quad - \left[ \frac{(1 + \alpha)\theta_1\theta_2(2\lambda + \theta_1 + \theta_2)}{(\lambda + \theta_1 + \theta_2)(\lambda + \theta_1)(\lambda + \theta_2)} + \frac{(8\alpha)\theta_1\theta_2(\lambda + \theta_1 + \theta_2)}{(\lambda + 2\theta_1 + 2\theta_2)(\lambda + 2\theta_1)(\lambda + 2\theta_2)} \right. \\
 &\quad \left. - 2\alpha\theta_1\theta_2 \left\{ \frac{(2\lambda + 2\theta_1 + \theta_2)}{(\lambda + 2\theta_1 + \theta_2)(\lambda + 2\theta_1)(\lambda + \theta_2)} + \frac{(2\lambda + \theta_1 + 2\theta_2)}{(\lambda + \theta_1 + 2\theta_2)(\lambda + \theta_1)(\lambda + 2\theta_2)} \right\} \right]
 \end{aligned}$$

$$\begin{aligned}
 p_{i,2} &= P[X_i = 2] \\
 &= P[Y_i > Z_{2:2}] \\
 &= \int_0^\infty \bar{F}(z) dP\{Z_{2:2} \leq z\} \\
 &= \int_0^\infty e^{-\lambda z} dP\{Z_{2:2} \leq z\} \\
 &= \frac{\theta_1}{\lambda + \theta_1} + \frac{\theta_2}{\lambda + \theta_2} - \frac{\theta_1 + \theta_2}{\lambda + \theta_1 + \theta_2} - \frac{\alpha(\theta_1 + \theta_2)}{\lambda + \theta_1 + \theta_2} + \frac{\alpha(2\theta_1 + \theta_2)}{\lambda + 2\theta_1 + \theta_2} \\
 &\quad + \frac{\alpha(\theta_1 + 2\theta_2)}{\lambda + \theta_1 + 2\theta_2} - \frac{2\alpha(\theta_1 + \theta_2)}{\lambda + 2\theta_1 + 2\theta_2}
 \end{aligned}$$

$$\begin{aligned}
 R_1^{k,n} &= \sum_{m=k}^n \binom{n}{m} E[(F(Z_{1:2}))^{n-m}(\bar{F}(Z_{1:2}))^m] \\
 &= \sum_{m=k}^n \binom{n}{m} \int_0^\infty (1 - e^{-\lambda z})^{n-m} (e^{-\lambda z})^m dP(Z_{1:2} \leq z) \\
 &= \sum_{m=k}^n \binom{n}{m} \int_0^\infty \left( \sum_{i=0}^{n-m} \binom{n-m}{i} (-e^{-\lambda z})^i \right) (e^{-\lambda z})^m \left[ (\theta_1 + \theta_2) e^{-(\theta_1 + \theta_2)z} \right. \\
 &\quad \left. + \alpha(\theta_1 + \theta_2) \{ e^{-(\theta_1 + \theta_2)z} - e^{-(2\theta_1 + \theta_2)z} - e^{-(\theta_1 + 2\theta_2)z} + e^{-2(\theta_1 + \theta_2)z} \} \right. \\
 &\quad \left. - \{ \alpha\theta_1 e^{-(2\theta_1 + \theta_2)z} - \alpha(\theta_1 + \theta_2) e^{-(2\theta_1 + \theta_2)z} + \alpha\theta_2 e^{-(\theta_1 + 2\theta_2)z} \} \right] dz \\
 &= \sum_{m=k}^n \binom{n}{m} \int_0^\infty \left( \sum_{i=0}^{n-m} \binom{n-m}{i} (-1)^i (e^{-\lambda iz}) \right) (e^{-\lambda mz}) \left[ (\alpha + 1)(\theta_1 + \theta_2) e^{-(\theta_1 + \theta_2)z} \right. \\
 &\quad \left. - \alpha(2\theta_1 + \theta_2) e^{-(2\theta_1 + \theta_2)z} - \alpha(\theta_1 + 2\theta_2) e^{-(\theta_1 + 2\theta_2)z} + 2\alpha(\theta_1 + \theta_2) e^{-2(\theta_1 + \theta_2)z} \right] dz \\
 &= \sum_{m=k}^n \sum_{l=0}^{n-m} \binom{n}{m} \binom{n-m}{l} (-1)^l \left[ \frac{(\alpha + 1)(\theta_1 + \theta_2)}{\lambda l + \lambda m + \theta_1 + \theta_2} - \frac{\alpha(2\theta_1 + \theta_2)}{\lambda l + \lambda m + 2\theta_1 + \theta_2} \right. \\
 &\quad \left. - \frac{\alpha(\theta_1 + 2\theta_2)}{\lambda l + \lambda m + \theta_1 + 2\theta_2} + \frac{2\alpha(\theta_1 + \theta_2)}{\lambda l + \lambda m + 2\theta_1 + 2\theta_2} \right]
 \end{aligned}$$

$$\begin{aligned}
 R_2^{k,n} &= \sum_{m=k}^n \binom{n}{m} E[(F(Z_{2:2}))^{n-m}(\bar{F}(Z_{2:2}))^m] \\
 &= \sum_{m=k}^n \binom{n}{m} \int_0^\infty (1 - e^{-\lambda z})^{n-m} (e^{-\lambda z})^m dP(Z_{2:2} \leq z) \\
 &= \sum_{m=k}^n \binom{n}{m} \int_0^\infty \left( \sum_{i=0}^{n-m} \binom{n-m}{i} (-e^{-\lambda z})^i \right) (e^{-\lambda z})^m [\theta_1 e^{-\theta_1 z} + \theta_2 e^{-\theta_2 z} \\
 &\quad - (\alpha + 1)(\theta_1 + \theta_2) e^{-(\theta_1 + \theta_2)z} + \alpha(2\theta_1 + \theta_2) e^{-(2\theta_1 + \theta_2)z} \\
 &\quad + \alpha(\theta_1 + 2\theta_2) e^{-(\theta_1 + 2\theta_2)z} - 2\alpha(\theta_1 + \theta_2) e^{-2(\theta_1 + \theta_2)z} \\
 &= \sum_{m=k}^n \sum_{l=0}^{n-m} \binom{n}{m} \binom{n-m}{l} (-1)^l \left[ \frac{\theta_1}{\lambda l + \lambda m + \theta_1} + \frac{\theta_2}{\lambda l + \lambda m + \theta_2} - \frac{(\alpha + 1)(\theta_1 + \theta_2)}{\lambda l + \lambda m + \theta_1 + \theta_2} \right. \\
 &\quad \left. - \frac{\alpha(2\theta_1 + \theta_2)}{\lambda l + \lambda m + 2\theta_1 + \theta_2} - \frac{\alpha(\theta_1 + 2\theta_2)}{\lambda l + \lambda m + \theta_1 + 2\theta_2} + \frac{2\alpha(\theta_1 + \theta_2)}{\lambda l + \lambda m + 2\theta_1 + 2\theta_2} \right]
 \end{aligned}$$

A couple of remarks are in order corresponding to Table 3.1 through Table 3.4 given below.

**Remark 3.3.** The survival and state probabilities for both the systems, 2-out-of-3 and 3-out-of-5, when  $\alpha = 0$  coincide with the values reported in Table 1 and Table 2 of Eryilmaz and İşçioğlu(2011).

**Remark 3.4.** For both the systems 2-out-of-3 and 3-out-of-5 systems and for various values of  $\lambda$ ,  $\theta_1$ ,  $\theta_2$ , and  $\alpha$  (which can be taken to represent the degree of positively quadrant dependence), survival probabilities in states 1, 2 and marginal probabilities corresponding to states 0,1, 2 have been tabulated in Table 3.1 and Table 3.2. These values reveal that the survival as well as marginal probabilities are significantly different for two extreme cases, namely,  $\alpha = 0$  (random variables  $Z_1$  and  $Z_2$



representing two stress levels are stochastically independent) and  $\alpha = 1$  (they being positively quadrant dependent). This goes to show that it is improper to disregard dependence structure between  $Z_1$  and  $Z_2$  and pretend that they are stochastically independent even though it lends itself to mathematical simplification. This observation about marginal probabilities becomes more pronounced as the value of  $\alpha$  approaches 1 especially when one considers expected number of components being in state  $j$  ( $j = 0,1,2$ ) for a given large value of  $n$ , the number of components.

### 3.1. Numerical results

Table 3.1: Tabulated values for a 2-out-of-3 system: PQD stresses

	$\alpha$	$R_1^{23}$	$R_2^{23}$	$r_0^{23}$	$r_1^{23}$	$r_2^{23}$	
$\lambda$	0	0.893	0.607	0.107	0.286	0.607	
=0.1	0.1	0.890	0.610	0.110	0.281	0.610	
	0.2	0.888	0.612	0.112	0.276	0.612	
	0.3	0.886	0.614	0.114	0.271	0.614	
$\theta_1$	0.4	0.883	0.617	0.117	0.266	0.617	
	=0.2	0.5	0.881	0.619	0.119	0.261	0.619
		0.6	0.878	0.622	0.122	0.256	0.622
$\theta_2$	0.7	0.876	0.624	0.124	0.252	0.624	
	=0.3	0.8	0.873	0.627	0.127	0.247	0.627
		0.9	0.871	0.629	0.129	0.242	0.629
	1	0.868	0.632	0.132	0.237	0.632	

Table 3.2: Tabulated values for a 2-out-of-3 system: PQD stresses

	$\alpha$	$R_1^{23}$	$R_2^{23}$	$r_0^{23}$	$r_1^{23}$	$r_2^{23}$	
$\lambda$	0	0.933	0.660	0.067	0.274	0.660	
=0.1	0.1	0.932	0.661	0.068	0.271	0.661	
	0.2	0.930	0.662	0.070	0.268	0.662	
	0.3	0.929	0.664	0.071	0.265	0.664	
$\theta_1$	0.4	0.927	0.665	0.073	0.262	0.665	
	=0.2	0.5	0.926	0.667	0.074	0.259	0.667
		0.6	0.925	0.668	0.075	0.256	0.668
$\theta_2$	0.7	0.923	0.670	0.077	0.253	0.670	
	=0.5	0.8	0.922	0.671	0.078	0.250	0.671
		0.9	0.920	0.673	0.080	0.248	0.673
	1	0.915	0.674	0.085	0.240	0.674	

Table 3.3: Tabulated values for a 3-out-of-5 system: PQD stresses

	$\alpha$	$R_1^{23}$	$R_2^{23}$	$r_0^{23}$	$r_1^{23}$	$r_2^{23}$	
$\lambda$	0	0.917	0.619	0.083	0.298	0.619	
=0.1	0.1	0.914	0.622	0.086	0.293	0.622	
	0.2	0.912	0.624	0.088	0.288	0.624	
	0.3	0.909	0.627	0.091	0.283	0.627	
$\theta_1$	0.4	0.907	0.629	0.093	0.278	0.629	
	=0.2	0.5	0.904	0.631	0.096	0.273	0.631
		0.6	0.902	0.634	0.098	0.268	0.634
$\theta_2$	0.7	0.899	0.636	0.101	0.263	0.636	
	=0.3	0.8	0.897	0.639	0.103	0.258	0.639
		0.9	0.894	0.641	0.106	0.253	0.641
	1	0.892	0.644	0.108	0.248	0.644	

Table 3.4: Tabulated values for a 3-out-of-5 system: PQD stresses

	$\alpha$	$R_1^{23}$	$R_2^{23}$	$r_0^{23}$	$r_1^{23}$	$r_2^{23}$	
$\lambda$	0	0.955	0.641	0.045	0.314	0.641	
=0.1	0.1	0.953	0.678	0.047	0.275	0.678	
	0.2	0.952	0.679	0.048	0.273	0.679	
	0.3	0.950	0.680	0.050	0.270	0.680	
$\theta_1$	0.4	0.949	0.682	0.051	0.267	0.682	
	=0.2	0.5	0.948	0.683	0.052	0.265	0.683
		0.6	0.946	0.685	0.054	0.262	0.685
$\theta_2$	0.7	0.945	0.686	0.055	0.259	0.686	
	=0.5	0.8	0.944	0.687	0.056	0.256	0.687
		0.9	0.942	0.689	0.058	0.254	0.689
	1	0.941	0.690	0.059	0.251	0.690	

### Conclusions

In a multi-state stress-strength model, when the lack of knowledge of the stress level, that causes a component to be in an intermediate state, is looked upon as missing data and an EM algorithm is used to estimate the unknown parameters and their parameter estimates are, in turn, used to estimate system reliabilities of interest, this study shows that these reliability estimates, by and large, turn out to be more efficient than their estimates based on the classical method of estimation, namely, maximum likelihood estimation. Another important point that this research article brings out is that the notion of stochastic dependence between two random variables representing two stress levels is more realistic than assuming statistical independence just for the sake mathematical simplification. It may further be noted that results for weaker forms of bivariate dependence like totally positive of order two ( $TP_2$ ), stochastically increasing (SI), right tail increasing (RTI) and left tail decreasing(LTD) etc. can be obtained on similar lines.

## References

- [1] Eryilmaz, S. (2008a) Consecutive k-out-of-n:G system in stress-strength setup. *Commun. Statist. Simul. Computat.* 37:579-589.
- [2] Eryilmaz, S. (2008b) Multivariate stress-strength reliability model and its evaluation for coherent structures. *J. Multivariate Anal.* 99:1878-1887.
- [3] Eryilmaz, S., İşçiöğlü, F. (2011) Reliability Evaluation for a Multi-state System Under Stress-Strength Setup. *Communications in Statistics-Theory and Methods* 40:547-558.
- [4] Kotz, S., Lumelskii, Y., and Pensky, M. (2003). *The Stress-Strength Model and its Generalizations*. World Scientific. Singapore.
- [5] Lai, C. , Xie, M. (2000). A new family of positive dependence bivariate distributions. *Statistics and Probability Letters* 46, 359-364.
- [6] Lai, C. , Xie, M. (2006). *Stochastic Aging and Dependence for Reliability*. Springer, New York.
- [7] McLachlan, G.J., Krishnan, T. (2008) *The EM Algorithm and Extensions*. 2nd Edition, Wiley & Sons, New Jersey.