Discriminating between the generalized Rayleigh and Weibull distributions: Some comparative studies

M. A. Ahmad [a*], M. Z. Raqab[a,b*], D. Kundu[c*]

[a]*Department of Mathematics, The University of Jordan, Amman 11942, JORDAN*
[b]*Department of Statistics and OR, Kuwait University, Safat 13060, Kuwait*
[c]*Department of Mathematics, Indian Institute of Technology Kanpur, Kanpur, Pin 208016, INDIA*

Abstract

The generalized Rayleigh distribution was introduced and studied quite effectively in the literature. The closeness and separation between the distributions are extremely important for analyzing any lifetime data. In this spirit, both the generalized Rayleigh and Weibull distributions can be used for analyzing skewed data sets. In this paper, we compare these two distributions based on the Fisher information measures, and use it for discrimination purposes. It is evident that the Fisher information measures play an important role in separating between the distributions. The total information measures and the variances of the different percentile estimators are computed and presented. A real life data set is analyzed for illustration purposes and a numerical comparison study is performed to assess our procedures in separating between these two distributions.

Keywords: Generalized Rayleigh distribution; Weibull distribution; Percentiles; Information measure matrix.

AMS Subject Classification (2010): 62E15; 62F10; 62E25.

1 Introduction

The generalized Weibull or exponentiated Weibull (EW) was originally proposed by Mudholkar and Srivastava (1993), see also Mudholkar, Srivastava and Freimer (1995). The EW family has the cumulative distribution function (cdf)

$$F(x;\alpha,\beta,\lambda) = (1-e^{-(\lambda x)^{\beta}})^{\alpha},\ x>0,\ \text{and}\ \alpha,\beta,\lambda>0,$$

and probability density function (pdf)

$$f(x;\alpha,\beta,\lambda) = \alpha\beta\lambda^{\beta}x^{\beta-1}e^{-(\lambda x)^{\beta}}(1-e^{-(\lambda x)^{\beta}})^{\alpha-1},\ x>0 \text{ and } \alpha,\beta,\lambda>0.$$

E-mail addresses: admiral451978@yahoo.com (M. A. Ahmad), *mraqab@stat.kuniv.edu* (M. Z. Raqab), *kundu@iitk.ac.in* (D. Kundu).

Here α and β are two shape parameters and λ is a scale parameter. It is observed that the EW family is a very flexible family and it can be used quite effectively for modelling skewed lifetime data. Several well known distributions like exponential ($\alpha = 1$, $\beta = 1$), Weibull ($\alpha = 1$), generalized Rayleigh (GR) ($\beta = 2$), Rayleigh ($\alpha = 1, \beta = 2$) are particular cases of the EW distribution.

Among several other right skewed distributions, the GR and WE distributions are used quite often in analyzing lifetime skewed data. The WE distribution is one of the most popular distributions used for analyzing skewed lifetime data (See, for example, Johnson, Kotz and Balakrishnan (1995, Chapter 21)). It is widely applied in reliability engineering, failure analysis and radar systems to model the dispersion of the received signals level produced by some types of clutters. The cdf and pdf of a Weibull distribution are

$$F_{WE}(x;\beta,\theta) = 1 - e^{-(\theta x)^{\beta}}, \; x > 0, \text{ and } \beta,\theta > 0, \tag{1}$$

and

$$f_{WE}(x;\beta,\theta) = \beta\theta^{\beta}x^{\beta-1}e^{-(\theta x)^{\beta}}, \; x > 0 \text{ and } \beta,\theta > 0, \tag{2}$$

respectively. Here β and θ are the shape and scale parameters, respectively.

The two-parameter GR distribution was introduced by Surles and Padgett (2001). They observed that the GR distribution can be used quite effectively in modelling strength data and also modelling general lifetime data. It has several properties which are quite common to the two-parameter gamma and Weibull distributions. The distribution function and the density function of a GR distribution have closed forms. Due to this, it can be used very conveniently for analyzing censored data set also. Unlike, gamma and Weibull distributions it can have non-monotone hazard function, which can be very useful in many practical applications. The cdf and pdf of the two-parameter GR distribution are of the forms:

$$F_{GR}(x;\alpha,\lambda) = (1 - e^{-(\lambda x)^2})^{\alpha}, \; x > 0, \text{and } \alpha,\lambda > 0, \tag{3}$$

and

$$f_{GR}(x;\alpha,\lambda) = 2\alpha\lambda^2 x e^{-(\lambda x)^2}(1 - e^{-(\lambda x)^2})^{\alpha-1}, \; x > 0 \text{ and } \alpha,\lambda > 0. \tag{4}$$

Here α and λ are two shape and scale parameters, respectively. For further details about the distributional properties and statistical inference of this distribution, one may refer to Ahmad et al. (1997), Raqab and Kundu (2006), Kundu and Raqab (2005, 2007) and the references therein.

From now on, the GR distribution with shape and scale parameters α and λ, respectively, will be denoted by $GR(\alpha,\lambda)$ and the Weibull distribution with shape and scale parameters β and θ, respectively will be denoted by $WE(\beta,\theta)$.

Now, let us consider the following problem. Let $X_1, X_2, ..., X_n$ be a random sample from a GR or WE distribution and we wish to compare these two distributions in terms of Fisher information (FI) measures. It may be mentioned that FI measure essentially describes the amount of information a data set has about the unknown parameters. It plays an important role in the estimation theory. It also serves in inference problems as well as in the interpretation of many physical processes. In this paper we compare the FI matrices of both distributions based on complete as well as censored samples. The total information measure is used to discriminate between the two distribution functions.

The rest of the paper is organized as follows. The FI matrices for complete and censored samples are presented in Sections 2 and 3, respectively. An illustrative example based on a real data set and the numerical comparisons between different discrimination procedures are presented in Section 4. Finally, we conclude the paper in Section 5.

2 FI MEASURE FOR COMPLETE SAMPLE

In this section, we present the FI measures for the WE and GR distributions based on a complete data. Let us denote the FI matrices of WE and GR distributions as follows:

$$I_W(\beta,\theta) = \begin{pmatrix} f_{11W} & f_{12W} \\ f_{21W} & f_{22W} \end{pmatrix} \text{ and } I_G(\beta,\theta) = \begin{pmatrix} f_{11G} & f_{12G} \\ f_{21G} & f_{22G} \end{pmatrix},$$

where

$$f_{11W} = \frac{1}{\beta^2}\left(\psi'(1) + \psi^2(2)\right), \; f_{12W} = f_{21W} = \frac{1}{\theta}\left(1 + \psi(1)\right), \; f_{22W} = \frac{\beta^2}{\theta^2},$$

and

$$f_{11G} = \frac{1}{\alpha^2}, \; f_{12G} = f_{21G} = \frac{2}{\lambda}\int_0^1 (1 + \alpha \ln x)\, \eta(x) x^{\alpha-1}\, dx,$$
$$f_{22G} = \frac{4\alpha}{\lambda^2}\int_0^1 \eta^2(x) x^{\alpha-1}\, dx,$$

where $\psi(x) = \Gamma'(x)/\Gamma(x)$ is the psi(or digamma) function and

$$\eta(x) = 1 + \ln(1-x) - (\alpha - 1)\left(\frac{1}{x} - 1\right)\ln(1-x).$$

We provide the numerical comparison of the two FI matrices based on two different criteria. One criterion is based on the total information measures of the corresponding FI matrices, as suggested by Gupta and Kundu (2006) along the line of the E-optimality in a design of experiment problem. Another related criterion is

Table 1: The TI and TV of FI matrices of $GR(\widetilde{\alpha}, \widetilde{\lambda})$ and $WE(\beta, 1)$ for different β

$\beta \downarrow$	$\widetilde{\alpha}$	$\widetilde{\lambda}$	TI(GR)	TI(WE)	TV(GR)	TV(WE)
0.5	0.2522	0.3608	24.2579	7.5447	0.2185	4.5866
1.0	0.3543	0.4767	14.7338	2.8237	0.3512	1.7166
1.5	0.6397	0.7851	6.7794	3.0605	0.9505	1.8606
2.0	1.0000	1.0000	5.0000	4.4559	2.1402	2.7089
2.5	1.4465	1.1593	4.5811	6.54179	4.5837	3.9769
3.0	1.9931	1.2851	4.6190	9.2026	9.3861	5.5945
3.5	2.6563	1.3891	4.8530	12.3989	18.3675	7.5376
4.0	3.4553	1.4781	5.1859	16.1140	34.4757	9.7961

the sum of the asymptotic variances of the MLEs of the shape and scale parameters (see, for example, Alshunnar et al. (2010)). For comparison purposes, it is more appropriate to consider the total information (TI) and total variance (TV) measures of both the distributions at their closest values. Recently, Raqab(2013) obtained the misclassified GR parameters $\widetilde{\alpha}$ and $\widetilde{\lambda}$ for a given β so that $GR(\widetilde{\alpha}, \widetilde{\lambda})$ is closest to $WE(\beta, 1)$ in terms of the Kullback-Leibler distance. Similarly, it is also obtained the miss-classified WE parameters $\widetilde{\beta}$ and $\widetilde{\theta}$ such that $WE(\widetilde{\beta}, \widetilde{\theta})$ is closest to $GR(\alpha, 1)$.

Now we compute different information measures of the two distribution functions at their closest values. The total information measures as well as the total variances of $GR(\widetilde{\alpha}, \widetilde{\lambda})$ and $WE(\beta, 1)$ are presented in Table 1. Similarly the corresponding values for $WE(\widetilde{\beta}, \widetilde{\theta})$ and $GR(\alpha, 1)$ are also presented in Table 2. From Tables 1 and 2, it is quite interesting to observe that even if the two distribution functions are quite close to each other, the corresponding Fisher information measures can be quite different. From both tables, one can observe the variation of the trace and the total variances of FI matrices of GR and WE distributions. As the shape parameter increases, the absolute value of the difference between traces or between the variances is decreasing-increasing. That is, the absolute difference reaches the minimum difference at some shape parameter value and then starts increasing as α or β increases. In fact, the minimum difference occurs at $\beta = 2$ when the underlying distribution is WE or at $\alpha = 1$ when the underlying distribution is GR. In these cases, both distributions are similar and consequently the trace and the total variance of FI matrices for both distributions are quite close. It is clear that as the value of α (or β) increases, the difference between the traces increases and the amount of information in the WE distribution becomes larger than the GR distribution. But the total variances of the GR distribution is larger than that of the WE distribution as α (or β) increases.

Generally, it is reasonable to adopt some function of the parameters of these distributions for comparison purposes. Here, we use the percentile function as one such function. Hence, we compare the asymptotic variances of the corresponding

Table 2: The TI and TV of FI matrices of $WE(\widetilde{\beta}, \widetilde{\theta})$ and $GR(\alpha, 1)$ for different α

$\alpha \downarrow$	$\widetilde{\beta}$	$\widetilde{\theta}$	TI(GR)	TI(WE)	TV(GR)	TV(WE)
0.5	1.3728	1.4483	6.1273	1.8661	1.0009	2.3797
1.0	2.0000	1.0000	5.0000	4.4559	2.1402	2.7088
1.5	2.4790	0.8520	6.1322	8.7627	4.8279	3.8669
2.0	2.8348	0.7773	7.4829	13.5274	9.2584	4.9687
2.5	3.1171	0.7298	8.8237	18.4317	15.6658	5.9680
3.0	3.3508	0.6964	10.1111	23.3139	24.2667	6.8736
3.5	3.5500	0.6712	11.3381	28.1186	35.2585	7.7010
4.0	3.7235	0.6514	12.5069	32.8059	48.8223	8.4625

percentile estimators. The $100p$-th percentiles $(0 < p < 1)$ of the GR and WE distributions are obtained as

$$P_{GR}(\alpha, \lambda) = \frac{1}{\lambda}\left(-ln(1 - p^{1/\alpha})\right)^{1/2}, \text{ and } P_{WE}(\beta, \theta) = \frac{1}{\theta}\left(-ln(1 - p)\right)^{1/\beta}.$$

Therefore the asymptotic variances of the $100p$-th percentiles estimator of the GR and WE distributions are, respectively,

$$Var_{GR}(p) = \begin{pmatrix} \frac{\partial P_{GR}}{\partial \alpha} & \frac{\partial P_{GR}}{\partial \lambda} \end{pmatrix} \begin{pmatrix} f_{11G} & f_{12G} \\ f_{21G} & f_{22G} \end{pmatrix}^{-1} \begin{pmatrix} \frac{\partial P_{GR}}{\partial \alpha} \\ \frac{\partial P_{GR}}{\partial \lambda} \end{pmatrix},$$

and

$$Var_{WE}(p) = \begin{pmatrix} \frac{\partial P_{WE}}{\partial \beta} & \frac{\partial P_{WE}}{\partial \theta} \end{pmatrix} \begin{pmatrix} f_{11W} & f_{12W} \\ f_{21W} & f_{22W} \end{pmatrix}^{-1} \begin{pmatrix} \frac{\partial P_{WE}}{\partial \beta} \\ \frac{\partial P_{WE}}{\partial \theta} \end{pmatrix}.$$

Now to compare the information measures of the two distributions, we propose to compare the average asymptotic variances of GR and WE distributions defined as follows:

$$AV_{GR}(p) = \int_0^1 Var_{GR}(p)\, dp, \text{ and } AV_{WE}(p) = \int_0^1 Var_{WE}(p)\, dp.$$

Under the GR and WE parent distributions, the average asymptotic variances over all percentiles are presented in Tables 3 and 4.

One can observe that from Table 3, the average asymptotic variance over all percentile estimators for both distributions are some how close when the shape parameter tends to be large and it has a significant difference when the shape parameter is getting small. Generally, while the average asymptotic variance of GR distribution is greater than that of the GR distribution under the WE parent, it becomes smaller under WE parent for large values of the shape parameter. It is not surprising to note that the shape parameter of the underlying distribution plays an important

Table 3: The average asymptotic variances over all percentile estimators under $GR(\alpha, 1)$ and $WE(\beta, 1)$

	GR parent		WE parent	
α or $\beta \downarrow$	GR	WE	GR	WE
0.5	0.3523	0.4391	2.7901	3.9150
1.0	0.3240	0.3480	1.5824	2.7842
1.5	0.2995	0.3047	0.5588	0.7129
2.0	0.2802	0.2823	0.3240	0.3480
2.5	0.2648	0.2686	0.2246	0.2175
3.0	0.2523	0.2588	0.1698	0.1535
3.5	0.2419	0.2514	0.1351	0.1161
4.0	0.2332	0.2454	0.1111	0.0919

role in determining the information amount and then the respective variance available from each distribution. From the closeness and deviation between the two distributions shown in Table 3, it would be interesting to investigate how the the information measure or the average asymptotic variance over all percentile estimators discriminate between WE and GR distributions. The closeness and separation would make a follow-up discussion in Section 4.

3 FI measure for censored sample

In this section, we first provide the FI matrix for the GR distribution and then for the WE distribution based on censored samples. Let us denote the FI matrices for the middle part (T_1, T_2), right censored at T_2 and left censored at T_1 for GR distribution by $I_{MG}(\alpha, \lambda, T_1, T_2)$, $I_{RG}(\alpha, \lambda, T_2)$ and $I_{LG}(\alpha, \lambda, T_1)$, respectively. Analogous notations can be given for the WE distribution. Therefore, the FI for complete sample or for fixed right censored (at time T_2) sample or for fixed left censored (at time T_1) sample with vector of parameters $\boldsymbol{\theta}$ can be obtained as

$$I_M(\boldsymbol{\theta}; 0, \infty), I_M(\boldsymbol{\theta}; 0, T_2) + I_R(\boldsymbol{\theta}; T_2) \text{ and } I_M(\boldsymbol{\theta}; T_1, \infty) + I_M(\boldsymbol{\theta}; T_1),$$

respectively, by observing the fact that $I_R(\boldsymbol{\theta}; \infty) = 0$ and $I_L(\boldsymbol{\theta}; 0) = 0$. The FI matrices for GR distribution based on censored data in terms of the parameters α and λ are described in Theorems 1 and 2 given below. The proofs can be seen easily via differentiation techniques and straight-forward algebra.

Theorem 1: *Let $p_1 = F_{GR}(T_1)$ and $p_2 = F_{GR}(T_2)$. Then the FI matrices for the GR distribution are obtained as follows:*

$$I_{MG}(\alpha, \lambda, T_1, T_2) = \begin{pmatrix} a_{11G} & a_{12G} \\ a_{21G} & a_{22G} \end{pmatrix}, \quad I_{RG}(\alpha, \lambda, T_2) = \begin{pmatrix} b_{11G} & b_{12G} \\ b_{21G} & b_{22G} \end{pmatrix},$$

$$I_{LG}(\alpha,\lambda,T_1)=\begin{pmatrix} c_{11G} & c_{12G} \\ c_{21G} & c_{22G} \end{pmatrix},$$

where

$$\begin{aligned}
a_{11G} &= \frac{1}{\alpha^2}\int_{p_1}^{p_2}(1+\ln y)^2\,dy=\frac{1}{\alpha^2}\left[p_2\left(1+\ln^2 p_2\right)-p_1\left(1+\ln^2 p_1\right)\right],\\
a_{22G} &= \frac{4\alpha}{\lambda^2}\int_{p_1^{1/\alpha}}^{p_2^{1/\alpha}}\eta^2(y)\,y^{\alpha-1}\,dy,\\
a_{12G} &= a_{21G}=\frac{2}{\lambda}\int_{p_1^{1/\alpha}}^{p_2^{1/\alpha}}(1+\alpha\ \ln y)\,\eta(y)\,y^{\alpha-1}\,dy,
\end{aligned}$$

$$\begin{aligned}
b_{11G} &= \frac{p_2^2(\ln p_2)^2}{\alpha^2(1-p_2)},\ b_{22G}=\frac{4\alpha^2}{\lambda^2(1-p_2)}p_2^{\frac{2(\alpha-1)}{\alpha}}\left(1-p_2^{1/\alpha}\right)^2\ln^2\left(1-p_2^{1/\alpha}\right),\\
b_{12G} &= b_{21G}=\frac{2}{\lambda(1-p_2)}p_2^{\frac{2\alpha-1}{\alpha}}\ln p_2\left(1-p_2^{1/\alpha}\right)\left[-ln\left(1-p_2^{1/\alpha}\right)\right],
\end{aligned}$$

$$\begin{aligned}
c_{11G} &= \frac{p_1(\ln p_1)^2}{\alpha^2},\\
c_{22G} &= \frac{4\alpha^2}{\lambda^2}p_1^{(1-\frac{2}{\alpha})}\left(1-p_1^{1/\alpha}\right)^2\ln^2\left(1-p_1^{1/\alpha}\right),\\
c_{12G} &= c_{21G}=\frac{2}{\lambda}p_1^{1-\frac{1}{\alpha}}\ \ln p_1\left(1-p_1^{1/\alpha}\right)\left[-\ln\left(1-p_1^{1/\alpha}\right)\right].
\end{aligned}$$

Proof: Taking the natural logarithm for (4), we get

$$\ln f_{GR}(x;\alpha,\lambda)=\ln\ 2+\ln\ \alpha+2\ln\ \lambda+\ln\ x-\lambda^2x^2+(\alpha-1)\ln(1-e^{-(\lambda\,x)^2}). \quad (5)$$

Differentiating both sides of (5) with respect to α and λ, respectively, we have

$$\frac{\partial\ln f_{GR}(x;\alpha,\lambda)}{\partial\alpha}=\frac{1}{\alpha}+\ln(1-e^{-(\lambda\,x)^2}),$$

and

$$\frac{\partial\ln f_{GR}(x;\alpha,\lambda)}{\partial\lambda}=\frac{2}{\lambda}-2\lambda x^2+(\alpha-1)\frac{2\lambda x^2e^{-(\lambda\,x)^2}}{1-e^{-(\lambda\,x)^2}}.$$

Then

$$a_{11G}=\int_{T_1}^{T_2}\left(\frac{\partial\ln f_{GR}(x;\alpha,\lambda)}{\partial\alpha}\right)^2 f_{GR}(x;\alpha,\lambda)dx,$$

and

$$a_{22G}=\int_{T_1}^{T_2}\left(\frac{\partial\ln f_{GR}(x;\alpha,\lambda)}{\partial\lambda}\right)^2 f_{GR}(x;\alpha,\lambda)dx.$$

On simplifications, a_{11G} and a_{22G} can be represented as

$$a_{11G} = \frac{1}{\alpha^2}\int_{p_1}^{p_2} (1+\ln\ y)^2 dy,$$

and

$$a_{22G} = \frac{4\alpha}{\lambda^2}\int_{p_1^{1/\alpha}}^{p_2^{1/\alpha}} \eta^2(y)\ y^{\alpha-1}dy.$$

Arguments similar to the above techniques, we also conclude the exact expression of a_{12G}. Now, consider the survival function of the GR distribution,

$$\bar{F}_{GR}(x;\alpha,\lambda) = 1-(1-e^{-(\lambda x)^2})^{\alpha}. \tag{6}$$

Differentiating both sides of (6) with respect to α and λ, respectively, we obtain

$$\frac{\partial \bar{F}_{GR}(x;\alpha,\lambda)}{\partial \alpha} = -(1-e^{-(\lambda x)^2})^{\alpha}\ \ln(1-e^{-(\lambda x)^2}),$$

and

$$\frac{\partial \bar{F}_{GR}(x;\alpha,\lambda)}{\partial \lambda} = -2\alpha\lambda x^2 e^{-(\lambda x)^2}(1-e^{-(\lambda x)^2})^{\alpha-1}.$$

Therefore

$$b_{11G} = \frac{1}{\bar{F}_{GR}(T_2;\alpha,\lambda)}\left(\frac{\partial \bar{F}_{GR}(T_2;\alpha,\lambda)}{\partial \alpha}\right)^2,$$

$$b_{22G} = \frac{1}{\bar{F}_{GR}(T_2;\alpha,\lambda)}\left(\frac{\partial \bar{F}_{GR}(T_2;\alpha,\lambda)}{\partial \lambda}\right)^2,$$

and

$$b_{12G} = \frac{1}{\bar{F}_{GR}(T_2;\alpha,\lambda)}\left(\frac{\partial \bar{F}_{GR}(T_2;\alpha,\lambda)}{\partial \alpha}\right)\left(\frac{\partial \bar{F}_{GR}(T_2;\alpha,\lambda)}{\partial \lambda}\right),$$

which leads to the exact expressions of b_{11G}, b_{22G} and b_{12G} in Theorem 1. For the left censoring at T_1, we just replace $\bar{F}(x;\alpha,\lambda)$ appeared in b_{ij}'s by $F(x;\alpha,\lambda)$ and then the exact expressions of c_{11G}, c_{22G} and c_{12G} follow readily. ∇

Based on the hazard rate $h(x) = f(x)/(1-F(x))$, the FI matrix for censored sample can be expressed as follows:

$$I_W(\boldsymbol{\theta}, T_1, T_2) = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

where

$$a_{ij} = \int_{T_1}^{T_2} \left(\frac{\partial}{\partial\theta_i}\ln h(x,\theta)\right)\left(\frac{\partial}{\partial\theta_j}\ln h(x,\theta)\right)\ dx.$$

This turns out that the FI matrix of $WE(\beta,\theta)$ can be displayed in the following theorem.

Theorem 2: *Let* $p_1 = F_{WE}(T_1)$ *and* $p_2 = F_{WE}(T_2)$*. Then the FI matrix of the WE distribution for censored sample is obtained as follows:*

$$I_W(\alpha, \lambda, T_1, T_2) = \begin{pmatrix} a_{11W} & a_{12W} \\ a_{21W} & a_{22W} \end{pmatrix},$$

where

$$a_{11W} = \frac{1}{\beta^2} \int_{-\ln(1-p_1)}^{-\ln(1-p_2)} (1 + \ln u)^2 e^{-u} \, du, \quad a_{22W} = \frac{\beta^2}{\theta^2}(p_2 - p_1),$$

and

$$a_{12W} = a_{21W} = \frac{1}{\theta} \int_{-\ln(1-p_1)}^{-\ln(1-p_2)} (1 + \ln u) e^{-u} \, du.$$

The values of the TI and TV of GR and WE distributions for censored data are computed for three different censoring schemes, precisely:

- scheme I: $p_1 = 0, p_2 = 0.75$
- scheme II: $p_1 = 0.15, p_2 = 0.90$
- scheme III: $p_1 = 0.25, p_2 = 1$

These schemes represent left censored, interval censored and right censored, respectively. The results are presented in Table 4 and Table 5 when the data come from $GR(\alpha, 1)$ and $WE(\beta, 1)$, respectively. From both tables, the WE is preferred distribution in terms of TI's and TV's when the shape parameter (α or β) exceeds the point where both distributions are closest to each other. For example, under GR parent, the TI's of WE distribution exceed the ones of GR distribution when $\alpha > 1$ while under WE parent, the TI's of WE are superior when $\beta > 2$. For $\alpha \leq 1$ or $\beta \leq 2$, this observation is reversed. It is also noticed that the TI's for the GR distribution based on the right censored exceed the TI's based on the left censored in most cases considered. This means that the loss of information becomes higher due to missing data at the beginning. For WE distribution, the Fisher information under right censored data is similar to the one under left censored data and deviations between both TI's vanish as the shape parameter β gets large. The information for WE distribution competes the information for GR distribution when the interval censored data are available for $\alpha > 1$ or $\beta > 2$. This feature is also assured in the sense of the total variance. It can be checked that the TV's for WE distribution become smaller than those for GR distribution as the shape parameter increases.

Table 4: The TI and TV of FI matrices of $WE(\widetilde{\beta},\widetilde{\theta})$ and $GR(\alpha,1)$ for different censoring schemes

$\alpha\downarrow$	$\widetilde{\beta}$	$\widetilde{\theta}$	scheme	TI(GR)	TI(WE)	TV(GR)	TV(WE)
0.5	1.3728	1.4483	I	5.3291	11.1946	1.7022	3.40768
			II	1.2790	1.0333	4.9683	7.9623
			III	5.1271	1.2687	1.2733	9.6923
1.0	2.0000	1.0000	I	3.9983	3.2454	2.8139	4.4134
			II	1.1892	3.1694	10.2396	11.6427
			III	4.7431	3.2803	3.3381	11.9467
1.5	2.4790	0.8520	I	5.0135	6.5091	6.0709	6.4255
			II	1.6004	6.4597	24.1318	17.2255
			III	5.9936	6.5319	8.1248	17.2680
2.0	2.8348	0.7773	I	6.28345	10.0975	11.5273	8.2965
			II	2.0456	10.0597	47.4047	22.3277
			III	7.3592	10.1149	16.1682	22.2575
2.5	3.1171	0.7298	I	7.5631	13.7832	19.4452	9.9830
			II	2.47382	13.7519	81.2632	26.9062
			III	8.6787	13.7976	27.9811	26.7639
3.0	3.3508	0.6964	I	8.8015	17.4511	30.0834	11.5092
			II	2.8775	17.4240	126.8180	31.0418
			III	9.9271	17.4635	44.0203	30.8452
3.5	3.5500	0.6712	I	9.9879	21.0583	43.6819	12.9013
			II	3.2569	21.0342	185.0730	34.8106
			III	11.1047	21.0694	64.6955	34.5696
4.0	3.7235	0.6514	I	11.1219	24.5766	60.4614	14.1815
			II	3.6143	24.5546	256.9360	38.2746
			III	12.2175	24.5866	90.3790	37.9956

Table 5: The TI and TV of FI matrices of $GR(\widetilde{\alpha}, \widetilde{\beta})$ and $WE(\beta, 1)$ for different censoring schemes

$\beta \downarrow$	$\widetilde{\alpha}$	$\widetilde{\lambda}$	scheme	TI(GR)	TI(WE)	TV(GR)	TV(WE)
0.5	0.2522	0.3608	I	24.9965	4.1134	0.5078	5.5938
			II	7.1451	2.8974	1.7654	10.6436
			III	36.4421	4.6720	0.5756	17.0153
1.0	0.3543	0.4767	I	11.6968	1.7315	0.6181	2.3546
			II	2.7985	1.4275	1.8350	5.2439
			III	12.7421	1.8711	0.4466	6.8146
1.5	0.6397	0.7851	I	5.3700	2.1237	1.3698	2.8880
			II	1.4445	1.9886	4.4661	7.3052
			III	6.1670	2.1858	1.3503	7.9606
2.0	1.0000	1.0000	I	3.9983	3.2454	2.8139	4.4134
			II	1.1892	3.1694	10.2396	11.6427
			III	4.7431	3.2803	3.33807	11.9467
2.5	1.4465	1.1593	I	3.7562	4.8445	5.8105	6.5881
			II	1.1824	4.7959	22.7364	17.6178
			III	4.4433	4.8669	7.6240	17.7251
3.0	1.9931	1.2851	I	3.8931	6.8591	11.7347	9.3276
			II	1.2568	6.8253	47.8464	25.0728
			III	4.5193	6.8746	16.3013	25.0371
3.5	2.6563	1.3891	I	4.1911	9.2676	22.8382	12.603
			II	1.36389	9.2428	95.2475	33.9536
			III	4.7552	9.2790	32.8748	33.7941
4.0	3.4553	1.4781	I	4.5694	12.0613	42.7642	16.4022
			II	1.4856	12.0423	180.5390	44.2378
			III	5.0699	12.0701	63.0723	43.9591

4 Illustrative Example and Comparison

4.1 Illustrative Example

In this section, based on a real data set we explain how the different concepts proposed here works in practice. The data set represents the the total seasonal annual rainfall (in inches) appeared at Los Angeles Civic Center during the last 28 years, from 1985 to 2012. The data set is already available on the Los Angeles Civic Website: http://www.laalmanac.com/ weather/we13.htm. For easy reference we present the data set below.

Data Set:

12.82, 17.86, 7.66, 12.48, 8.08, 7.35, 11.99, 21.00, 27.36, 8.11, 24.35, 12.44, 12.40, 31.01, 9.09, 11.57, 17.94, 4.42, 16.42, 9.25, 37.96, 13.19, 3.21, 13.53, 9.08, 16.36, 20.20, 8.69.

Before progressing further we compute some basic statistics of this data set. The mean, variance and the coefficient of skewness are 14.49, 62.59 and 1.203, respectively. Since the data is positively skewed, the GR and WE models may be used to analyze this data set. Recently Raqab (2013) has assessed the goodness of fit of the GR and WE distributions to the seasonal annual rainfall from 1985 to 2009 and it provides a good fit. Adding three more observations, it is observed that both the distributions fit the data reasonably well.

For this new data set, the MLEs of the unknown parameters become $\widehat{\alpha} = 1.0839, \widehat{\lambda} = 0.0622, \widehat{\beta} = 1.9584$, and $\widehat{\theta} = 0.0609$. The Kolmogorov-Smirnov (K-S) distance between the empirical distribution function and the fitted GR ($K-S = 0.1634, p-value = 0.4430$) and fitted Weibull ($K-S = 0.1472, p-value = 0.5785$). Although both the distributions fit the data reasonably well but Weibull is the preferred one. The empirical survival function and the two fitted survival functions are presented in Figure 1. The computer program codes used for all the numerical computations in this paper are available on the website: http://eacademic.ju.edu.jo/mraqab.

The log-likelihood values corresponds to GR and WE distributions are -94.69 and -94.73. Therefore, the ratio of maximized likelihood (RML) is

$$T_n = LogL(\hat{\alpha}, \hat{\lambda}) - L(\hat{\beta}, \hat{\theta}) = -94.69 + 94.73 \approx 0.$$

This turns out that both distributions are very similar in terms of the RML criterion. The estimated Fisher information matrices for the GR and WE models are, respectively,

$$I_G(\hat{\alpha}, \hat{\lambda}) = \begin{pmatrix} 0.8512 & -20.2103 \\ -20.2103 & 1110.08 \end{pmatrix}, \quad \text{and} \quad I_W(\hat{\beta}, \hat{\theta}) = \begin{pmatrix} 0.4755 & 6.9423 \\ 6.9423 & 1034.11 \end{pmatrix}.$$

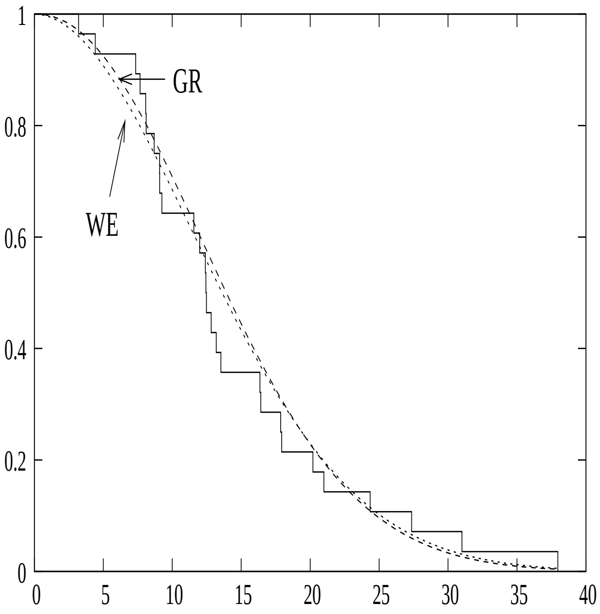

Figure 1: Empirical and fitted survival functions.

The total information measures and the total variance are as follows:

$$TI_G = 1110.9312,\ TI_W = 1034.5855, TV_G = 2.0710,\ TV_W = 2.3327.$$

Further, the average asymptotic variance over all percentile estimators for both distributions are $AV_{GR} = 82.5795,\ AV_W = 98.404$. This concludes that the average asymptotic variance of the percentile estimators for GR distribution is smaller than that of the Weibull distribution. Another point can be considered here is the differences between the variances of the MLEs of shape and scale parameters marginally. For example, the asymptotic variances of MLEs of the shape parameters are $Var(\widehat{\alpha}) = 2.0694, Var(\widehat{\beta}) = 2.3316$, while the asymptotic variances of MLEs of the scale parameters are $Var(\widehat{\lambda}) = 0.0016, Var(\widehat{\theta}) = 0.0011$. It is clear that the asymptotic variance of $\widehat{\alpha}$ is smaller than that of $\widehat{\beta}$ and the asymptotic variance of $\widehat{\lambda}$ is close to the one for $\widehat{\theta}$. This behavior may force us to arise another related question. How does this affect the profile likelihood function? For this purpose, we present plots for the profile likelihoods of α and β for $\lambda = \widehat{\lambda}$ and $\theta = \widehat{\theta}$ in Figures 1 and 2, respectively. Clearly, Figure 1 shows that the profile likelihood of α is higher than the one of β. Moreover, the asymptotic variances of the scale parameters reflect on the similarity of the profile likelihoods of λ and θ as presented in Figure 2.

Now we compare the variance of the pth percentile estimators of both the distributions for various choices of p, Figure 3 shows that the GR distribution has less variances than the WE distribution.

Now let us consider the case where we have a censored data set. For example, we assume $p_1 = 0.15$ and $p_2 = 0.90$. The FI matrices for GR and WE distributions are

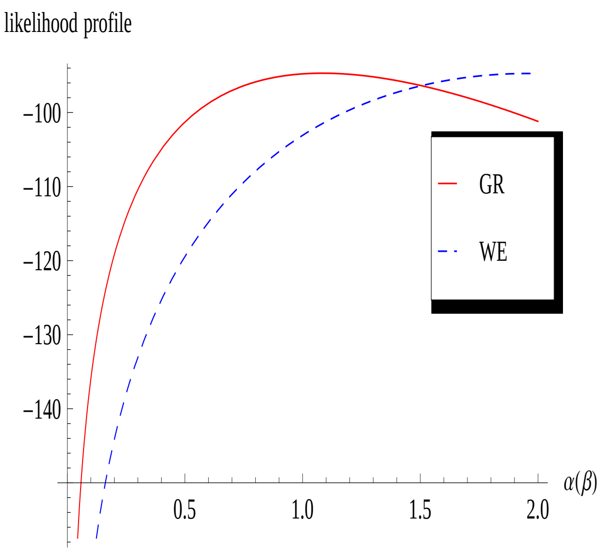

Figure 2: Profile likelihood of α and β when λ and θ are fixed.

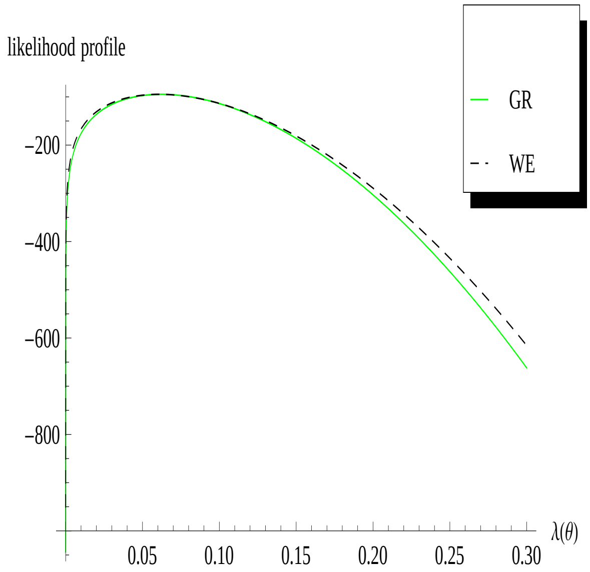

Figure 3: Profile likelihood of λ and θ when α and β are fixed.

computed to be

$$I_G(\hat{\alpha},\hat{\lambda}) = \begin{pmatrix} 0.1874 & -4.9439 \\ -4.9439 & 274.4570 \end{pmatrix} \text{ and } I_W(\hat{\beta},\hat{\theta}) = \begin{pmatrix} 0.1766 & 7.9751 \\ 7.9751 & 775.5850 \end{pmatrix}.$$

respectively. Based on the FI matrices for complete and censored data sets, it is observed the loss of information due to truncation for GR distribution is much more than WE distribution. Another additional tool, we have plotted the asymptotic variances of the pth percentile estimators for censored samples in Figure 3. From Figure 3, it is clear that the loss of information due to truncation for GR is much more than WE.

4.2 Comparison

In this section we carried out an extensive Monte Carlo simulation experiments to see how the proposed Fisher information measures behave compared to other well-known methods for discriminating between the GR and WE distributions. The MLE's of α and λ as well as the MLE's of β and θ are determined using the Newton's Raphson method under the GR and WE parent distributions, respectively. In our case, the information distance between GR and WE distributions is defined as

$$FD_n = TI_G(\widehat{\alpha},\widehat{\lambda}) - TI_W(\widehat{\beta},\widehat{\theta}). \tag{7}$$

Then based on (7), we choose GR or WE distribution as the preferred fit if $FD_n > 0$ or $FD_n < 0$, respectively. Similarly, one can define the distance between the average

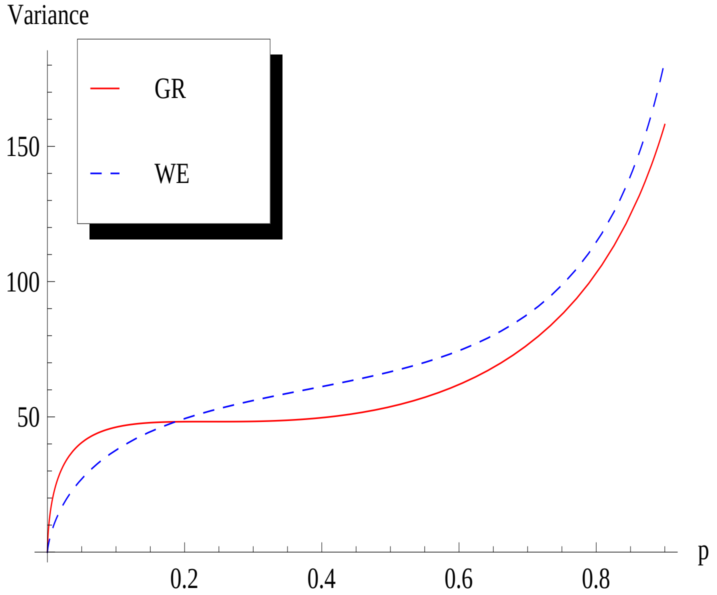

Figure 4: The variances of the pth percentile estimators for GR and WE distributions for complete sample.

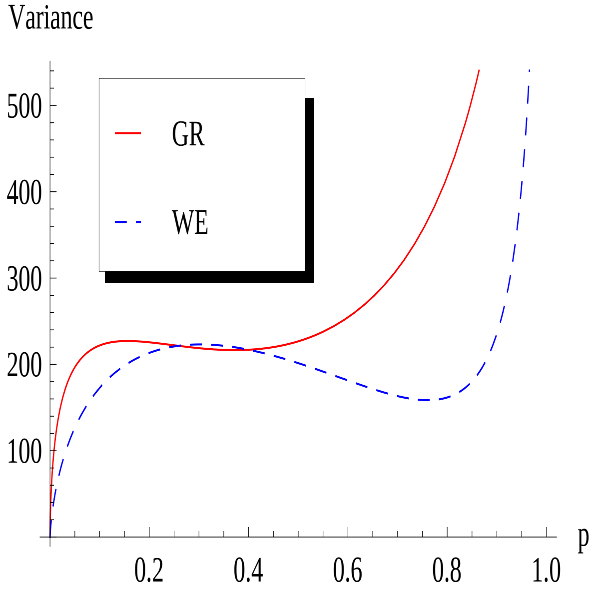

Figure 5: The variances of the pth percentile estimators for GR and WE distributions for censored sample.

asymptotic variances of GR and WE distributions as

$$AVD_n = \int_0^1 Var_{GR}(p)\, dp - \int_0^1 Var_{WE}(p)\, dp. \tag{8}$$

On the basis of (8), the GR or WE is the preferred fitted distribution if $AVD_n < 0$ or $AVD_n > 0$. For a particular sample size n, we generate samples from the GR and WE with different shape parameters. In each case, we compute the MLE's and substitute the respective values of the parameters into FD_n and AVD_n in (7) and (8). We replicate the process M=1000 times and compute the probability of correct selection (PCS) in each case. Under the assumption that the true distribution is $GR(\alpha, 1)$, the approximate PCS based on Fisher information and average asymptotic variance can be described as follows:

$$PCS_{FD} = P(FD_n > 0) \approx \frac{\#\text{ of } FD_n \text{ values} > 0}{M},$$

and

$$PCS_{AVD} = P(AVD_n < 0) \approx \frac{\#\text{ of } AVD_n \text{ values} < 0}{M}.$$

The PCSs based on the RML and K-S distance are also computed and compared with the respective PCSs values obtained using Fisher information and average asymptotic variance of the pth percentiles discussed here. The results, up to 2 decimal places, are reported in Tables 6 and 7. It is worth mentioning that under GR (WE) parent, the shape parameter value $\alpha = 1 (\beta = 2)$ is a critical value, respectively. At these values, both distributions are close in terms of Kullback-Leibler distance. From Table 6, it can be seen, from these numerical results that the Fisher information and asymptotic variance procedures outperform RML and K-S distances in discriminating between the two distributions when α is below 1. Interestingly, the discrimination procedure using the variance of the median method outperforms RML or K-S procedure for all parameter values. In addition, the results in Table 7, show that the Fisher information and asymptotic variance do quite well when compared to RML and K-S distances with shape parameter values β tend to be large ($\beta > 2$). For $\beta < 2$, the result is reversed. The performance based on the variance of the median outperforms other method when the shape parameter of the Weibull distribution is less than one. This assures that there are situations where Fisher information and asymptotic variance criteria work effectively in discriminating between both distributions. We have the following recommendations in discriminating between Weibull and GR distribution. If the empirical evidence, say for example scaled TTT plot, indicates that the parent distribution has a decreasing or bathtub shaped hazard function, then use variance of the median method, otherwise use RML. The natural question does arise: what happens if we want to discriminate between other types of distributions. More work is needed along that direction.

5 Conclusions

In this article, we have compared the Fisher information matrices of GR and WE distributions. The asymptotic variances as well as the average asymptotic variances of the percentile estimators are also computed and compared for these two distributions. We have proposed to use the Fisher information measures to discriminate between these two distributions. It is observed that the proposed method works very well in general, and it outperforms some of the existing methods for certain cases.

Acknowledgement: The authors are thankful to the referees for their valuable suggestions on the previous version of this manuscript.

References

[1] Ahmad, K.E., Fakhry, M.E. and Jaheen, Z.F. (1997), Empirical Bayes estimation of $P(X < Y)$ and characterization of Burr Type X model, *Journal of Statistical Planning and Inference* 64, 297 - 308.

[2] Aarset, M. V. (1987). "How to identify a bathtub hazard rate", *IEEE Transactions on Reliability* 36, 106 - 108.

[3] Alshunnar, F.S., Raqab, M.Z. and Kundu, D. (2010). On the comparison of the Fisher information of the log-normal and Weibull distributions, *Journal of Applied Statistics* 37, 391 - 404.

[4] Bain, L.J. and Engelhardt, M. (1991). *Statistical Analysis of Reliability and Life-Testing Models*, 2nd. Edition, Marcel and Decker, New York.

[5] Dumonceaux, R. and Antle, C. E. (1973). Discrimination between the log-normal and the Weibull distributions, *Technometrics* 15, 923-926.

[6] Gupta, R.D. and Kundu, D. (2006). On comparison of the Fisher information of the Weibull and GE distributions, *Journal of the Statistical Planning and Inference*, 136, 3130 - 3144.

[7] Johnson, N.L., Kotz, S. and Balakrishnan, N. (1995). *Continuous Univariate Distribution Vol. 1*, 2nd Ed., New York, Wiley.

[8] Kundu, D. and Raqab, M. Z. (2005). Generalized Rayleigh distribution: different methods of estimation, *Computational Statistics & Data Analysis* 49, 187-200.

[9] Kundu, D. and Raqab, M. Z. (2007). Discriminating between the generalized Rayleigh and log-normal distribution, *Statistics* 41, 505-515.

[10] Mudholkar, G.S. and Srivastava, D.K. (1993). Exponentiated Weibull family for analyzing bathtub failure data, *IEEE Transactions on Reliability* 42, 299-302.

[11] Mudholkar, G.S., Srivastava, D.K. and Freimer, M. (1995). The exponentiated Weibull family: A re-analysis of the bus motor failure data, *Technometrics* 37, 436-445.

[12] Pascual, F. G. (2005). Maximum likelihood estimation under misspecified log-normal and Weibull distributions, *Communications in Statistics- Simulation & Computations* 34, 503-524.

[13] Raqab, M. Z. (2013). Discriminating between the generalized Rayleigh and Weibull distributions, *Journal of Applied Statistics* 40(7), 1480-1493.

[14] Raqab, M. Z. and Kundu, D. (2006). Burr type X distribution: revisited, *Journal of Probability and Statistical Sciences* 4, 179 - 193.

[15] Surles, J.G. and Padgett, W.J. (2001). Inference for reliability and stress-strength for a scaled Burr Type X distribution, *Lifetime Data Analysis*, vol. 7, 187-200.

[16] Yu, H. F. (2007). Mis-specification analysis between normal and extreme value distributions for a linear regression model, *Communications in Statistics-Theory & Methods* 36, 499-521.

Table 6: The simulated PCSs when the null distribution is GR.

		Disc. Proc.				
α	n	RML	K-S	FI	AV	VM
0.5	20	0.68	0.40	0.99	0.97	0.76
	30	0.68	0.45	1.00	0.99	0.76
	40	0.72	0.47	1.00	1.00	0.72
	50	0.74	0.51	1.00	1.00	0.72
0.6	20	0.63	0.40	0.96	0.97	0.87
	30	0.63	0.37	0.99	0.98	0.90
	40	0.67	0.43	1.00	0.99	0.94
	50	0.68	0.44	1.00	1.00	0.93
0.7	20	0.60	0.39	0.88	0.92	0.93
	30	0.59	0.36	0.96	0.95	0.96
	40	0.60	0.37	0.97	0.98	0.98
	50	0.64	0.39	0.99	0.99	0.98
0.8	20	0.54	0.41	0.77	0.82	0.97
	30	0.53	0.37	0.85	0.89	0.98
	40	0.57	0.39	0.91	0.94	0.99
	50	0.56	0.38	0.93	0.96	1.00
1.0	20	0.50	0.47	0.50	0.69	0.94
	30	0.46	0.45	0.56	0.75	0.97
	40	0.50	0.45	0.57	0.77	0.99
	50	0.51	0.42	0.61	0.81	0.99
1.5	20	0.50	0.64	0.08	0.45	0.88
	30	0.56	0.64	0.06	0.51	0.94
	40	0.58	0.64	0.04	0.51	0.97
	50	0.58	0.62	0.03	0.52	0.98
2.0	20	0.57	0.69	0.01	0.45	0.85
	30	0.60	0.72	0.01	0.44	0.91
	40	0.62	0.67	0.00	0.48	0.95
	50	0.65	0.73	0.00	0.47	0.97
5.0	20	0.69	0.79	0.00	0.55	0.84
	30	0.70	0.79	0.00	0.54	0.87
	40	0.78	0.83	0.00	0.63	0.92
	50	0.83	0.85	0.00	0.63	0.96
10.0	20	0.73	0.84	0.00	0.62	0.82
	30	0.80	0.85	0.00	0.70	0.88
	40	0.84	0.88	0.00	0.74	0.93
	50	0.87	0.90	0.00	0.76	0.95
20.0	20	0.75	0.84	0.00	0.69	0.82
	30	0.83	0.87	0.00	0.75	0.90
	40	0.89	0.90	0.00	0.82	0.93
	50	0.91	0.92	0.00	0.85	0.95

Table 7: The simulated PCSs when the null distribution is WE.

		Disc. Proc.				
β	n	RML	K-S	FI	AV	VM
0.5	20	0.72	0.89	0.00	0.00	0.99
	30	0.78	0.91	0.00	0.00	0.99
	40	0.87	0.94	0.00	0.01	0.99
	50	0.89	0.95	0.00	0.01	1.00
0.8	20	0.66	0.84	0.00	0.00	0.94
	30	0.71	0.88	0.00	0.00	0.98
	40	0.75	0.90	0.00	0.00	0.99
	50	0.81	0.92	0.00	0.00	1.00
1.0	20	0.59	0.79	0.00	0.00	0.75
	30	0.69	0.85	0.00	0.00	0.87
	40	0.73	0.86	0.00	0.00	0.93
	50	0.76	0.87	0.00	0.00	0.96
1.5	20	0.50	0.71	0.07	0.04	0.12
	30	0.54	0.74	0.02	0.01	0.11
	40	0.57	0.74	0.01	0.01	0.08
	50	0.61	0.77	0.00	0.00	0.08
2.0	20	0.50	0.50	0.48	0.31	0.06
	30	0.50	0.52	0.46	0.26	0.03
	40	0.49	0.56	0.41	0.23	0.01
	50	0.54	0.59	0.39	0.21	0.01
2.5	20	0.56	0.39	0.90	0.60	0.17
	30	0.60	0.41	0.91	0.60	0.11
	40	0.63	0.44	0.93	0.64	0.08
	50	0.62	0.47	0.96	0.67	0.08
3.0	20	0.65	0.37	0.99	0.76	0.27
	30	0.68	0.43	1.00	0.80	0.25
	40	0.65	0.42	1.00	0.81	0.17
	50	0.69	0.50	1.00	0.85	0.17
3.5	20	0.68	0.38	1.00	0.81	0.38
	30	0.70	0.44	1.00	0.84	0.35
	40	0.74	0.50	1.00	0.89	0.34
	50	0.77	0.55	1.00	0.90	0.33
4.0	20	0.72	0.41	1.00	0.83	0.48
	30	0.76	0.51	1.00	0.89	0.47
	40	0.78	0.55	1.00	0.90	0.47
	50	0.81	0.59	1.00	0.92	0.46
5.0	20	0.74	0.47	1.00	0.84	0.62
	30	0.78	0.54	1.00	0.88	0.60
	40	0.85	0.64	1.00	0.94	0.66
	50	0.88	0.69	1.00	0.95	0.71
6.0	20	0.78	0.49	1.00	0.85	0.67
	30	0.84	0.61	1.00	0.91	0.73
	40	0.89	0.69	1.00	0.95	0.77
	50	0.90	0.74	1.00	0.97	0.80
8.0	20	0.80	0.54	0.96	0.84	0.75
	30	0.87	0.65	0.99	0.91	0.81
	40	0.91	0.74	1.00	0.95	0.87
	50	0.94	0.81	1.00	0.97	0.90
10.0	20	0.85	0.61	1.00	0.87	0.82
	30	0.89	0.69	0.99	0.93	0.86
	40	0.93	0.78	1.00	0.95	0.91
	50	0.96	0.86	1.00	0.97	0.94
15.0	20	0.88	0.66	1.00	0.89	0.87
	30	0.93	0.77	1.00	0.94	0.92
	40	0.96	0.85	1.00	0.97	0.96
	50	0.98	0.88	1.00	0.99	0.98