# Goodness-of-Fit Test Based on Kullback-Leibler Information for Progressively Type-II Censored Data

Arezou Habibi Rad, Fatemeh Yousefzadeh, and Narayanaswamy Balakrishnan

***Abstract*—We express the joint entropy of progressively Type-II censored order statistics in terms of an incomplete integral of the hazard function, and use it to develop a simple estimate of the joint entropy of progressively Type-II censored data, considered earlier by Balakrishnan *et al.*, *IEEE Trans. Reliability*, vol. 56, pp. 349–356. We then construct a goodness-of-fit test statistic based on the Kullback-Leibler information for Pareto, log-normal, and Weibull distributions by using maximum likelihood estimates and approximate maximum likelihood estimates of the model parameters. Finally, we use Monte Carlo simulations to evaluate the power of the proposed test for several alternatives under different sample sizes and progressive censoring schemes.**

***Index Terms*—Approximate maximum likelihood estimate, entropy, goodness-of-fit test, hazard function, Kullback-Leibler information, log-normal distribution, maximum likelihood estimate, Monte Carlo simulation, Pareto distribution, progressively type-II censored data, Weibull distribution.**

## ACRONYMS

| | |
|---|---|
| AMLE | approximate maximum likelihood estimate |
| c.d.f. | cumulative distribution function |
| KL | Kullback–Leibler |
| MLE | maximum likelihood estimate |
| p.d.f. | probability density function |

## NOTATION

| | |
|---|---|
| $n$ | sample size |
| $X_{i:n}$ | $i$-th order statistic from a sample of size $n$ |
| $m(m \leq n)$ | the number of complete failures in a progressively censored sample |
| $R_i$ | the number of surviving units censored at random at the time of the $i$-th failure |
| $X_{i:m:n}^{(R_1,\ldots,R_m)}$ | $i$-th progressively censored order statistic; for convenience, we will use the simplified notation $X_{i:m:n}$, $i = 1, 2, \ldots, m$ |
| $f_{X_{r:m:n}}$ | the probability density function (p.d.f.) of $X_{r:m:n}$ |
| $f_{X_{1:m:n},\ldots,X_{m:m:n}}$ | the joint probability density function of all $m$ progressively Type-II censored order statistics |
| $h(x)$ | the hazard function (rate), $f(x)/(1-F(x))$ |
| $H_{1\cdots m:m:n}$ | the joint entropy of $X_{1:m:n}, \ldots, X_{m:m:n}$ |
| $I_{1\cdots m:m:n}(f:g)$ | the Kullback-Leibler information of $X_{1:m:n}, \ldots, X_{m:m:n}$ |
| $f_{X_{r:m:n}\mid X_{r-1:m:n}}(x_r\mid x_{r-1})$ | the conditional p.d.f. of $X_{r:m:n}$, given $X_{r-1:m:n} = x_{r-1:m:n}$ |
| $H_{r\mid r-1:m:n}$ | the expectation of conditional entropy of $X_{r:m:n}$, given $X_{r-1:m:n} = x_{r-1:m:n}$. |

Manuscript received April 19, 2010; revised December 12, 2010; accepted March 23, 2011. Date of publication August 01, 2011; date of current version August 31, 2011. The research of the first two authors was supported by a Grant (No. MS87173HAB) from Ferdowsi University of Mashhad, Iran. The last author acknowledges the support of a Discovery Research Grant from the Natural Sciences and Engineering Research Council of Canada for this research through an Individual Discovery grant, and the support of Research Project Number KSU-VPP-107 from King Saud University, Riyadh, Saudi Arabia. Associate Editor: L. Walls.

A. Habibi Rad is with the Department of Statistics, School of Mathematical Sciences, Ferdowsi University of Mashhad, Iran (e-mail: ahabibi@ferdowsi.um.ac.ir).

F. Yousefzadeh is with the Department of Statistics, School of Sciences, Birjand University, Iran.

N. Balakrishnan is with the Department of Mathematics and Statistics, McMaster University, Hamilton, ON L8S 4K1 Canada , and also with King Saud University, Faculty of Science, Riyadh, Saudi Arabia (e-mail: bala@mcmaster.ca).

Digital Object Identifier 10.1109/TR.2011.2162470

## I. INTRODUCTION

DATA arising from life-testing and reliability experiments are often censored based on cost and time considerations. Conventional Type-I and Type-II censoring are the most common forms of censoring. Because parametric inferential procedures are developed with the assumption of a special lifetime distribution, such as exponential, Weibull, log-normal, gamma, and Pareto, it becomes important to test the validity of the assumed lifetime model based on the observed censored data. Many such goodness-of-fit procedures have been developed in the literature for the cases when the available samples are either Type-I or Type-II censored. Recently, a more flexible and efficient form of censoring, called *progressive censoring*, has been studied quite extensively; see, for example, Balakrishnan [2] and Balakrishnan & Aggarwala [3] for elaborate reviews on various developments regarding progressive

TABLE I
POWERS FOR DIFFERENT HAZARD ALTERNATIVES AT 5% LEVEL OF SIGNIFICANCE FOR SEVERAL PROGRESSIVE CENSORING SCHEMES WHEN THE SAMPLE SIZES ARE $n = 10, 20, 30$ WHILE TESTING FOR THE PARETO DISTRIBUTION

| | | | monotone increasing hazard alternatives | | monotone decreasing hazard alternatives | | nonmonotone hazard alternatives | |
|---|---|---|---|---|---|---|---|---|
| $n$ | $m$ | schemes $(R_1, \cdots, R_m)$ | Gamma shape 2 | Gexp shape 2 | Gamma shape 0.5 | Gexp shape 0.5 | Beta shape 0.5 | Log-logistic shape 0.5 |
| 10 | 5 | 5,0,0,0,0 | 0.3458 | 0.3286 | 0.1520 | 0.1574 | 0.2981 | 0.2457 |
| | 5 | 0,5,0,0,0 | 0.4164 | 0.4011 | 0.1844 | 0.1894 | 0.3093 | 0.3302 |
| | 5 | 1,1,1,1,1 | 0.2875 | 0.2790 | 0.1327 | 0.1343 | 0.1782 | 0.2527 |
| | 5 | 0,0,0,5,0 | 0.2120 | 0.2059 | 0.0969 | 0.0974 | 0.1277 | 0.1888 |
| | 5 | 0,0,0,0,5 | 0.2213 | 0.2162 | 0.1176 | 0.1178 | 0.1380 | 0.2075 |
| | 7 | 3,0,0,0,0,0,0 | 0.5790 | 0.5565 | 0.2986 | 0.3117 | 0.5524 | 0.3957 |
| | 7 | 0,3,0,0,0,0,0 | 0.5985 | 0.5779 | 0.3127 | 0.3228 | 0.5465 | 0.4363 |
| | 7 | 1,0,0,1,0,0,1 | 0.4843 | 0.4701 | 0.2509 | 0.2566 | 0.3920 | 0.3925 |
| | 7 | 0,0,0,0,0,3,0 | 0.3872 | 0.3725 | 0.1840 | 0.1886 | 0.3127 | 0.3081 |
| | 7 | 0,0,0,0,0,0,3 | 0.4464 | 0.4382 | 0.2467 | 0.2498 | 0.3284 | 0.3955 |
| 20 | 5 | 15,0,0,0,0 | 0.4317 | 0.4101 | 0.1444 | 0.1481 | 0.2818 | 0.3175 |
| | 5 | 0,15,0,0,0 | 0.5134 | 0.5004 | 0.1862 | 0.1907 | 0.2858 | 0.4483 |
| | 5 | 3,3,3,3,3 | 0.2925 | 0.2873 | 0.1162 | 0.1168 | 0.1265 | 0.2875 |
| | 5 | 0,0,0,15,0 | 0.1613 | 0.1574 | 0.0613 | 0.0615 | 0.0657 | 0.1591 |
| | 5 | 0,0,0,0,15 | 0.2015 | 0.1993 | 0.0934 | 0.0936 | 0.0984 | 0.2028 |
| | 10 | 10,0,0,...,0,0,0 | 0.8362 | 0.8178 | 0.4889 | 0.5040 | 0.7884 | 0.6069 |
| | 10 | 0,10,0,...,0,0,0 | 0.8546 | 0.8365 | 0.5219 | 0.5345 | 0.7834 | 0.6759 |
| | 10 | 1,1,1,...,1,1,1 | 0.7172 | 0.6990 | 0.3436 | 0.3486 | 0.4910 | 0.6257 |
| | 10 | 0,0,0,...,0,10,0 | 0.4006 | 0.3880 | 0.1768 | 0.1791 | 0.2477 | 0.3499 |
| | 10 | 0,0,0,...,0,0,10 | 0.6322 | 0.6239 | 0.3515 | 0.3531 | 0.4002 | 0.6130 |
| 30 | 15 | 15,0,0,...,0,0,0 | 0.9692 | 0.9585 | 0.6704 | 0.6898 | 0.9436 | 0.7520 |
| | 15 | 0,15,0,...,0,0,0 | 0.9754 | 0.9674 | 0.6933 | 0.7122 | 0.9422 | 0.8004 |
| | 15 | 1,1,1,...,1,1,1 | 0.8854 | 0.8719 | 0.4286 | 0.4362 | 0.6420 | 0.7917 |
| | 15 | 0,0,0,...,0,15,0 | 0.5227 | 0.5068 | 0.2059 | 0.2098 | 0.3010 | 0.4508 |
| | 15 | 0,0,0,...,0,0,15 | 0.8347 | 0.8260 | 0.5077 | 0.5110 | 0.5667 | 0.8198 |
| | 20 | 10,0,0,...,0,0,0 | 0.9931 | 0.9898 | 0.8115 | 0.8293 | 0.9912 | 0.8025 |
| | 20 | 0,10,0,...,0,0,0 | 0.9925 | 0.9879 | 0.8218 | 0.8397 | 0.9889 | 0.8347 |
| | 20 | 1,0,1,...,0,1,0 | 0.9660 | 0.9534 | 0.6435 | 0.6629 | 0.9356 | 0.7903 |
| | 20 | 0,0,0,...,0,10,0 | 0.7588 | 0.7360 | 0.4078 | 0.4171 | 0.6538 | 0.6169 |
| | 20 | 0,0,0,...,0,0,10 | 0.9425 | 0.9369 | 0.6404 | 0.6471 | 0.7588 | 0.9156 |

censoring. However, not many goodness-of-fit tests have been developed for progressively censored data. We therefore develop here an information-based goodness-of-fit test for Pareto, log-normal, and Weibull distributions based on progressively Type-II censored data.

Suppose a random variable $X$ has distribution function $F(x)$, and continuous density function $f(x)$. Then, the differential entropy $H(f)$ of the random variable $X$ is defined by Shannon [18] to be

$$H(f) = -\int_{-\infty}^{\infty} f(x) \log f(x) dx. \quad (1)$$

Vasicek [19] was the first to propose a test for normality based on sample entropy, who also compared its power with those of leading test statistics for complete samples. The entropy difference $H(f) - H(g)$ was considered by Dudewicz & van der Meulen [10] and Gokhale [12] for establishing goodness-of-fit tests for the class of maximum entropy distributions.

The Kullback-Leibler (KL) information in favor of $g(x)$ against $f(x)$ was defined by Kullback [14] to be

$$I(g:f) = \int_{-\infty}^{\infty} g(x) \log \frac{g(x)}{f(x)} dx,$$

which is in fact an extended concept of entropy. Because $I(g:f)$ is such that $I(g:f) \geq 0$ with the equality holding if $g = f$, the sample estimate of the KL information has also been considered as a goodness-of-fit test statistic; see, for example, Arizono & Ohta [1], and Ebrahimi & Habibullah [11] for complete samples; and Park [17], and Balakrishnan *et al.* [5] for Type-II censored data, and progressively Type-II censored data, respectively.

In this paper, we extend the goodness-of-fit test based on KL information for progressively Type-II censored data for Pareto, log-normal, and Weibull distributions. The rest of this paper is organized as follows. In Section II, we describe briefly the progressive Type-II censoring, the joint entropy of these progressively censored order statistics in terms of the hazard function, and the nonparametric estimate of this joint entropy. In Section III, we define the KL information for progressively Type-II censored data. Then we use that result to propose a goodness-of-fit test for Pareto, log-normal, and Weibull distributions in Sections IV Sections V Sections VI, respectively. Finally, in Section VII, we use Monte Carlo simulations to evaluate the power of these tests for several alternatives under different sample sizes and progressive Type-II censoring schemes. Finally, in Section VIII, we present an example to illustrate the goodness-of-fit test developed here.

TABLE II
POWERS FOR DIFFERENT HAZARD ALTERNATIVES AT 10% LEVEL OF SIGNIFICANCE FOR SEVERAL PROGRESSIVE CENSORING SCHEMES WHEN THE SAMPLE SIZES ARE $n = 10, 20, 30$ WHILE TESTING FOR THE PARETO DISTRIBUTION

| | | | monotone increasing hazard alternatives | | monotone decreasing hazard alternatives | | nonmonotone hazard alternatives | |
|---|---|---|---|---|---|---|---|---|
| $n$ | $m$ | schemes $(R_1, \cdots, R_m)$ | Gamma shape 2 | Gexp shape 2 | Gamma shape 0.5 | Gexp shape 0.5 | Beta shape 0.5 | Log-logistic shape 0.5 |
| 10 | 5 | 5,0,0,0,0 | 0.5198 | 0.5034 | 0.2416 | 0.2480 | 0.4167 | 0.3957 |
| | 5 | 0,5,0,0,0 | 0.5475 | 0.5339 | 0.2700 | 0.2776 | 0.4135 | 0.4525 |
| | 5 | 1,1,1,1,1 | 0.4338 | 0.4219 | 0.2106 | 0.2133 | 0.2754 | 0.3907 |
| | 5 | 0,0,0,5,0 | 0.3147 | 0.3058 | 0.1421 | 0.1435 | 0.1868 | 0.2828 |
| | 5 | 0,0,0,0,5 | 0.3239 | 0.3191 | 0.1665 | 0.1680 | 0.1909 | 0.3119 |
| | 7 | 3,0,0,0,0,0,0 | 0.6918 | 0.6699 | 0.3884 | 0.3987 | 0.6464 | 0.5118 |
| | 7 | 0,3,0,0,0,0,0 | 0.7110 | 0.6956 | 0.4005 | 0.4102 | 0.6375 | 0.5492 |
| | 7 | 1,0,0,1,0,0,1 | 0.6273 | 0.6105 | 0.3320 | 0.3388 | 0.4797 | 0.5261 |
| | 7 | 0,0,0,0,0,3,0 | 0.4943 | 0.4771 | 0.2355 | 0.2409 | 0.3824 | 0.3972 |
| | 7 | 0,0,0,0,0,0,3 | 0.5514 | 0.5421 | 0.3078 | 0.3120 | 0.3971 | 0.5028 |
| 20 | 5 | 15,0,0,0,0 | 0.5773 | 0.5582 | 0.2075 | 0.2136 | 0.3687 | 0.4529 |
| | 5 | 0,15,0,0,0 | 0.6441 | 0.6335 | 0.2699 | 0.2741 | 0.3981 | 0.5785 |
| | 5 | 3,3,3,3,3 | 0.4233 | 0.4178 | 0.1728 | 0.1736 | 0.1877 | 0.4200 |
| | 5 | 0,0,0,15,0 | 0.2855 | 0.2782 | 0.1043 | 0.1050 | 0.1135 | 0.2742 |
| | 5 | 0,0,0,0,15 | 0.2874 | 0.2850 | 0.1411 | 0.1411 | 0.1457 | 0.2905 |
| | 10 | 10,0,0,...,0,0,0 | 0.9156 | 0.8995 | 0.5795 | 0.5951 | 0.8483 | 0.7184 |
| | 10 | 0,10,0,...,0,0,0 | 0.9161 | 0.9033 | 0.5907 | 0.6049 | 0.8423 | 0.7544 |
| | 10 | 1,1,1,....,1,1,1 | 0.7995 | 0.7858 | 0.4057 | 0.4130 | 0.5628 | 0.7188 |
| | 10 | 0,0,0,...,0,10,0 | 0.4904 | 0.4756 | 0.2202 | 0.2230 | 0.3072 | 0.4248 |
| | 10 | 0,0,0,...,0,0,10 | 0.7080 | 0.7003 | 0.3935 | 0.3959 | 0.4454 | 0.6895 |
| 30 | 15 | 15,0,0,...,0,0,0 | 0.9884 | 0.9828 | 0.7532 | 0.7695 | 0.9640 | 0.8239 |
| | 15 | 0,15,0,...,0,0,0 | 0.9881 | 0.9825 | 0.7672 | 0.7826 | 0.9656 | 0.8650 |
| | 15 | 1,1,1,....,1,1,1 | 0.9407 | 0.9295 | 0.5125 | 0.5209 | 0.7249 | 0.8647 |
| | 15 | 0,0,0,...,0,15,0 | 0.6097 | 0.5914 | 0.2614 | 0.2652 | 0.3758 | 0.5302 |
| | 15 | 0,0,0,...,0,0,15 | 0.8710 | 0.8642 | 0.5176 | 0.5204 | 0.5766 | 0.8604 |
| | 20 | 10,0,0,...,0,0,0 | 0.9974 | 0.9949 | 0.8680 | 0.8841 | 0.9941 | 0.8623 |
| | 20 | 0,10,0,...,0,0,0 | 0.9976 | 0.9956 | 0.8764 | 0.8903 | 0.9949 | 0.8868 |
| | 20 | 1,0,1,....,0,1,0 | 0.9832 | 0.9755 | 0.7062 | 0.7224 | 0.9576 | 0.8402 |
| | 20 | 0,0,0,...,0,10,0 | 0.8235 | 0.8013 | 0.4638 | 0.4737 | 0.7211 | 0.6779 |
| | 20 | 0,0,0,...,0,0,10 | 0.9640 | 0.9609 | 0.6918 | 0.6965 | 0.8042 | 0.9475 |

## II. PRELIMINARIES

### A. Progressively Type-II Censored Data

Suppose $n$ identical units are placed on a life-testing experiment. Assume that their life-times are $s$-independent and identically distributed with cumulative distribution function (c.d.f.) $F(x;\underline{\theta})$, and probability density function (p.d.f.) $f(x;\underline{\theta})$, where $\underline{\theta}$ is a vector of parameters.

Quite often, in life-testing and reliability experiments, some units that are subject to test are removed from the experiment before their failure either intentionally or unintentionally. Such units result in censored lifetimes. The two most common censoring schemes are the conventional Type-I and Type-II censoring schemes, which have been studied extensively in the statistical and reliability literature; see Balakrishnan & Cohen [4], and Cohen [9]. Briefly, they can be described as follows. Consider $n$ items under observation in a life-testing experiment. In the conventional Type-I censoring scheme, the experiment continues up to a pre-specified time T. The conventional Type-II censoring scheme requires the experiment to continue until a pre-specified number of failures $m(\leq n)$ occur. One of the drawbacks of the conventional Type-I and Type-II censoring schemes is that they do not allow for removal of units at points other than the termination point of the experiment. The progressive Type-II censoring scheme, which does facilitate the removal of units at many stages of the experiment, can be described as follows. Consider $n$ units in a study, and suppose $m(\leq n)$ is fixed prior to the experiment. Moreover, $m$ non-negative integers, $R_1, \ldots R_m$, are also fixed prior to the experiment such that $R_1 + \cdots + R_m + m = n$. At the time of the first failure, say $X_{1:m:n}$, $R_1$ of the surviving units are randomly removed. Next, at the time of the second failure, say $X_{2:m:n}$, $R_2$ of the surviving units are randomly removed, and so on. Finally, at the time of the $m$-th failure, say $X_{m:m:n}$, all remaining $R_m$ surviving units are removed. For further details on progressive Type-II censoring, one may refer to the monograph by Balakrishnan & Aggarwala [3], and the elaborate review article by Balakrishnan [2].

The joint p.d.f. of all $m$ progressively Type-II censored order statistics $(X_{1:m:n}, \ldots, X_{m:m:n})$ is given by

$$f_{X_{1:m:n},X_{2:m:n},\ldots,X_{m:m:n}}(x_1, x_2, \ldots, x_m) = c\prod_{i=1}^{m} f(x_i)\left\{1 - F(x_i)\right\}^{R_i}, \quad x_1 < x_2 < \cdots < x_m,$$

where

$c = n(n - R_1 - 1)\cdots(n - R_1 - R_2 - \cdots - R_{m-1} - m + 1).$

TABLE III
POWERS FOR DIFFERENT HAZARD ALTERNATIVES AT 5% LEVEL OF SIGNIFICANCE FOR SEVERAL PROGRESSIVE CENSORING SCHEMES WHEN THE SAMPLE SIZES ARE $n = 10, 20, 30$ WHILE TESTING FOR THE LOG-NORMAL DISTRIBUTION

| | | | monotone increasing hazard alternatives | | monotone decreasing hazard alternatives | | nonmonotone hazard alternatives | |
|---|---|---|---|---|---|---|---|---|
| $n$ | $m$ | schemes $(R_1, \cdots, R_m)$ | Gamma shape 2 | Gexp shape 2 | Gamma shape 0.5 | Gexp shape 0.5 | Beta shape 0.5 | Log-logistic shape 0.5 |
| 10 | 5 | 5,0,0,0,0 | 0.1202 | 0.1355 | 0.4444 | 0.4340 | 0.2465 | 0.2519 |
| | 5 | 0,5,0,0,0 | 0.1147 | 0.1262 | 0.4447 | 0.4359 | 0.2579 | 0.2277 |
| | 5 | 1,1,1,1,1 | 0.0933 | 0.1010 | 0.4060 | 0.4016 | 0.3086 | 0.1263 |
| | 5 | 0,0,0,5,0 | 0.0664 | 0.0691 | 0.3547 | 0.3483 | 0.2300 | 0.1095 |
| | 5 | 0,0,0,0,5 | 0.0945 | 0.0992 | 0.3499 | 0.3487 | 0.3073 | 0.0995 |
| | 7 | 3,0,0,0,0,0,0 | 0.1620 | 0.1861 | 0.5988 | 0.5848 | 0.2925 | 0.3799 |
| | 7 | 0,3,0,0,0,0,0 | 0.1812 | 0.2085 | 0.6187 | 0.6043 | 0.3171 | 0.3910 |
| | 7 | 1,0,0,1,0,0,1 | 0.1674 | 0.1846 | 0.6025 | 0.5939 | 0.3928 | 0.2832 |
| | 7 | 0,0,0,0,0,3,0 | 0.1006 | 0.1107 | 0.4704 | 0.4605 | 0.3040 | 0.1821 |
| | 7 | 0,0,0,0,0,0,3 | 0.1646 | 0.1763 | 0.5566 | 0.5526 | 0.4312 | 0.2161 |
| 20 | 5 | 15,0,0,0,0 | 0.1153 | 0.1301 | 0.4972 | 0.4876 | 0.2928 | 0.2472 |
| | 5 | 0,15,0,0,0 | 0.1422 | 0.1580 | 0.5531 | 0.5470 | 0.3904 | 0.2483 |
| | 5 | 3,3,3,3,3 | 0.0897 | 0.0937 | 0.4670 | 0.4655 | 0.4323 | 0.0918 |
| | 5 | 0,0,0,15,0 | 0.0591 | 0.0633 | 0.4665 | 0.4625 | 0.3669 | 0.0894 |
| | 5 | 0,0,0,0,15 | 0.0910 | 0.0925 | 0.3722 | 0.3718 | 0.3621 | 0.0852 |
| | 10 | 10,0,0,....,0,0,0 | 0.2500 | 0.2947 | 0.8526 | 0.8411 | 0.5385 | 0.5476 |
| | 10 | 0,10,0,....,0,0,0 | 0.2964 | 0.3408 | 0.8670 | 0.8595 | 0.5814 | 0.5720 |
| | 10 | 1,1,1,....,1,1,1 | 0.2286 | 0.2541 | 0.8675 | 0.8630 | 0.7518 | 0.3271 |
| | 10 | 0,0,0,....,0,10,0 | 0.0945 | 0.0988 | 0.5670 | 0.5619 | 0.4145 | 0.1343 |
| | 10 | 0,0,0,....,0,0,10 | 0.2232 | 0.2387 | 0.7999 | 0.7980 | 0.7448 | 0.2369 |
| 30 | 15 | 15,0,0,....,0,0,0 | 0.3143 | 0.3724 | 0.9568 | 0.9522 | 0.7532 | 0.6768 |
| | 15 | 0,15,0,....,0,0,0 | 0.3682 | 0.4296 | 0.9644 | 0.9605 | 0.7869 | 0.7153 |
| | 15 | 1,1,1,....,1,1,1 | 0.3173 | 0.3527 | 0.9654 | 0.9635 | 0.9058 | 0.4371 |
| | 15 | 0,0,0,....,0,15,0 | 0.1202 | 0.1234 | 0.5334 | 0.5251 | 0.4019 | 0.1370 |
| | 15 | 0,0,0,....,0,0,15 | 0.3573 | 0.3792 | 0.9635 | 0.9624 | 0.9395 | 0.3751 |
| | 20 | 10,0,0,....,0,0,0 | 0.3961 | 0.4669 | 0.9859 | 0.9842 | 0.8930 | 0.7883 |
| | 20 | 0,10,0,....,0,0,0 | 0.4136 | 0.4759 | 0.9831 | 0.9806 | 0.8924 | 0.7920 |
| | 20 | 1,0,1,....,0,1,0 | 0.2702 | 0.3063 | 0.9573 | 0.9546 | 0.8616 | 0.5103 |
| | 20 | 0,0,0,....,0,10,0 | 0.1936 | 0.2013 | 0.7446 | 0.7377 | 0.6356 | 0.2297 |
| | 20 | 0,0,0,....,0,0,10 | 0.5284 | 0.5714 | 0.9953 | 0.9948 | 0.9816 | 0.6222 |

### *B. Entropy of Progressively Censored Data in Terms of the Hazard Function*

The joint entropy of $X_{1:m:n}, \ldots, X_{m:m:n}$ is given by (Park [17])

$$\begin{aligned} H_{1\cdots m:m:n} &= -\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{x_{2:m:n}} f_{X_{1:m:n},X_{2:m:n},\ldots,X_{m:m:n}} \\ &\quad \times (x_1, x_2, \ldots, x_m) \times \log f_{X_{1:m:n},X_{2:m:n},\ldots,X_{m:m:n}} \\ &\quad \times (x_1, x_2, \ldots, x_m) dx_1 \cdots dx_m, \end{aligned}$$

where $f_{X_{1:m:n},X_{2:m:n},\ldots,X_{m:m:n}}(x_1, x_2, \ldots, x_m)$ is the joint density of all $m$ progressively Type-II censored order statistics. Because $H_{1\cdots m:m:n}$ is an $m$-dimensional integral, we need to obtain a simpler expression for this multiple integral.

The calculation of the entropy of the single and consecutive order statistics has been discussed by Wong & Chan [21], and Park [16]. The multiple integral of the entropy for Type-II censored data has been simplified to a single-integral by Park [17], and the joint entropy of progressively Type-II censored order statistics has been simplified by Balakrishnan *et al.* [5] in terms of an integral involving the hazard function , $h(x)$, as

$$H_{1\cdots m:m:n} = -\log c + n\overline{H}_{1\cdots m:m:n},$$

where

$$\overline{H}_{1\cdots m:m:n} = \frac{m}{n} - \frac{1}{n}\int_{-\infty}^{\infty} \sum_{i=1}^{m} f_{X_{i:m:n}}(x) \log h(x) dx.$$

### *C. Nonparametric Entropy Estimate*

The nonparametric estimate of the joint entropy $H_{1\cdots m:m:n}$ was obtained by Balakrishnan *et al.* [5] as

$$H_{1\cdots m:m:n}(w, n, m) = -\log c + nH(w, n, m),$$

where

$$\begin{aligned} H(w, n, m) = \frac{1}{n}\sum_{i=1}^{m} \log\left(\frac{x_{i+w:m:n} - x_{i-w:m:n}}{E(U_{i+w:m:n}) - E(U_{i-w:m:n})}\right) \\ - \left(1 - \frac{m}{n}\right)\log\left(1 - \frac{m}{n}\right). \end{aligned}$$

TABLE IV
POWERS FOR DIFFERENT HAZARD ALTERNATIVES AT 10% LEVEL OF SIGNIFICANCE FOR SEVERAL PROGRESSIVE CENSORING SCHEMES WHEN THE SAMPLE SIZES ARE $n = 10, 20, 30$ WHILE TESTING FOR THE LOG-NORMAL DISTRIBUTION

| | | | monotone increasing hazard alternatives | | monotone decreasing hazard alternatives | | nonmonotone hazard alternatives | |
|---|---|---|---|---|---|---|---|---|
| $n$ | $m$ | schemes $(R_1, \cdots, R_m)$ | Gamma shape 2 | Gexp shape 2 | Gamma shape 0.5 | Gexp shape 0.5 | Beta shape 0.5 | Log-logistic shape 0.5 |
| 10 | 5 | 5,0,0,0,0 | 0.1947 | 0.2127 | 0.5421 | 0.5341 | 0.3391 | 0.3370 |
| | 5 | 0,5,0,0,0 | 0.2006 | 0.2196 | 0.5607 | 0.5526 | 0.3685 | 0.3314 |
| | 5 | 1,1,1,1,1 | 0.1745 | 0.1829 | 0.5198 | 0.5155 | 0.4241 | 0.2172 |
| | 5 | 0,0,0,5,0 | 0.1184 | 0.1236 | 0.4323 | 0.4268 | 0.3043 | 0.1682 |
| | 5 | 0,0,0,0,5 | 0.1750 | 0.1809 | 0.4758 | 0.4738 | 0.4276 | 0.1825 |
| | 7 | 3,0,0,0,0,0,0 | 0.2496 | 0.2820 | 0.6965 | 0.6835 | 0.4077 | 0.4757 |
| | 7 | 0,3,0,0,0,0,0 | 0.2751 | 0.3065 | 0.7111 | 0.7007 | 0.4165 | 0.4920 |
| | 7 | 1,0,0,1,0,0,1 | 0.2568 | 0.2799 | 0.7028 | 0.6944 | 0.5063 | 0.3769 |
| | 7 | 0,0,0,0,0,3,0 | 0.1620 | 0.1706 | 0.5678 | 0.5609 | 0.4059 | 0.2506 |
| | 7 | 0,0,0,0,0,0,3 | 0.2667 | 0.2828 | 0.6791 | 0.6747 | 0.5607 | 0.3264 |
| 20 | 5 | 15,0,0,0,0 | 0.1936 | 0.2133 | 0.5912 | 0.5844 | 0.3912 | 0.3330 |
| | 5 | 0,15,0,0,0 | 0.2218 | 0.2390 | 0.6252 | 0.6181 | 0.4709 | 0.3316 |
| | 5 | 3,3,3,3,3 | 0.1565 | 0.1613 | 0.5562 | 0.5545 | 0.5229 | 0.1587 |
| | 5 | 0,0,0,15,0 | 0.1242 | 0.1318 | 0.5772 | 0.5740 | 0.4917 | 0.1644 |
| | 5 | 0,0,0,0,15 | 0.1659 | 0.1689 | 0.4876 | 0.4873 | 0.4750 | 0.1583 |
| | 10 | 10,0,0,...,0,0,0 | 0.3595 | 0.4046 | 0.9093 | 0.9033 | 0.6752 | 0.6434 |
| | 10 | 0,10,0,...,0,0,0 | 0.4237 | 0.4699 | 0.9157 | 0.9095 | 0.7135 | 0.6785 |
| | 10 | 1,1,1,...,1,1,1 | 0.3408 | 0.3712 | 0.9227 | 0.9196 | 0.8444 | 0.4398 |
| | 10 | 0,0,0,...,0,10,0 | 0.1529 | 0.1578 | 0.6430 | 0.6350 | 0.5047 | 0.1976 |
| | 10 | 0,0,0,...,0,0,10 | 0.3475 | 0.3646 | 0.8833 | 0.8817 | 0.8413 | 0.3640 |
| 30 | 15 | 15,0,0,...,0,0,0 | 0.4245 | 0.4873 | 0.9750 | 0.9723 | 0.8539 | 0.7531 |
| | 15 | 0,15,0,...,0,0,0 | 0.4949 | 0.5493 | 0.9804 | 0.9783 | 0.8853 | 0.7873 |
| | 15 | 1,1,1,...,1,1,1 | 0.4205 | 0.4538 | 0.9807 | 0.9800 | 0.9477 | 0.5225 |
| | 15 | 0,0,0,...,0,15,0 | 0.1862 | 0.1888 | 0.6288 | 0.6222 | 0.5019 | 0.1993 |
| | 15 | 0,0,0,...,0,0,15 | 0.5073 | 0.5325 | 0.9804 | 0.9801 | 0.9660 | 0.5245 |
| | 20 | 10,0,0,...,0,0,0 | 0.5117 | 0.5755 | 0.9922 | 0.9908 | 0.9388 | 0.8364 |
| | 20 | 0,10,0,...,0,0,0 | 0.5450 | 0.6038 | 0.9934 | 0.9930 | 0.9499 | 0.8515 |
| | 20 | 1,0,1,...,0,1,0 | 0.4663 | 0.5166 | 0.9933 | 0.9927 | 0.9556 | 0.6946 |
| | 20 | 0,0,0,...,0,10,0 | 0.2698 | 0.2790 | 0.8082 | 0.8028 | 0.7062 | 0.3036 |
| | 20 | 0,0,0,...,0,0,10 | 0.6870 | 0.7204 | 0.9987 | 0.9984 | 0.9947 | 0.7618 |

## III. GOODNESS-OF-FIT TEST BASED ON THE KULLBACK-LEIBLER INFORMATION

For a null density function $f^0(x;\underline{\theta})$ whose adequacy for the data at hand that we wish to assess, the KL information from a progressively Type-II censored data is given by (See the equation at the bottom of the page) where $f_{X_{1:m:n},X_{2:m:n},\ldots,X_{m:m:n}}(x_1,x_2,\ldots,x_m)$ is the joint density of all $m$ progressively Type-II censored order statistics. By using an estimate of the above KL information based on the given progressively Type-II censored sample, a test statistic based on $(1/n)I_{1\cdots m:m:n}(f:f^0)$ can be constructed as

$$T(w,n,m) = -H(w,n,m) - \frac{1}{n}\left[\sum_{i=1}^{m}\log f^0(x_i;\widehat{\underline{\theta}}) + \sum_{i=1}^{m} R_i \log\left(1 - F^0(x_i;\widehat{\underline{\theta}})\right)\right], \quad (2)$$

where $\widehat{\underline{\theta}}$ is a consistent estimate of $\underline{\theta}$.

The idea and motivation behind the test statistic proposed in (2) is quite simple and intuitive. Based on the KL information measure, we are evaluating the discrepancy between the observed vector of progressively Type-II censored order statistics

$$I_{1\cdots m:m:n}(f:f^0) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{x_{2:m:n}} f_{X_{1:m:n},X_{2:m:n},\ldots,X_{m:m:n}}(x_1,x_2,\ldots,x_m;\underline{\theta}) \times \log\frac{f_{X_{1:m:n},X_{2:m:n},\ldots,X_{m:m:n}}(x_1,x_2,\ldots,x_m;\underline{\theta})}{f^0_{X_{1:m:n},X_{2:m:n},\ldots,X_{m:m:n}}(x_1,x_2,\ldots,x_m;\underline{\theta})}dx_1\cdots dx_m,$$

TABLE V
POWERS FOR DIFFERENT HAZARD ALTERNATIVES AT 5% LEVEL OF SIGNIFICANCE FOR SEVERAL PROGRESSIVE CENSORING SCHEMES WHEN THE SAMPLE SIZES ARE $n = 10, 20, 30$ WHILE TESTING FOR THE WEIBULL DISTRIBUTION

| | | | monotone increasing hazard alternatives | | monotone decreasing hazard alternatives | | nonmonotone hazard alternatives | |
|---|---|---|---|---|---|---|---|---|
| $n$ | $m$ | schemes $(R_1, \cdots, R_m)$ | Gamma shape 2 | Gexp shape 2 | Gamma shape 0.5 | Gexp shape 0.5 | Beta shape 0.5 | Glogistic shape 0.5 |
| 10 | 5 | 5,0,0,0,0 | 0.1813 | 0.2005 | 0.5285 | 0.5186 | 0.3027 | 0.3273 |
| | 5 | 0,5,0,0,0 | 0.2503 | 0.2713 | 0.6148 | 0.6059 | 0.4109 | 0.3841 |
| | 5 | 1,1,1,1,1 | 0.1574 | 0.1670 | 0.5069 | 0.5039 | 0.4071 | 0.2046 |
| | 5 | 0,0,0,5,0 | 0.2039 | 0.2213 | 0.6073 | 0.6034 | 0.4851 | 0.2836 |
| | 5 | 0,0,0,0,5 | 0.1178 | 0.1216 | 0.3976 | 0.3950 | 0.3522 | 0.1249 |
| | 7 | 3,0,0,0,0,0,0 | 0.3389 | 0.3725 | 0.7472 | 0.7376 | 0.4275 | 0.5626 |
| | 7 | 0,3,0,0,0,0,0 | 0.4013 | 0.4337 | 0.7932 | 0.7840 | 0.5083 | 0.5990 |
| | 7 | 1,0,0,1,0,0,1 | 0.3051 | 0.3282 | 0.7330 | 0.7246 | 0.5350 | 0.4354 |
| | 7 | 0,0,0,0,0,3,0 | 0.2792 | 0.2989 | 0.7211 | 0.7141 | 0.5382 | 0.4013 |
| | 7 | 0,0,0,0,0,0,3 | 0.2573 | 0.2758 | 0.6820 | 0.6775 | 0.5576 | 0.3238 |
| 20 | 5 | 15,0,0,0,0 | 0.1545 | 0.1724 | 0.5557 | 0.5469 | 0.3310 | 0.3011 |
| | 5 | 0,15,0,0,0 | 0.3114 | 0.3331 | 0.7090 | 0.7038 | 0.5454 | 0.4321 |
| | 5 | 3,3,3,3,3 | 0.1387 | 0.1444 | 0.5483 | 0.5464 | 0.5150 | 0.1395 |
| | 5 | 0,0,0,15,0 | 0.2966 | 0.3150 | 0.7679 | 0.7666 | 0.7331 | 0.3426 |
| | 5 | 0,0,0,0,15 | 0.1045 | 0.1065 | 0.3900 | 0.3898 | 0.3783 | 0.0970 |
| | 10 | 10,0,0,...,0,0,0 | 0.4370 | 0.4856 | 0.9076 | 0.9022 | 0.6032 | 0.7005 |
| | 10 | 0,10,0,...,0,0,0 | 0.5529 | 0.5994 | 0.9454 | 0.9411 | 0.7287 | 0.7734 |
| | 10 | 1,1,1,...,1,1,1 | 0.4901 | 0.5222 | 0.9511 | 0.9496 | 0.8979 | 0.5874 |
| | 10 | 0,0,0,...,0,10,0 | 0.2769 | 0.2981 | 0.8443 | 0.8417 | 0.7437 | 0.3621 |
| | 10 | 0,0,0,...,0,0,10 | 0.3384 | 0.3546 | 0.8793 | 0.8775 | 0.8353 | 0.3546 |
| 30 | 15 | 15,0,0,...,0,0,0 | 0.7017 | 0.7536 | 0.9961 | 0.9956 | 0.8728 | 0.9053 |
| | 15 | 0,15,0,...,0,0,0 | 0.6837 | 0.7348 | 0.9920 | 0.9904 | 0.8672 | 0.8937 |
| | 15 | 1,1,1,...,1,1,1 | 0.7540 | 0.7851 | 0.9983 | 0.9982 | 0.9934 | 0.8224 |
| | 15 | 0,0,0,...,0,15,0 | 0.2871 | 0.3096 | 0.9085 | 0.9057 | 0.8253 | 0.3753 |
| | 15 | 0,0,0,...,0,0,15 | 0.5949 | 0.6165 | 0.9883 | 0.9880 | 0.9777 | 0.6087 |
| | 20 | 10,0,0,...,0,0,0 | 0.8895 | 0.9182 | 0.9999 | 0.9999 | 0.9823 | 0.9780 |
| | 20 | 0,10,0,...,0,0,0 | 0.8573 | 0.8934 | 0.9999 | 0.9998 | 0.9757 | 0.9698 |
| | 20 | 1,0,1,...,0,1,0 | 0.8439 | 0.8715 | 0.9991 | 0.9990 | 0.9940 | 0.9334 |
| | 20 | 0,0,0,...,0,10,0 | 0.4520 | 0.4852 | 0.9789 | 0.9776 | 0.9205 | 0.5809 |
| | 20 | 0,0,0,...,0,0,10 | 0.8243 | 0.8481 | 0.9993 | 0.9993 | 0.9979 | 0.8694 |

and what we would expect from such a data if they were to arise from the hypothesized model $f^0$.

It may also be noted that the statistic $T(w,n,m)$ in (2) may take on negative values as we have left out the constant term, and also have based the statistic on $(1/n)I_{1\cdots m:m:n}(f : f^0)$ rather than on $I_{1\cdots m:m:n}(f : f^0)$ itself. This approach does not pose any problem because the statistic proposed in (2) is for the purpose of goodness-of-fit, and the critical values get adjusted accordingly. One could also instead use the notion of maximum entropy to develop a goodness-of-fit test.

## IV. GOODNESS-OF-FIT TEST FOR PARETO

Suppose we are interested in testing

$$H_0 : f^0 = \frac{\alpha\beta^\alpha}{x^{\alpha+1}} I(x>\beta) \quad \text{vs.} \quad H_A : f^0 \neq \frac{\alpha\beta^\alpha}{x^{\alpha+1}} I(x>\beta),$$

where $\underline{\theta} = (\alpha, \beta)$ is unknown, and $I(\cdot)$ is the indicator function. Then the KL information for a progressively Type-II censored data can be estimated, and the corresponding test statistic can be computed from (2) for which we may estimate the unknown parameters $(\alpha, \beta)$ by maximum likelihood estimates (MLEs). Of course, as mentioned above, one could use instead the notion of maximum entropy for developing a goodness-of-fit test. In the case of the Pareto distribution, this approach will be quite meaningful because it is known that Pareto distributions have maximal entropy under some conditions. We are currently working on this problem, and hope to report these findings in a future paper.

### A. MLEs for the Pareto Distribution

Here, we provide the MLEs of $\alpha$, and $\beta$ for the Pareto distribution. Let $x_{1:m:n}, \ldots, x_{m:m:n}$ be a progressively Type-II censored sample from the two-parameter Pareto distribution under $H_0$, with censoring scheme $R_1, \ldots, R_m$. Then the likelihood function is

$$L(\alpha,\beta) = k\alpha^m \prod_{i=1}^{m} \frac{\beta^\alpha}{x_{i:m:n}^{\alpha+1}} \left(\frac{\beta}{x_{i:m:n}}\right)^{\alpha R_i}, \quad \beta \leq \min(x_{i:m:n}), \tag{3}$$

where $k$ is a normalizing constant. So the MLE of $\beta$ is

$$\hat{\beta} = x_{1:m:n}. \tag{4}$$

TABLE VI
POWERS FOR DIFFERENT HAZARD ALTERNATIVES AT 10% LEVEL OF SIGNIFICANCE FOR SEVERAL PROGRESSIVE CENSORING SCHEMES WHEN THE SAMPLE SIZES ARE $n = 10, 20, 30$ WHILE TESTING FOR THE WEIBULL DISTRIBUTION

| | | | monotone increasing hazard alternatives | | monotone decreasing hazard alternatives | | nonmonotone hazard alternatives | |
|---|---|---|---|---|---|---|---|---|
| $n$ | $m$ | schemes $(R_1, \cdots, R_m)$ | Gamma shape 2 | Gexp shape 2 | Gamma shape 0.5 | Gexp shape 0.5 | Beta shape 0.5 | Glogistic shape 0.5 |
| 10 | 5 | 5,0,0,0,0 | 0.2791 | 0.3051 | 0.6306 | 0.6228 | 0.4047 | 0.4287 |
| | 5 | 0,5,0,0,0 | 0.3697 | 0.3902 | 0.6997 | 0.6941 | 0.5141 | 0.4954 |
| | 5 | 1,1,1,1,1 | 0.2701 | 0.2845 | 0.6334 | 0.6294 | 0.5385 | 0.3204 |
| | 5 | 0,0,0,5,0 | 0.3094 | 0.3239 | 0.6956 | 0.6920 | 0.5947 | 0.3773 |
| | 5 | 0,0,0,0,5 | 0.2078 | 0.2127 | 0.5144 | 0.5123 | 0.4673 | 0.2162 |
| | 7 | 3,0,0,0,0,0,0 | 0.4845 | 0.5207 | 0.8452 | 0.8377 | 0.5709 | 0.6710 |
| | 7 | 0,3,0,0,0,0,0 | 0.5337 | 0.5627 | 0.8648 | 0.8589 | 0.6340 | 0.7013 |
| | 7 | 1,0,0,1,0,0,1 | 0.4243 | 0.4511 | 0.8078 | 0.8025 | 0.6507 | 0.5458 |
| | 7 | 0,0,0,0,0,3,0 | 0.3899 | 0.4134 | 0.8010 | 0.7960 | 0.6495 | 0.5018 |
| | 7 | 0,0,0,0,0,0,3 | 0.4030 | 0.4203 | 0.7894 | 0.7858 | 0.6841 | 0.4712 |
| 20 | 5 | 15,0,0,0,0 | 0.2329 | 0.2516 | 0.6215 | 0.6138 | 0.4115 | 0.3808 |
| | 5 | 0,15,0,0,0 | 0.4198 | 0.4414 | 0.7764 | 0.7718 | 0.6383 | 0.5275 |
| | 5 | 3,3,3,3,3 | 0.2321 | 0.2395 | 0.6522 | 0.6510 | 0.6241 | 0.2322 |
| | 5 | 0,0,0,15,0 | 0.4109 | 0.4261 | 0.8168 | 0.8161 | 0.7920 | 0.4424 |
| | 5 | 0,0,0,0,15 | 0.1829 | 0.1864 | 0.5064 | 0.5058 | 0.4945 | 0.1741 |
| | 10 | 10,0,0,...,0,0,0 | 0.5755 | 0.6195 | 0.9563 | 0.9521 | 0.7378 | 0.7978 |
| | 10 | 0,10,0,...,0,0,0 | 0.6673 | 0.7045 | 0.9745 | 0.9722 | 0.8382 | 0.8387 |
| | 10 | 1,1,1,...,1,1,1 | 0.6273 | 0.6536 | 0.9759 | 0.9751 | 0.9482 | 0.6968 |
| | 10 | 0,0,0,...,0,10,0 | 0.3940 | 0.4178 | 0.8995 | 0.8976 | 0.8358 | 0.4746 |
| | 10 | 0,0,0,...,0,0,10 | 0.4700 | 0.4904 | 0.9319 | 0.9304 | 0.9022 | 0.4837 |
| 30 | 15 | 15,0,0,...,0,0,0 | 0.7951 | 0.8306 | 0.9973 | 0.9969 | 0.9316 | 0.9364 |
| | 15 | 0,15,0,...,0,0,0 | 0.8060 | 0.8416 | 0.9976 | 0.9974 | 0.9518 | 0.9435 |
| | 15 | 1,1,1,...,1,1,1 | 0.8422 | 0.8583 | 0.9990 | 0.9990 | 0.9967 | 0.8812 |
| | 15 | 0,0,0,...,0,15,0 | 0.4023 | 0.4307 | 0.9524 | 0.9516 | 0.9035 | 0.4853 |
| | 15 | 0,0,0,...,0,0,15 | 0.7475 | 0.7639 | 0.9958 | 0.9958 | 0.9925 | 0.7543 |
| | 20 | 10,0,0,...,0,0,0 | 0.9369 | 0.9556 | 1.0000 | 0.9999 | 0.9917 | 0.9888 |
| | 20 | 0,10,0,...,0,0,0 | 0.9154 | 0.9378 | 0.9997 | 0.9997 | 0.9917 | 0.9828 |
| | 20 | 1,0,1,...,0,1,0 | 0.9094 | 0.9261 | 0.9999 | 0.9999 | 0.9976 | 0.9645 |
| | 20 | 0,0,0,...,0,10,0 | 0.5607 | 0.5899 | 0.9876 | 0.9866 | 0.9516 | 0.6675 |
| | 20 | 0,0,0,...,0,0,10 | 0.9110 | 0.9250 | 1.0000 | 1.0000 | 0.9992 | 0.9352 |

TABLE VII
VALUES OF THE WINDOW SIZE FOR PARETO, LOG-NORMAL, AND WEIBULL DISTRIBUTIONS, RESPECTIVELY, AT $\alpha = 0.05$ AND $\alpha = 0.10$ LEVELS OF SIGNIFICANCE

| $n$ | $m$ | $w$ | | |
|---|---|---|---|---|
| 10 | 5 | 2 | 2 | 2 |
| | 7 | 3 | 3 | 3 |
| 20 | 5 | 2 | 2 | 2 |
| | 10 | 4 | 4 | 4 |
| 30 | 15 | 7 | 6 | 7 |
| | 20 | 9 | 8 | 9 |

Next, we can find $\hat{\alpha}$ by solving the likelihood equation

$$\frac{\partial \ln L}{\partial \alpha} = \frac{m}{\alpha} + m \ln \beta - \sum_{i=1}^{m} \ln x_{i:m:n} + \ln \beta \sum_{i=1}^{m} R_i - \sum_{i=1}^{m} R_i \ln x_{i:m:n} = 0,$$

which yields

$$\hat{\alpha} = \frac{m}{\sum_{i=1}^{m} (\ln x_{i:m:n} - \ln \hat{\beta})(1 + R_i)}. \tag{5}$$

## V. GOODNESS-OF-FIT TEST FOR LOG-NORMAL

Suppose we are interested in testing

$$H_0 : f^0 = \frac{1}{\sqrt{2\pi}\sigma y} e^{-(\ln y - \mu)^2 / 2\sigma^2} I(y > 0) \quad \text{vs.}$$
$$H_A : f^0 \neq \frac{1}{\sqrt{2\pi}\sigma y} e^{-(\ln y - \mu)^2 / 2\sigma^2} I(y > 0),$$

where $\mu$, and $\sigma$ are the scale, and shape parameters, respectively. If the random variable $Y$ has the density function under $H_0$, then $X = \ln Y$ has the normal distribution with location, and scale parameters as $\mu$, and $\sigma$, respectively. The KL information for a progressively Type-II censored data can be estimated, and the corresponding test statistic can be computed from (2) for which we may estimate the unknown parameters by MLEs. Because the MLEs cannot be obtained explicitly, we can instead use the approximate maximum likelihood estimates (AMLEs) proposed by Balakrishnan *et al.* [6] which are simple explicit estimators, and are nearly as efficient as the MLEs.

## VI. GOODNESS-OF-FIT TEST FOR WEIBULL

Suppose we are interested in testing

$$H_0 : f^0 = \alpha \lambda y^{\alpha - 1} e^{-\lambda y^{\alpha}} I(y > 0) \quad \text{vs.}$$
$$H_A : f^0 \neq \alpha \lambda y^{\alpha - 1} e^{-\lambda y^{\alpha}} I(y > 0), \tag{6}$$

TABLE VIII
CRITICAL VALUES OF THE TEST STATISTIC

| $n$ | $m$ | schemes $(R_1,\cdots,R_m)$ | Pareto distribution $\alpha=0.05$ | Pareto distribution $\alpha=0.1$ | Log-normal distribution $\alpha=0.05$ | Log-normal distribution $\alpha=0.1$ | Weibull distribution $\alpha=0.05$ | Weibull distribution $\alpha=0.1$ |
|---|---|---|---|---|---|---|---|---|
| 10 | 5 | 5,0,0,0,0 | -0.2442473 | -0.3189446 | -0.07591125 | -0.1203591 | 0.03230927 | -0.01556996 |
| | 5 | 0,5,0,0,0 | -0.1616068 | -0.2426989 | -0.01348657 | -0.06391083 | 0.1021960 | 0.05256394 |
| | 5 | 1,1,1,1,1 | -0.0480302 | -0.1162841 | 0.0697365 | 0.01734272 | 0.3067945 | 0.2496497 |
| | 5 | 0,0,0,5,0 | 0.01563756 | -0.03772818 | 0.1125063 | 0.07491981 | 0.3583388 | 0.3133855 |
| | 5 | 0,0,0,0,5 | -0.005274668 | -0.04817759 | 0.1103262 | 0.05839243 | 0.3583388 | 0.3831224 |
| | 7 | 3,0,0,0,0,0,0 | -0.2908719 | -0.3573684 | -0.1173803 | -0.1600330 | -0.01003275 | -0.06295593 |
| | 7 | 0,3,0,0,0,0,0 | -0.2548138 | -0.3354644 | -0.09529554 | -0.1330588 | 0.02145127 | -0.02007891 |
| | 7 | 1,0,0,1,0,0,1 | -0.0991883 | -0.1728702 | 0.01672273 | -0.01979114 | 0.2050205 | 0.1599670 |
| | 7 | 0,0,0,0,0,3,0 | -0.02234244 | -0.08870372 | 0.0841374 | 0.04880206 | 0.2679201 | 0.2263587 |
| | 7 | 0,0,0,0,0,0,3 | -0.01043444 | -0.0621921 | 0.1008859 | 0.05718918 | 0.3342688 | 0.2854822 |
| 20 | 5 | 15,0,0,0,0 | -0.1567559 | -0.1870175 | -0.00987938 | -0.0342441 | 0.1218705 | 0.09690488 |
| | 5 | 0,15,0,0,0 | -0.07519543 | -0.1258857 | 0.03949301 | 0.01766325 | 0.2171737 | 0.1949909 |
| | 5 | 3,3,3,3,3 | 0.002362862 | -0.0274577 | 0.1091025 | 0.08721069 | 0.6117008 | 0.5845867 |
| | 5 | 0,0,0,15,0 | 0.07806294 | 0.03087266 | 0.1521573 | 0.1197406 | 0.7344277 | 0.7112053 |
| | 5 | 0,0,0,0,15 | 0.007549286 | -0.01231158 | 0.09060948 | 0.06629348 | 0.7335832 | 0.7094029 |
| | 10 | 10,0,0,...,0,0,0 | -0.3167694 | -0.3603417 | -0.1546463 | -0.1762428 | -0.0001366207 | -0.03219522 |
| | 10 | 0,10,0,...,0,0,0 | -0.2860902 | -0.3369770 | -0.1308312 | -0.1504816 | 0.05841994 | 0.03534434 |
| | 10 | 1,1,1,...,1,1,1 | -0.07539401 | -0.1133254 | 0.01547473 | -0.002979384 | 0.3601191 | 0.3351809 |
| | 10 | 0,0,0,...,0,10,0 | 0.05493798 | 0.01575643 | 0.1247215 | 0.09863766 | 0.5402122 | 0.5072429 |
| | 10 | 0,0,0,...,0,0,10 | 0.000716371 | -0.02218124 | 0.07354464 | 0.04733289 | 0.5449744 | 0.5235059 |
| 30 | 15 | 15,0,0,...,0,0,0 | -0.3784075 | -0.4258246 | -0.1870219 | -0.1999568 | -0.02241883 | -0.03985451 |
| | 15 | 0,15,0,...,0,0,0 | -0.3650036 | -0.413455 | -0.1665779 | -0.1785467 | 0.05852065 | 0.03958939 |
| | 15 | 1,1,1,...,1,1,1 | -0.0991323 | -0.1404370 | 0.003488712 | -0.00827531 | 0.4184222 | 0.04013981 |
| | 15 | 0,0,0,...,0,15,0 | 0.06338891 | 0.02541773 | 0.1222373 | 0.09615582 | 0.647837 | 0.6198293 |
| | 15 | 0,0,0,...,0,0,15 | 0.003392069 | -0.01323904 | 0.05799687 | 0.03999982 | 0.627888 | 0.6104206 |
| | 20 | 10,0,0,...,0,0,0 | -0.4595367 | -0.5234324 | -0.2168436 | -0.231033 | -0.03978316 | -0.05837345 |
| | 20 | 0,10,0,...,0,0,0 | -0.4598187 | -0.5187202 | -0.2006536 | -0.2169001 | 0.023179135 | 0.005484341 |
| | 20 | 1,0,1,...,0,1,0 | -0.2467683 | -0.2942354 | -0.05003602 | -0.06774027 | 0.26997 | 0.2475547 |
| | 20 | 0,0,0,...,0,10,0 | 0.01026771 | -0.02888746 | 0.0971043 | 0.07380307 | 0.5268236 | 0.5016643 |
| | 20 | 0,0,0,...,0,0,10 | -0.003466335 | -0.02701657 | 0.06108307 | 0.04101777 | 0.5279432 | 0.5085816 |

where $\alpha > 0$, and $\lambda > 0$ are the shape, and scale parameters, respectively. If the random variable $Y$ has the density function under $H_0$, then $X = \ln Y$ has the extreme value distribution with p.d.f.

$$f^0 = \frac{1}{\sigma}\exp\left\{\left(\frac{x-\mu}{\sigma}\right) - \exp\left(\frac{(x-\mu)}{\sigma}\right)\right\}, \quad -\infty < x < \infty, \tag{7}$$

where $\mu = -\sigma \ln \lambda$, and $\sigma = 1/\alpha$. The density function in (7) is the extreme value density with location, and scale parameters as $\mu$, and $\sigma$, respectively; one may refer to Johnson *et al.* [13] for elaborate details on this distribution. Models (6) and (7) are equivalent models in the sense that a procedure developed under one model can be easily used for the other model. Although they are equivalent models, it is often easier to work with the extreme value model in (7) than the Weibull model in (6) because in model (7) the two parameters $\mu$, and $\sigma$ appear respectively as location, and scale parameters. Then, the KL-based test statistic for a progressively Type-II censored data can be computed from (2) by using the MLEs for the unknown parameters. Because these MLEs are not available explicitly, we may alternatively use the AMLEs proposed by Balakrishnan *et al.* [7] which are simple, explicit estimators, and are nearly as efficient as the MLEs.

## VII. IMPLEMENTATION OF TESTS

Because the sampling distribution of $T(w,n,m)$ is intractable, we determine the percentage points using 10,000 Monte Carlo simulations from Pareto, log-normal, and Weibull distributions. In determining the window size $w$, which depends on $n$, $m$ and $\alpha$, we define the optimal window size $w$ to be one which gives minimum critical points. However, we find from the simulated percentage points that the optimal window size $w$ varies much according to $m$ rather than $n$, and does not vary much according to $\alpha$, if $\alpha \leq 0.1$. In view of these observations, our recommended values of $w$ for different $m$ are presented in Table VII.

Then, after choosing the value of $w$, we simulated progressively Type-II censored samples from Pareto, log-normal, and Weibull distributions, and then calculated the value of $T(w,n,m)$, 10,000 times. Critical values were then determined as the percentage points of the obtained (empirical) distribution of the test statistic $T$, and these are presented in Table VIII. This empirical procedure of finding critical values of the proposed goodness-of-fit procedure is simple and straightforward, and hence it can be easily adopted to determine the critical values required for any other sample size and progressive censoring scheme.

### A. Power Results

As the proposed test statistic is related closely to the hazard function of the distribution, we consider the following alternatives according to the type of hazard function:

a) Monotone increasing hazard including Gamma and Gexp (shape parameter 2),

TABLE IX
PROGRESSIVELY CENSORED SAMPLE GENERATED FROM THE TIMES TO BREAKDOWN DATA ON INSULATING FLUIDS TESTED AT 34 KILOVOLTS, GIVEN BY VIVEROS AND BALAKRISHNAN [20]

| i | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| $x_{i:8:19}$ | 0.189987 | 0.77997 | 0.959925 | 1.30996 | 2.77986 | 4.84962 | 6.49999 | 7.35 |
| $R_i$ | 0 | 0 | 3 | 0 | 3 | 0 | 0 | 5 |

b) Monotone decreasing hazard including Gamma and Gexp (shape parameter 0.5), and
c) Nonmonotone hazard including Beta and Log-logistic (shape parameter 0.5).

We used 10,000 Monte Carlo simulations for $n = 10, 20$, and 30 to estimate the power of the proposed test statistic. The simulation results are summarized in Tables I and II for the Pareto distribution, Tables III and IV for the log-normal distribution, and Tables V and VI for Weibull distribution at $\alpha = 0.05$, and $\alpha = 0.10$ levels of significance.

We can see from Tables I –VI that the censoring schemes $(R_1 = n - m, R_2 = 0, \ldots, R_m = 0)$ and $(R_1 = 0, R_2 = n-m, R_3 = 0, \ldots, R_m = 0)$ show in general better power than other censoring schemes. Thus, early censoring situations seem to possess better power. In particular, early censoring seems to be especially useful when the alternative is either monotone increasing hazard functions, or monotone decreasing hazard functions. Also, in Tables I –VI, four out of the five progressive censoring schemes considered within each case correspond to single step censoring schemes. The variation in these four single step censoring schemes indeed show how the test performs in a stable manner when the nature of progressive censoring varies. In addition, we have included an equal censoring scheme in each case to demonstrate how the results for single step censoring compare to this equal censoring scheme.

## VIII. ILLUSTRATIVE EXAMPLE

In this section, we present an example to illustrate the use of the test $T(w, n, m)$ for testing the validity of the log-normal, Weibull, and Pareto distributions based on a progressively Type-II right censored sample. These data have been used in the progressive censoring literature quite extensively; see Balakrishnan [2].

Nelson [15] reported data on times to breakdown of an insulating fluid in an accelerated test conducted at various test voltages (see Tables VII and VIII). From these data, Viveros & Balakrishnan [20] produced a progressively Type-II censored sample of size $m = 8$ from $n = 19$ observations recorded at 34 kilovolts. These progressively censored data are presented in Table IX.

For these data, the estimates of the parameters (the AMLEs of the parameters for log-normal and Weibull distributions, and the MLEs for the Pareto distribution), of the test statistics, critical values, and $p$-values determined corresponding to the three distributions are presented in Table X. Clearly, neither the log-normal model nor the Pareto model is suitable for these data, while the Weibull distribution seems to provide an excellent fit ($p$-value being 0.7396) for these progressively Type-II censored data. This conclusion agrees with a similar finding of Viveros & Balakrishnan [20], and Balakrishnan *et al.* [8].

TABLE X
CRITICAL VALUES, ESTIMATES OF THE PARAMETERS, TEST STATISTICS, AND THE $p$-VALUES, WITH $w = 3$

| Distributions | Critical value | $\underline{\hat{\theta}}$ | $T(w, n, m)$ | $p$-value |
|---|---|---|---|---|
| Log-normal | 0.06669 | (1.64268, 1.71928) | 2.39503 | 0.0000 |
| Pareto | -0.01140 | (0.16241, 0.18998) | 0.93839 | 0.0000 |
| Weibull | 0.51512 | (0.43561, 0.46495) | 0.39397 | 0.7396 |

It is worth mentioning here that, if different items are withdrawn from the original sample due to the progressive censoring, then the resulting data will be different, and so the value of the test statistic will also be different. However, because the test statistic has been developed based on the joint distribution of the progressively Type-II censored sample, and that every set of realized failures produced by progressive censoring is equally likely, we would expect the result produced by one realization to be similar to the average one.

## ACKNOWLEDGMENT

The authors would like to thank the Associate Editor, Dr. L. Walls, and three anonymous referees for their constructive comments and useful suggestions on an earlier version of this manuscript which led to this improved version.

## REFERENCES

[1] I. Arizono and H. Ohta, "A test for normality based on Kullback-Leibler information," *The American Statistician*, vol. 43, pp. 20–23, 1989.
[2] N. Balakrishnan, "Progressive censoring methodology: An appraisal (with discussions)," *Test*, vol. 16, pp. 211–296, 2007.
[3] N. Balakrishnan and R. Aggarwala, *Progressive Censoring: Theory, Methods, and Applications*. Boston: Birkhäuser, 2000.
[4] N. Balakrishnan and A. C. Cohen, *Order Statistics and Inference: Estimation Methods*. San Diego: Academic Press, 1991.
[5] N. Balakrishnan, A. Habibi Rad, and N. R. Arghami, "Testing exponentiality based on Kullback-Leibler information with progressively Type-II censored data," *IEEE Trans. Reliability*, vol. 56, pp. 349–356, 2007.
[6] N. Balakrishnan, N. Kannan, C.-T. Lin, and H. K. T. Ng, "Point and interval estimation for Gaussian distribution, based on progressively Type-II censored samples," *IEEE Trans. Reliability*, vol. 52, pp. 90–95, 2003.
[7] N. Balakrishnan, N. Kannan, C.-T. Lin, and S. J. S. Wu, "Inference for the extreme value distribution under progressive Type-II censoring," *Journal of Statistical Computation and Simulation*, vol. 74, pp. 25–45, 2004.
[8] N. Balakrishnan, H. K. T. Ng, and N. Kannan, "Goodness-of-fit tests based on spacings for progressively Type-II censored data from a general location-scale distribution," *IEEE Trans. Reliability*, vol. 53, pp. 349–356, 2004.
[9] A. C. Cohen, *Truncated and Censored Samples: Theory and Applications*. New York: Marcel Dekker, 1991.
[10] E. J. Dudewicz and E. C. van der Meulen, "Entropy-based tests of uniformity," *Journal of the American Statistical Association*, vol. 76, pp. 967–974, 1981.
[11] N. Ebrahimi and M. Habibullah, "Testing exponentiality based on Kullback-Leibler information," *Journal of the Royal Statistical Society, Series B*, vol. 54, pp. 739–748, 1992.
[12] D. V. Gokhale, "On entropy-based goodness-of-fit tests," *Computational Statistics & Data Analysis*, vol. 1, pp. 157–165, 1983.

[13] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, Second ed. New York: John Wiley & Sons, 1995.
[14] S. Kullback, *Information Theory and Statistics*. New York: John Wiley & Sons, 1959.
[15] W. Nelson, *Applied Life Data Analysis*. New York: John Wiley & Sons, 1982.
[16] S. Park, "The entropy of consecutive order statistics," *IEEE Trans. Information Theory*, vol. 41, pp. 2003–2007, 1995.
[17] S. Park, "Testing exponentiality based on the Kullback-Leibler information with the Type-II censored data," *IEEE Trans. Reliability*, vol. 54, pp. 22–26, 2005.
[18] C. E. Shannon, "A mathematical theory of communications," *Bell System Technical Journal*, vol. 27, pp. 379–423, 1948.
[19] O. Vasicek, "A test for normality based on sample entropy," *Journal of the Royal Statistical Society, Series B*, vol. 38, pp. 730–737, 1976.
[20] R. Viveros and N. Balakrishnan, "Interval estimation of parameters of life from progressively censored data," *Technometrics*, vol. 36, pp. 84–91, 1994.
[21] K. M. Wong and S. Chan, "The entropy of ordered sequences and order statistics," *IEEE Trans. Information Theory*, vol. 36, pp. 276–284, 1990.

**Arezou Habibi Rad** is an Assistant Professor of Statistics at Ferdowsi University of Mashhad, Iran. Her research interests include censored data analysis, record values, and evidential inference.

**Fatemeh Yousefzadeh** is an Assistant Professor of Statistics at Birjand University, Iran. Her research interests include censored data analysis, record values, and Monte Carlo simulations.

**Narayanaswamy Balakrishnan** is a Professor of Statistics at McMaster University, Hamilton, Ontario, Canada. He received his B.Sc., and M.Sc. degrees in Statistics from the University of Madras, India, in 1976, and 1978, respectively. He finished his Ph.D. in Statistics from Indian Institute of Technology, Kanpur, India, in 1981. He is a *Fellow of the American Statistical Association*, and a *Fellow of the Institute of Mathematical Statistics*. He is currently the Editor-in-Chief of *Communications in Statistics*, and Executive Editor of *Journal of Statistical Planning and Inference*. His research interests include distribution theory, ordered data analysis, censoring methodology, reliability, survival analysis, nonparametric inference, and statistical quality control.