APPLICATIONES MATHEMATICAE
28,2 (2001), pp. 151–168

Kerwin Morris (Adelaide)
Dominik Szynal (Lublin)

GOODNESS-OF-FIT TESTS USING CHARACTERIZATIONS OF CONTINUOUS DISTRIBUTIONS

Abstract. Using characterization conditions of continuous distributions in terms of moments of order statistics and moments of record values we present new goodness-of-fit techniques.

1. Introduction and preliminaries. Let $(X_1, \ldots, X_n)$ be a sample from a continuous distribution $F(x) = P[X \le x]$, $x \in \mathbb{R}$, and let $X_{k:n}$ denote the kth smallest order statistic of the sample. We construct goodness-of-fit tests for continuous distributions using characterizations of distributions via moments of order statistics and moments of record values (cf. [2]–[5], [10]). The results presented extend the tests for uniformity and exponentiality discussed in [6] and [7]. Moreover, we give the proof of statements on tests for exponentiality announced in [7]. We include a theorem on the asymptotic effect of substituting estimators for parameters in the tests proposed here. It can be used, among other things, to construct a test for normality.

(O) *Characterizations in terms of moments of order statistics.* We use the characterization conditions contained in the following theorems.

Theorem 1 (cf. [10], [3]). *Let n, k, l be given integers such that $n \ge k \ge l \ge 1$. Assume that G is a nondecreasing right-continuous function from $\mathbb{R}$ to $\mathbb{R}$. Then $F(x) = G(x)$ on $I(F)$ (the minimal interval containing the support of F) and F is continuous on $\mathbb{R}$ iff*

$$\frac{(k-l)!}{(n-l+1)!} EG^{2l}(X_{k+1-l:n+1-l}) - \frac{2k!}{(n+1)!} EG^{l}(X_{k+1:n+1}) + \frac{(k+l)!}{(n+l+1)!} = 0. \tag{1.1}$$

2000 *Mathematics Subject Classification*: Primary 62E10, 62F03.
Key words and phrases: goodness-of-fit tests, characterizations, order statistics, record values, uniform, exponential, Weibull, Pareto, geometric and logarithmic distributions.

Theorem 2 (cf. [5]). *Under the assumptions of Theorem* 1, $F(x) = G(x)$ *on* $I(F)$ *and* F *is continuous on* $\mathbb{R}$ *iff*

$$EG^l(X_{k+1:n+1}) = \frac{(k+l)!(n+1)!}{k!(n+l+1)!},$$
$$EG^{2l}(X_{k+1-l:n+1-l}) = \frac{(k+l)!(n-l+1)!}{(k-l)!(n+l+1)!}. \tag{1.2}$$

Note that Theorem 2 is a consequence of Theorem 1, since (1.1) implies $F = G$ implies (1.2) implies (1.1).

Corollary 1. $X \sim F$ *and* F *is continuous iff*

$$EF(X_{2:2}) - EF^2(X) = \tfrac{1}{3} \tag{1.3}$$

or

$$EF(X_{2:2}) = \tfrac{2}{3}, \quad EF^2(X) = \tfrac{1}{3}. \tag{1.4}$$

In particular:

(a) $X \sim U(\alpha, \beta)$ (uniform distribution), i.e. $F(x) = (x-\alpha)/(\beta-\alpha)$, $\alpha < x < \beta$, iff

$$E[(X_{2:2}-\alpha)/(\beta-\alpha)] - E[(X-\alpha)/(\beta-\alpha)]^2 = \tfrac{1}{3}$$

or

$$E[(X_{2:2}-\alpha)/(\beta-\alpha)] = \tfrac{2}{3}, \quad E[(X-\alpha)/(\beta-\alpha)]^2 = \tfrac{1}{3},$$

(b) $X \sim \mathrm{Exp}(\alpha)$ (exponential distribution), i.e. $F(x) = 1 - \exp(-\alpha x)$, $x > 0$, $\alpha > 0$, iff

$$E(1-\exp(-\alpha X_{2:2})) - E(1-\exp(-\alpha X))^2 = \tfrac{1}{3}$$

or

$$E(1-\exp(-\alpha X_{2:2})) = \tfrac{2}{3}, \quad E(1-\exp(-\alpha X))^2 = \tfrac{1}{3}.$$

(R) *Characterization conditions in terms of moments of record values.* Let $\{X_n,\ n \ge 1\}$ be a sequence of i.i.d. random variables with cdf F and pdf f. For a fixed $k \ge 1$ we define the sequence $U_k(1), U_k(2), \ldots$ of k-(*upper*) *record times* of $X_1, X_2, \ldots$ as follows:

$$U_k(1) = 1,$$
$$U_k(n) = \min\{j > U_k(n-1) : X_{j:j+k-1} > X_{U_k(n-1):U_k(n-1)+k-1}\}, \quad n = 2, 3, \ldots$$

Write

$$Y_n^{(k)} := X_{U_k(n):U_k(n)+k-1}, \quad n \ge 1.$$

The sequence $\{Y_n^{(k)},\ n \ge 1\}$ is called the sequence of k-(*upper*) *record values* of the above sequence. For convenience we also take $Y_0^{(k)} = 0$ and note that $Y_1^{(k)} = X_{1:k} = \min(X_1, \ldots, X_k)$ (cf. [1]).

We see that for $k = 1, 2, \ldots$, the sequences $\{Y_n^{(k)},\ n \ge 1\}$ of kth record values can be obtained from $\{X_n,\ n \ge 1\}$ by inspecting successively the samples X_1, (X_1, X_2), (X_1, X_2, X_3), and so on. For $k = 1$, $Y_1^{(1)} = X_1$, and the following terms are obtained by looking at the maxima of the successive samples; $Y_2^{(1)}$ is the first maximum that exceeds $Y_1^{(1)}$, $Y_3^{(1)}$ is the first maximum that exceeds $Y_2^{(1)}$, and so on. For $k = 2$, $Y_1^{(2)} = \min(X_1, X_2)$, and the following terms are obtained by looking at the next-to-largest values in the successive samples: $Y_2^{(2)}$ is the first next-to-largest value that exceeds $Y_1^{(2)}$, $Y_3^{(2)}$ is the next-to-largest value that exceeds $Y_2^{(2)}$, and so on. And generally, $Y_1^{(k)} = \min(X_1, \ldots, X_k) = X_{1:k}$, and the following kth record values are obtained by looking at the kth largest values in successive samples, i.e., looking at the order statistics $X_{2:k+1}$ from $(X_1, \ldots, X_{k+1})$, $X_{3:k+2}$ from $(X_1, \ldots, X_{k+2})$, and so on.

We have the following characterizations.

THEOREM 3 (cf. [4]). *Let $\{X_n,\ n \ge 1\}$ be a sequence of i.i.d. random variables with cdf F. Assume that G is a nondecreasing right-continuous function from $\mathbb{R}$ to $(-\infty, 1]$, and let n, k, l be given integers such that $k \ge 1$ and $n \ge l \ge 1$. Then $F(x) = G(x)$ on $I(F)$ iff*

$$(1.5)\quad k^{2l}(n-l)!E[-\log(1 - G(Y_{n-l+1}^{(k)}))]^{2l} - 2n!k^l E[-\log(1 - G(Y_{n+1}^{(k)}))]^l + (n+l)! = 0.$$

THEOREM 4 (cf. [5], [4]). *Under the assumptions of Theorem 3, $F(x) = G(x)$ on $I(F)$ iff*

$$E[-\log(1 - G(Y_{n+1}^{(k)}))]^l = \frac{(n+l)!}{n!k^l},$$

$$E[-\log(1 - G(Y_{n-l+1}^{(k)}))]^{2l} = \frac{(n+l)!}{(n-l)!k^{2l}}.$$

Following the observation after Theorem 2 we see that Theorem 4 is a consequence of Theorem 3.

COROLLARY 2. *$X \sim F$ and F is continuous iff*

$$(1.6)\quad E[-\log(1 - F(Y_1^{(k)}))]^2 - \frac{2}{k}E[-\log(1 - F(Y_2^{(k)}))] + \frac{2}{k^2} = 0$$

or

$$(1.7)\quad E[-\log(1 - F(Y_2^{(k)}))] = \frac{2}{k}, \quad E[-\log(1 - F(Y_1^{(k)}))]^2 = \frac{2}{k^2}.$$

In particular:

(a) $X \sim U(0,1)$ iff

$$E[-\log(1-Y_1^{(k)})]^2 - \frac{2}{k}E[-\log(1-Y_2^{(k)})] + \frac{2}{k^2} = 0$$

or

$$E[-\log(1-Y_2^{(k)})] = \frac{2}{k}, \quad E[-\log(1-Y_1^{(k)})]^2 = \frac{2}{k^2},$$

(b) $X \sim \mathrm{Exp}(\alpha)$ iff

$$\alpha^2 E(Y_1^{(k)})^2 - \frac{2}{k^2}\alpha E(Y_2^{(k)}) + \frac{2}{k^2} = 0$$

or

$$EY_2^{(k)} = \frac{2}{\alpha k}, \quad E(Y_1^{(k)})^2 = \frac{2}{\alpha^2 k^2}.$$

2. Goodness-of-fit tests based on characterizations via moments of order statistics. The cases when parameters of F are specified and unknown will be treated separately.

(A) *Parameters of F are specified.* First we construct goodness-of-fit tests based on the characterization in (1.1) (see also (1.3)) which we can write in the form

$$E(F(X_{2:2})) - \tfrac{1}{2}(E(F^2(X_1) + F^2(X_2))) = \tfrac{1}{3}$$

where X_1 and X_2 are i.i.d. as X.

Let $(X_1,\ldots,X_{2n})$ be a sample from F, where F is continuous and strictly increasing. Define

$$\begin{aligned} Y_j &= F^2(X_{2j-1}) + F^2(X_{2j}), \\ Z_j &= F(\max(X_{2j-1}, X_{2j})), \quad j = 1,\ldots,n. \end{aligned}$$

Then $Y_1,\ldots,Y_n$ are i.i.d. and $Z_1,\ldots,Z_n$ are i.i.d. Writing $Y := Y_1 = F^2(X_1)+F^2(X_2)$, $Z := Z_1 = F(\max(X_1,X_2))$ we state the following result.

LEMMA 1. *Under the above assumptions, the density function of (Y,Z) is given by*

$$f(y,z) = \begin{cases} 1/\sqrt{y-z^2}, & 0 \le z \le 1,\ z^2 < y \le 2z^2, \\ 0, & \text{otherwise}, \end{cases}$$

and

$$EY = \tfrac{2}{3},\ \mathrm{Var}(Y) = \tfrac{8}{45},\ EZ = \tfrac{2}{3},\ \mathrm{Var}(Z) = \tfrac{1}{18},\ \mathrm{Cov}(Y,Z) = \tfrac{4}{45}.$$

Now we define

$$R_j = Z_j - \tfrac{1}{2}Y_j, \quad j = 1,\ldots,n.$$

We see that

$$ER_j = EZ_j - \tfrac{1}{2}EY_j = \tfrac{1}{3},$$
$$\operatorname{Var} R_j = \operatorname{Var} Z_j + \tfrac{1}{4}\operatorname{Var} Y_j - \operatorname{Cov}(Z_j, Y_j) = \tfrac{1}{90}, \qquad j = 1, \dots, n.$$

Write

$$\overline{R_n} = \frac{1}{n}\sum_{j=1}^{n} R_j;$$

then by the CLT

$$3\sqrt{10n}\left(\overline{R_n} - \tfrac{1}{3}\right) \xrightarrow{D} V \sim N(0,1),$$

and hence

(2.1) $$D_n^{(1)} := 45 \cdot 2n\left(\overline{R_n} - \tfrac{1}{3}\right)^2 \xrightarrow{D} \chi^2(1),$$

and so $D_n^{(1)}$ provides a simple asymptotic test of the hypothesis $X \sim F$.

Setting $X_j^* = \max(X_{2j-1}, X_{2j})$, $j = 1, \dots, n$, we note that $D_n^{(1)}$ in (2.1) has the form

(2.2) $$D_n^{(1)} = 45 \cdot 2n\left(\frac{1}{n}\sum_{j=1}^{n} F(X_j^*) - \frac{1}{2n}\sum_{j=1}^{2n} F^2(X_j) - \frac{1}{3}\right)^2.$$

Next we construct goodness-of-fit tests based on the characterization in (1.2) (see also (1.4)), which we write in the form

$$EF(\max(X_1, X_2)) = \tfrac{2}{3}, \quad EF^2(X_1) = \tfrac{1}{3}.$$

Define

$$\mathbf{W}_j = \begin{pmatrix} Y_j \\ Z_j \end{pmatrix}, \qquad j = 1, \dots, n,$$
$$\mu = E\mathbf{W}_1 = \begin{pmatrix} 2/3 \\ 2/3 \end{pmatrix} = \frac{2}{3}\begin{pmatrix} 1 \\ 1 \end{pmatrix},$$
$$\Sigma := \operatorname{Var}(\mathbf{W}_1) = E(\mathbf{W}_1 - E\mathbf{W}_1)(\mathbf{W}_1 - E\mathbf{W}_1)' = \begin{pmatrix} 8/45 & 4/45 \\ 4/45 & 1/18 \end{pmatrix},$$

and write $\overline{\mathbf{W}_n} = n^{-1}\sum_{j=1}^{n} \mathbf{W}_j$. The CLT says that

(2.3) $$\sqrt{n}\,(\overline{\mathbf{W}_n} - \mu) \xrightarrow{D} \mathbf{V} \sim N(0, \Sigma),$$

whence

$$D_n^{(2)} := n(\overline{\mathbf{W}_n} - \mu)'\Sigma^{-1}(\overline{\mathbf{W}_n} - \mu) \xrightarrow{D} \mathbf{V}'\Sigma^{-1}\mathbf{V} \sim \chi^2(2).$$

But $D_n^{(2)}$ is a reasonable measure of the "size" of $(\overline{\mathbf{W}_n} - \mu)$ and so by (2.3) provides a test of the hypothesis that X has the distribution function F. And since

$$\Sigma^{-1} = 45\begin{pmatrix} 5/8 & -1 \\ -1 & 2 \end{pmatrix},$$

it follows that in extended form

$$D_n^{(2)} = 45n\big[\tfrac{5}{8}\big(\overline{Y_n} - \tfrac{2}{3}\big)^2 + 2\big(\overline{Z_n} - \tfrac{2}{3}\big)^2 - 2\big(\overline{Y_n} - \tfrac{2}{3}\big)\big(\overline{Z_n} - \tfrac{2}{3}\big)\big]. \tag{2.4}$$

In terms of X_j^*, $D_n^{(2)}$ in (2.4) has the form

$$D_n^{(2)} = 45 \cdot 2n\big[\tfrac{1}{4}\big(\overline{F^2(X_{2n})} - \tfrac{1}{3}\big)^2 + \big(\overline{F^2(X_{2n})} - \overline{F(X_n^*)} + \tfrac{1}{3}\big)^2\big]. \tag{2.5}$$

By (2.5) and (2.2) we have

Lemma 2.

$$D_n^{(2)} = \tfrac{45}{4} \cdot 2n\big(\overline{F^2(X_{2n})} - \tfrac{1}{3}\big)^2 + D_n^{(1)}.$$

Special cases:

(a) If $X \sim U(\alpha, \beta)$ then

$$D_n^{(1)} = 45 \cdot 2n\bigg(\frac{1}{(\beta-\alpha)^2}\overline{X_{2n}^2} - \frac{\beta+\alpha}{(\beta-\alpha)^2}\overline{X_{2n}} - \frac{1}{\beta-\alpha}\overline{X_n^+} + \frac{\alpha\beta}{\beta-\alpha} + \frac{1}{3}\bigg)^2,$$

$$D_n^{(2)} = \frac{45}{4} \cdot 2n\bigg(\frac{\overline{X_{2n}^2}}{(\beta-\alpha)^2} - 2\alpha\frac{\overline{X_{2n}}}{(\beta-\alpha)^2} + \frac{\alpha^2}{(\beta-\alpha)^2} - \frac{1}{3}\bigg)^2 + D_n^{(1)}.$$

Remark. If $X \sim U(0, \beta)$ then

$$D_n^{(1)} = 45 \cdot 2n\big(\overline{X_{2n}^2}/\beta^2 - \overline{X_{2n}}/\beta - \overline{X_n^+}/\beta + \tfrac{1}{3}\big)^2,$$
$$D_n^{(2)} = \tfrac{45}{4} \cdot 2n\big(\overline{X_{2n}^2}/\beta^2 - \tfrac{1}{3}\big)^2 + D_n^{(1)}.$$

(b) If $X \sim \mathrm{Pow}(\alpha)$ (power distribution), i.e. $F(x) = 1 - (1 - x/\alpha)^\alpha$, $0 \le x \le \alpha$, $0 < \alpha \le 1$, then

$$D_n^{(1)} = 45 \cdot 2n\bigg(\frac{1}{2n}\sum_{j=1}^{2n}(1 - (1 - X_j/\alpha)^\alpha)^2 + \frac{1}{n}\sum_{j=1}^{n}(1 - X_j^*/\alpha)^\alpha - \frac{2}{3}\bigg)^2,$$

$$D_n^{(2)} = \frac{45}{4} \cdot 2n\bigg(\frac{1}{2n}\sum_{j=1}^{2n}(1 - (1 - X_j/\alpha)^\alpha)^2 - \frac{1}{3}\bigg)^2 + D_n^{(1)}.$$

(c) If $X \sim \mathrm{Exp}(\alpha)$ then

$$D_n^{(1)} = 45 \cdot 2n\bigg(\frac{1}{2n}\sum_{j=1}^{2n}(1 - e^{-\alpha X_j})^2 + \frac{1}{n}\sum_{j=1}^{n}e^{-\alpha X_j^*} - \frac{2}{3}\bigg)^2,$$

$$D_n^{(2)} = \frac{45}{4} \cdot 2n\bigg(\frac{1}{2n}\sum_{j=1}^{2n}(1 - e^{-\alpha X_j})^2 - \frac{1}{3}\bigg)^2 + D_n^{(1)}.$$

(d) If $X \sim W(\beta,\alpha)$ (Weibull distribution), i.e. $F(x) = 1 - \exp(-\alpha x^{\beta})$, $x > 0$, $\alpha > 0$, $\beta > 0$, then

$$D_n^{(1)} = 45 \cdot 2n\biggl(\frac{1}{2n}\sum_{j=1}^{2n}(1 - e^{-\alpha X_j^{\beta}})^2 + \frac{1}{n}\sum_{j=1}^{n} e^{-\alpha (X_j^*)^{\beta}} - \frac{2}{3}\biggr)^2,$$

$$D_n^{(2)} = \frac{45}{4} \cdot 2n\biggl(\frac{1}{2n}\sum_{j=1}^{2n}(1 - e^{-\alpha X_j^{\beta}})^2 - \frac{1}{3}\biggr)^2 + D_n^{(1)}.$$

(e) If $X \sim \mathrm{Par}_S(\alpha,\sigma)$ (single-parameter Pareto distribution), i.e. $F(x) = 1 - (\sigma/x)^{\alpha}$, $x > \sigma$, $\alpha > 0$, $\sigma > 0$, then

$$D_n^{(1)} = 45 \cdot 2n\biggl(\frac{1}{2n}\sum_{j=1}^{2n}\biggl(1 - \biggl(\frac{\sigma}{X_j}\biggr)^{\alpha}\biggr)^2 + \frac{1}{n}\sum_{j=1}^{n}(\sigma/X_j^*)^{\alpha} - \frac{2}{3}\biggr)^2,$$

$$D_n^{(2)} = \frac{45}{4} \cdot 2n\biggl(\frac{1}{2n}\sum_{j=1}^{2n}\biggl(1 - \biggl(\frac{\sigma}{X_j}\biggr)^{\alpha}\biggr)^2 - \frac{1}{3}\biggr)^2 + D_n^{(1)}.$$

(f) If $X \sim \mathrm{Par}_T(\alpha,\theta)$ (two-parameter Pareto distribution), i.e. $F(x) = 1 - \bigl(\frac{\theta}{x+\theta}\bigr)^{\alpha}$, $x > 0$, $\alpha > 0$, $\theta > 0$, then

$$D_n^{(1)} = 45 \cdot 2n\biggl(\frac{1}{2n}\sum_{j=1}^{2n}\biggl(1 - \biggl(\frac{\theta}{X_j + \theta}\biggr)^{\alpha}\biggr)^2 + \frac{1}{n}\sum_{j=1}^{n}\biggl(\frac{\theta}{X_j^* + \theta}\biggr)^{\alpha} - \frac{2}{3}\biggr)^2,$$

$$D_n^{(2)} = \frac{45}{4} \cdot 2n\biggl(\frac{1}{2n}\sum_{j=1}^{2n}\biggl(1 - \biggl(\frac{\theta}{X_j + \theta}\biggr)^{\alpha}\biggr)^2 - \frac{1}{3}\biggr)^2 + D_n^{(1)}.$$

(g) If $X \sim \mathrm{Log}(\alpha,\beta)$ (logistic distribution), i.e.

$$F(x) = [1 + \exp(-(x-\alpha)/\beta)]^{-1}, \qquad -\infty < x < \infty,\ \alpha \in \mathbb{R},\ \beta > 0,$$

then

$$\begin{aligned} D_n^{(1)} = 45 \cdot 2n\biggl(&\frac{1}{2n}\sum_{j=1}^{2n}(1 + \exp(-(X_j - \alpha)/\beta))^{-2} \\ &- \frac{1}{n}\sum_{j=1}^{n}(1 + \exp(-(X_j^* - \alpha)/\beta))^{-1} + \frac{1}{3}\biggr)^2, \end{aligned}$$

$$D_n^{(2)} = \frac{45}{4} \cdot 2n\biggl(\frac{1}{2n}\sum_{j=1}^{2n}(1 + \exp(-(X_j - \alpha)/\beta))^{-2} - \frac{1}{3}\biggr)^2 + D_n^{(1)}.$$

(B) *Unknown parameters.* We discuss asymptotic tests obtained from $D_n^{(1)}$ and $D_n^{(2)}$ in (A) when parameters are replaced by estimators.

PROPOSITION 1. *Goodness-of-fit tests for* $F(x) = x/\beta$, $x \in (0, \beta)$, $\beta > 0$, *are given by*

$$\widehat{D}_n^{(1)} := D_n^{(1)}(\widehat{\beta}_n) = 45 \cdot 2n\left(\frac{\overline{X_{2n}^2}}{\widehat{\beta}_n^2} - \frac{\overline{X_{2n}}}{\widehat{\beta}_n} - \frac{\overline{X_n^+}}{\widehat{\beta}_n} + \frac{1}{3}\right)^2 \xrightarrow{D} \chi^2(1),$$

$$\widehat{D}_n^{(2)} := D_n^{(2)}(\widehat{\beta}_n) = \frac{45}{4} \cdot 2n\left(\frac{\overline{X_{2n}^2}}{\widehat{\beta}_n^2} - \frac{1}{3}\right)^2 + D_n^{(1)}(\widehat{\beta}_n) \xrightarrow{D} \chi^2(2),$$

where $\widehat{\beta}_n = \max(X_1, \ldots, X_{2n})$.

PROPOSITION 2. *Goodness-of-fit tests for* $F(x) = \frac{x-\alpha}{\beta-\alpha}$, $x \in (\alpha, \beta)$, $\alpha < \beta$, *are given by*

$$\widehat{D}_n^{(1)} := D_n^{(1)}(\widehat{\alpha}_n, \widehat{\beta}_n) = 45 \cdot 2n\left(\frac{\overline{X_{2n}^2}}{(\widehat{\beta}_n - \widehat{\alpha}_n)^2} - (\widehat{\beta}_n + \widehat{\alpha}_n)\frac{\overline{X_{2n}}}{(\widehat{\beta}_n - \widehat{\alpha}_n)^2} - \frac{\overline{X_n^+}}{(\widehat{\beta}_n - \widehat{\alpha}_n)} + \frac{\widehat{\alpha}_n \widehat{\beta}_n}{(\widehat{\beta}_n - \widehat{\alpha}_n)^2} + \frac{1}{3}\right)^2 \xrightarrow{D} \chi^2(1),$$

$$\widehat{D}_n^{(2)} := D_n^{(2)}(\widehat{\alpha}_n, \widehat{\beta}_n) = \frac{45}{4} \cdot 2n\left(\frac{\overline{X_{2n}^2}}{(\widehat{\beta}_n - \widehat{\alpha}_n)^2} - 2\frac{\widehat{\alpha}_n \overline{X_{2n}}}{(\widehat{\beta}_n - \widehat{\alpha}_n)} + \frac{\widehat{\alpha}_n^2}{(\widehat{\beta}_n - \widehat{\alpha}_n)^2} - \frac{1}{3}\right)^2 + D_n^{(1)}(\widehat{\alpha}_n, \widehat{\beta}_n) \xrightarrow{D} \chi^2(2),$$

where $\widehat{\beta}_n = \max(X_1, \ldots, X_{2n})$ *and* $\widehat{\alpha}_n = \min(X_1, \ldots, X_{2n})$.

The proofs of Propositions 1 and 2 are given in [6] and [7]. For the following propositions concerning exponential and normal distributions we use a general theorem based on results in [8] and [9].

THEOREM 5 ([8]). *Let* $\widehat{T}_n = T_n(X_1, \ldots, X_n; \widehat{\lambda}_n)$, *where* $\widehat{\lambda}_n = \widehat{\lambda}_n(X_1, \ldots, X_n)$ *is an estimator of a parameter* λ *of the distribution of* X, *and let* $T_n = T_n(X_1, \ldots, X_n; \lambda)$ (*here* T_n, λ *and* $\widehat{\lambda}_n$ *may be vectors*). *Suppose that*:

(i) *For each* λ,

$$\sqrt{n}\begin{pmatrix} T_n \\ \widehat{\lambda}_n - \lambda \end{pmatrix} \xrightarrow{D} T \sim N(0, V),$$

where

$$V = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}$$

and V_{22} *is nonsingular.*

(ii) *There is a matrix* B, *possibly depending continuously on* λ, *such that*

$$\sqrt{n}\,\widehat{T}_n = \sqrt{n}\,T_n + B\sqrt{n}\,(\widehat{\lambda}_n - \lambda) + o_p(1).$$

(iii) $\widehat{\lambda}_n$ *is asymptotically efficient* (*cf.* [8]).

Then

$$\sqrt{n}\,\widehat{T}_n \xrightarrow{D} T^* \sim N(0, V_{11} - BV_{22}B'). \tag{2.6}$$

Note that (ii) is satisfied when T_n is differentiable in λ, and then

$$B = \lim_{n\to\infty} E\Big[\frac{\partial}{\partial\lambda} T_n\Big].$$

The following result is a consequence of Theorem 5.

THEOREM 6. *Let* $(X_1, \ldots, X_{2n})$ *be a sample with an absolutely continuous distribution function* $F(x;\lambda)$ *differentiable with respect to the* $m \times 1$ *vector* λ. *Set*

$$\overline{\mathbf{W}_n} := \begin{pmatrix} \overline{Y_n} \\ \overline{Z_n} \end{pmatrix} := \begin{pmatrix} \overline{Y_n(\lambda)} \\ \overline{Z_n(\lambda)} \end{pmatrix} = \overline{W_n(\lambda)},$$

where

$$\overline{Y_n} = \frac{1}{n}\sum_{j=1}^{2n} F^2(X_j;\lambda), \quad \overline{Z_n} = \frac{1}{n}\sum_{j=1}^{n} F(X_j^*;\lambda),$$

and $X_j^* = \max(X_{2j-1}, X_{2j})$, $j = 1, \ldots, n$. *Write*

$$\widehat{W}_n = \overline{W_n(\widehat{\lambda}_{2n})} = \begin{pmatrix} \widehat{Y}_n \\ \widehat{Z}_n \end{pmatrix},$$

where

$$\widehat{Y}_n := \overline{Y_n(\widehat{\lambda}_{2n})}, \quad \widehat{Z}_n := \overline{Z_n(\widehat{\lambda}_{2n})}.$$

and $\widehat{\lambda}_{2n}$ *is the MLE of* λ. *Suppose that* F *is such that the MLE* $\widehat{\lambda}_{2n}$ *is "regular" in the sense that*

$$\sqrt{2n}\,(\widehat{\lambda}_{2n} - \lambda) \xrightarrow{D} \gamma \sim N(0, I^{-1}),$$

where $I = I(\lambda)$ *is the information matrix for* λ *based on a single observation. Then*

$$\sqrt{n}\,(\overline{W_n}(\widehat{\lambda}_{2n}) - \mu) \xrightarrow{D} W \sim N(0, \Sigma_1), \tag{2.7}$$

$$\widehat{D}_n^{(1)} := 45 \cdot 2n\big(\overline{\widehat{F}(X_n^*)} - \overline{\widehat{F}^2(X_{2n})} - \tfrac{1}{3}\big)^2 \to \chi^2(1), \tag{2.8}$$

$$\widehat{D}_n^{(2)} := \frac{45}{4-b} 2n\big(\overline{\widehat{F}^2(X_{2n})} - \tfrac{1}{3}\big)^2 + \widehat{D}_n^{(1)} \to \chi^2(2), \tag{2.9}$$

where $\Sigma_1 = \Sigma - B(2I)^{-1}B$, μ *and* Σ *are taken from* (2.3), $\widehat{F}(x) := F(x, \widehat{\lambda})$, *and*

$$B = 2\begin{pmatrix} 2 \\ 1 \end{pmatrix} d',$$

where

$$d := E\left(F(X;\lambda)\frac{dF(X;\lambda)}{d\lambda}\right) \quad \textit{is } m\times 1,$$

and

$$b := b(\lambda) = 180 d' I^{-1} d.$$

Proof. The statement (2.7) follows directly from (2.3) and (2.6). Now note that

$$E\frac{\partial\overline{Y_n}}{\partial\lambda_j} = 2E\frac{\partial F^2(X;\lambda)}{\partial\lambda_j} = 4E\left(F\frac{\partial F}{\partial\lambda_j}\right)$$
$$= 4\int F(x;\lambda)\frac{\partial F}{\partial\lambda_j} f(x;\lambda)\,dx, \quad j=1,\dots,m,$$

and correspondingly

$$E\left(\frac{\partial\overline{Z_n}}{\partial\lambda_j}\right) = E\left(\frac{\partial F(\max(X_{2j-1},X_{2j});\lambda)}{\partial\lambda_j}\right) = \frac{1}{2}E\frac{\partial\overline{Y_n}}{\partial\lambda_j}, \quad j=1,\dots,m,$$

since the pdf of $X_i^* = \max(X_{2i-1},X_{2i})$ is $2F(x^*;\lambda)f(x^*;\lambda)$, $i=1,\dots,n$.

It follows that

$$B = \lim_{n\to\infty} E\left(\frac{\partial\overline{W_n}}{\partial\lambda}\right) = \binom{4}{2}\left(E\left(F\frac{\partial F}{\partial\lambda_1}\right)\dots\ E\left(F\frac{\partial F}{\partial\lambda_m}\right)\right) = 2\binom{2}{1}d',$$

and hence that

$$B(2I)^{-1}B' = 2d'I^{-1}d\begin{pmatrix}4 & 2\\ 2 & 1\end{pmatrix} = \frac{b}{90}\begin{pmatrix}4 & 2\\ 2 & 1\end{pmatrix}.$$

Thus we have

$$\Sigma_1 = \Sigma - B(2I)^{-1}B' = \frac{1}{90}\begin{pmatrix}16 & 8\\ 8 & 5\end{pmatrix} - \frac{b}{90}\begin{pmatrix}4 & 2\\ 2 & 1\end{pmatrix}$$
$$= \frac{1}{90}\begin{pmatrix}4(4-b) & 2(4-b)\\ 2(4-b) & 5-b\end{pmatrix},$$

and

$$\Sigma_1^{-1} = 90\begin{pmatrix}5-b/(4(4-b)) & -1/2\\ -1/2 & 1\end{pmatrix}.$$

Therefore

$$\widehat{D}_n^{(2)} = n(\widehat{W}-\mu)'\Sigma_1^{-1}(\widehat{W}-\mu) \xrightarrow{D} \chi^2(2),$$

which (in extended form) proves (2.9).

Finally, writing $a = \binom{-1/2}{1}$ we see that

$$\sqrt{n}\left(\widehat{Z}_n - \frac{1}{2}\widehat{Y}_n - \frac{1}{3}\right) = a'(\sqrt{n}\,(\widehat{W}-\mu)) \xrightarrow{D} a'W \sim N(0, a'\Sigma_1 a),$$

and $a'\Sigma_1 a = 1/90$, which shows (2.8) and completes the proof of Theorem 6.

PROPOSITION 3. *Goodness-of-fit tests for $X \sim \mathrm{Exp}(\alpha)$ are given by*

$$\begin{aligned}\widehat{D}_n^{(1)} &:= D_n^{(1)}(\widehat{\alpha}_{2n}) \\ &= 45 \cdot 2n\bigg(\frac{1}{2n}\sum_{j=1}^{2n}(1-e^{-\widehat{\alpha}_{2n}X_j})^2 + \frac{1}{n}\sum_{j=1}^{n} e^{-\widehat{\alpha}_{2n}X_j^*} - \frac{2}{3}\bigg)^2 \\ &\xrightarrow{D} \chi^2(1), \\ \widehat{D}_n^{(2)} &:= D_n^{(2)}(\widehat{\alpha}_{2n}) \\ &= \frac{45\cdot 36}{19} 2n\bigg(\frac{1}{2n}\sum_{j=1}^{2n}(1-e^{-\widehat{\alpha}_{2n}X_j})^2 - \frac{1}{3}\bigg)^2 + \widehat{D}_n^{(1)} \xrightarrow{D} \chi^2(2),\end{aligned}$$

where $\widehat{\alpha}_{2n} = 1/\overline{X_{2n}}$.

Proof. The first statement of Proposition 3 follows from (c) after Lemma 2 and Theorem 6. To prove the second statement it is enough to see that for $X \sim \mathrm{Exp}(\alpha)$ we have $I(\alpha) = 1/\alpha^2$,

$$d = \alpha \int_0^\infty (xe^{-2\alpha x} - xe^{-3\alpha x})\,dx = 5/(36\alpha),$$

and $b = 125/36$, which by (2.9) gives the test statistic $\widehat{D}_n^{(2)}$.

PROPOSITION 4. *Goodness-of-fit tests for $X \sim N(\mu,\sigma^2)$ with*

$$\begin{aligned}F(x) &= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x} e^{-(t-\mu)^2/(2\sigma^2)}\,dt, \\ f(x) &= \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)}, \qquad -\infty < x < \infty,\end{aligned}$$

are given by

$$\begin{aligned}\widehat{D}_n^{(1)} &:= D_n^{(1)}(\widehat{\mu}_{2n}, \widehat{\sigma}_{2n}^2) \\ &= 45 \cdot 2n\big(\overline{\Phi^2((X_{2n}-\widehat{\mu}_{2n})/\widehat{\sigma}_{2n})} - \overline{\Phi((X_n^*-\widehat{\mu}_{2n})/\widehat{\sigma}_{2n})} + \tfrac{1}{3}\big)^2 \\ &\xrightarrow{D} \chi^2(1), \\ \widehat{D}_n^{(2)} &:= D_n^{(2)}(\widehat{\mu}_{2n}, \widehat{\sigma}_{2n}^2) \\ &= \frac{45\cdot 8\pi^2}{32\pi^2 - 15(6\pi+1)} \cdot 2n\big(\overline{\Phi^2((X_{2n}-\widehat{\mu}_{2n})/\widehat{\sigma}_{2n})} - \tfrac{1}{3}\big)^2 + \widehat{D}_n^{(1)},\end{aligned}$$

where

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}\,dt,$$

$$\overline{\Phi^2((X_{2n}-\widehat{\mu}_{2n})/\widehat{\sigma}_{2n})} = \frac{1}{2n}\sum_{j=1}^{2n}\Phi^2((X_j-\widehat{\mu}_{2n})/\widehat{\sigma}_{2n}),$$

$$\overline{\Phi((X_n^*-\widehat{\mu}_{2n})/\widehat{\sigma}_{2n})} = \frac{1}{n}\sum_{j=1}^{n}\Phi((X_j^*-\widehat{\mu}_n)/\widehat{\sigma}_n),$$

and

$$\widehat{\mu}_{2n} = \overline{X_{2n}}, \quad \widehat{\sigma}^2_{2n} = \frac{1}{2n}\sum_{j=1}^{2n}(X_j-\overline{X_{2n}})^2.$$

Proof. Here

$$I^{-1} = \begin{pmatrix} \sigma^2 & 0 \\ 0 & 2\sigma^4 \end{pmatrix}$$

and

$$\frac{\partial F}{\sigma\mu} = -f, \quad \frac{\partial F}{\partial\sigma^2} = -\frac{1}{2\sigma^2}(x-\mu)f,$$

so

$$d_1 = -\int_{-\infty}^{\infty} F(x)f^2(x)\,dx, \quad d_2 = -\frac{1}{2\sigma^2}\int_{-\infty}^{\infty}(x-\mu)F(x)f^2(x)\,dx.$$

To evaluate the integrals, write

$$F(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x_1} e^{-y^2/2}\,dy = \frac{1}{2}+\psi(x_1),$$

where

$$x_1 = (x-\mu)/\sigma, \quad \psi(x_1) = \frac{1}{\sqrt{2\pi}}\int_0^{x_1} e^{-y^2/2}\,dy.$$

Changing variables to $x_1 = (x-\mu)/\sigma$ gives

$$d_1 = -\frac{1}{2\pi\sigma}\int_{-\infty}^{\infty}\left(\frac{1}{2}+\psi(x_1)\right)e^{-x_1^2}\,dx_1 = -\frac{1}{4\pi\sigma}\int_{-\infty}^{\infty} e^{-x_1^2}\,dx_1 = -\frac{1}{4\sqrt{\pi}\,\sigma},$$

where we have used the fact that ψ is an odd function. Similarly

$$d_2 = -\frac{1}{4\pi\sigma^2}\int_{-\infty}^{\infty}\left(\frac{1}{2}+\psi(x_1)\right)x_1e^{-x_1^2}\,dx_1 = -\frac{1}{4\pi\sigma^2}\int_{-\infty}^{\infty}\psi(x_1)x_1e^{-x_1^2}\,dx_1$$
$$= -\frac{1}{8\pi\sigma^2}\int_{-\infty}^{\infty}\psi'(x_1)e^{-x_1^2}\,dx_1 = -\frac{1}{8\sqrt{3}\,\pi\sigma^2},$$

where we have used integration by parts and the facts that $x_1e^{-x_1^2}$ is an odd function and

$$\psi'(x_1) = \frac{1}{\sqrt{2\pi}}e^{-x_1^2/2}.$$

Hence

$$b = 180(d_1, d_2)\begin{pmatrix} \sigma^2 & 0 \\ 0 & 2\sigma^4 \end{pmatrix}(d_1, d_2)' = \frac{15(6\pi + 1)}{8\pi^2},$$

which by (2.9) leads us to the $\widehat{D}_n^{(2)}$ test. Then $\widehat{D}_n^{(1)}$ is obtained immediately from (2.8).

3. Goodness-of-fit tests based on characterizations via moments of record values. Suppose that X has df F and pdf f. To simplify the notation we write

$$g(x) = 1 - F(x) \quad \text{and} \quad h(x) = -\log(g(x))$$

if $F(x) < 1$ and 0 otherwise.

Then Theorem 3 says (see (1.5)) that $X \sim F$ iff

$$k^{2l}(n-l)!Eh^{2l}(Y_{n-l+1}^{(k)}) - 2n!k^l Eh^l(Y_{n+1}^{(k)}) + (n+l)! = 0.$$

Since the definition of $Y_n^{(k)}$ requires an infinite sequence it is hard to see how a finite sample can be used to estimate $EY_n^{(k)}$. So our procedure is as follows.

We consider the special case $l = n$. Then $X \sim F$ iff

$$Eh^{2n}(X_{1:k}) - \frac{2n!}{k^n} Eh^n(Y_{n+1}^{(k)}) + \frac{(2n)!}{k^{2n}} = 0. \tag{3.1}$$

We know that the pdf of $Y_n^{(k)}$ is

$$f_{Y_n^{(k)}}(x) = \frac{k^n}{(n-1)!} h^{n-1}(x) g^{k-1}(x) f(x) \quad \text{(cf. [1])}$$

and that

$$\begin{aligned} F_{Y_{n+1}^{(k)}}(x) &= F_{Y_n^{(k)}}(x) - \frac{k^n}{n!} h^n(x) g^k(x) \\ &= 1 - g^k(x) \sum_{j=0}^{n} \frac{k^j}{j!} h^j(x) \quad \text{(cf. [2])}. \end{aligned} \tag{3.2}$$

Hence

$$\begin{aligned} Eh^n(Y_{n+1}^{(k)}) = Eh^n(Y_n^{(k)}) &- \frac{k^n}{(n-1)!} Eh^{2n-1}(X) g^{k-1}(X) \\ &+ \frac{k^{n+1}}{n!} Eh^{2n}(X) g^{k-1}(X). \end{aligned}$$

Taking into account that

$$Eg^{\alpha-1}(X) h^{\beta-1}(X) = \frac{\Gamma(\beta)}{\alpha^\beta} \quad \text{for } \alpha, \beta > 0$$

as X has df F, we get

$$Eh^n(Y_{n+1}^{(k)}) = Eh^n(Y_n^{(k)}) + \frac{(2n)!}{2n!k^n}.$$

Hence by (3.1) we obtain

$$Eh^{2n}(X_{1:k}) - \frac{2n!}{k^n} Eh^n(Y_n^{(k)}) = 0.$$

Letting $n = 1$ we have

$$Eh^2(X_{1:k}) - \frac{2}{k} Eh(X_{1:k}) = 0. \tag{3.3}$$

Similarly using the second equality in (3.2) we get

$$Eh^{2n}(X_{1:k}) - \frac{2n!}{k^n} Eh^n(X_{1:k}) - \frac{(2n)! - 2(n!)^2}{k^{2n}} = 0. \tag{3.4}$$

To verify $H: \ X \sim F$ we use (3.3). Consider first the case $k = 1$. Then

$$E(h^2(X_1) - 2h(X_1)) = 0.$$

The sample $(X_1, \ldots, X_n)$ provides an estimator of EW_1, where $W_1 = h^2(X_1) - 2h(X_1)$, of the form

$$\overline{W_n} = \overline{h^2(X_n)} - 2\overline{h(X_n)},$$

where

$$\overline{h^2(X_n)} = \frac{1}{n} \sum_{j=1}^n h^2(X_j), \quad \overline{h(X_n)} = \frac{1}{n} \sum_{j=1}^n h(X_j).$$

It follows from the CLT that

$$\sqrt{n}\ \overline{W_n} \xrightarrow{D} N(0, \mathrm{Var}(W_1)),$$

and hence that

$$T_n^{(1)} := n\overline{W_n}^2 / \mathrm{Var}(W_1) \xrightarrow{D} \chi^2(1),$$

and so provides a simple asymptotic test of the hypothesis $X \sim F$ when the parameters of F are known. Here

$$\mathrm{Var}\, W_1 = Eh^4(X_1) - 4Eh^3(X_1) + 4Eh^2(X_1) = 8$$

since $h(X_1) \sim \mathrm{Exp}(1)$ gives $Eh^m(X_1) = m!$, $m = 1, 2, \ldots$, and so

$$T_n^{(1)} = \frac{n}{8}(\overline{h^2(X_n)} - 2\overline{h(X_n)})^2.$$

We have proved

PROPOSITION 5. *If* $X_n \sim F$, $n \geq 1$, *are independent then*

$$T_n^{(1)} = \frac{n}{8}(\overline{h^2(X_n)} - 2\overline{h(X_n)})^2 \xrightarrow{D} \chi^2(1). \tag{3.5}$$

Now consider the case $k = 2$. Write $U_1 := X_{1:2} = \min(X_1, X_2)$. Here from (3.3) we have to estimate EW_1', where $W_1' = h^2(U_1) - h(U_1)$. The sample $X_1, \ldots, X_{2n}$ provides the sample $W_1', \ldots, W_n'$, where $W_j' = h^2(U_j) - h(U_j)$ and $U_j = \min(X_{2j-1}, X_{2j})$, $j = 1, \ldots, n$. Then EW_1' is estimated by

$$\overline{W_n'} = \overline{h^2(U_n)} - \overline{h(U_n)},$$

and

$$T_n^{(2)} := n(\overline{W_n'})^2/\mathrm{Var}(W_1') \xrightarrow{D} \chi^2(1).$$

Taking into account that $h(U_1) \sim \mathrm{Exp}(2)$ we see that $\mathrm{Var}(W_1') = 1/2$. Thus another simple asymptotic test is provided by

PROPOSITION 6. *If* $X_n \sim F$, $n \geq 1$, *are independent then*

$$T_n^{(2)} = 2n(\overline{h^2(U_n)} - \overline{h(U_n)})^2 \xrightarrow{D} \chi^2(1). \tag{3.6}$$

The same argument leads to a similar test for the case $k = 3, \ldots, n-1$ based on a sample of size kn.

We now consider the case $k = n$. Write $U_n = \min(X_1, \ldots, X_n)$. Then by (3.3) we have to estimate $E(h^2(U_n) - (2/n)h(U_n))$. The obvious estimate is $h^2(U_n) - (2/n)h(U_n)$ itself, and if the parameters of F are specified the test statistic is

$$T_n^{(n)} := \left(h^2(U_n) - \frac{2}{n}h(U_n)\right)^2.$$

As above, under H, $h(U_n) \sim \mathrm{Exp}(n)$, whence

$$R_n := nh(U_n) \sim U \sim \mathrm{Exp}(1), \quad n \geq 1. \tag{3.7}$$

It follows that

$$T_n^{(n)} = \frac{1}{n^4}(R_n^2 - 2R_n)^2$$

and so an equivalent test statistic is $T_n := (R_n^2 - 2R_n)^2 \sim T := (U^2 - 2U)^2$, $n \geq 1$, which provides an exact test for $H: \ X \sim F$.

PROPOSITION 7 (cf. [7]). *The significance probability of the test using* T_n *is*

$$\begin{aligned} P_t &:= P[T_n > t] \\ &= \begin{cases} e^{-1-\sqrt{1+\sqrt{t}}} + e^{-1+\sqrt{1-\sqrt{t}}} - e^{-1-\sqrt{1-\sqrt{t}}} & \text{if } 0 < t \leq 1, \\ e^{-1-\sqrt{1+\sqrt{t}}} & \text{if } t \geq 1. \end{cases} \end{aligned} \tag{3.8}$$

Proof. The significance probability $P[T_n > t]$ associated with an observed value t can be obtained by considering the graph of $u^2(u-2)^2 = t$ and using the fact that $P[U < u] = 1 - e^{-u}$. One finds readily that (3.8) holds true.

In particular we consider the 5% test of H, i.e. $P_t = 0.05$. But since

$$P[T > 1] = e^{-(1+\sqrt{2})} > 0.05$$

the 5% test rejects when $R_n > u_0$, where $e^{-u_0} = 0.05$, i.e. when $u_0 = 3.00$. Thus the exact 5% test rejects when $nh(U_n) > 3$.

Now we show that instead of $T_n = [R_n^2 - 2R_n]^2$ one can use more generally the statistics

$$T_n^{[m]} := \{(R_n^m - m!)^2 - ((2m)! - (m!)^2)\}^2, \quad m \geq 1.$$

We note that $T_n = T_n^{[1]}$.

Writing (3.4) in the form

$$Ek^{2m}h^{2m}(X_{1:k}) - 2m!Ek^m h^m(X_{1:k}) - ((2m)! - 2(m!)^2) = 0$$

and letting $k = n$ (sample size), we have

$$E\{((nh(X_{1:n}))^m - m!)^2 - ((2m)! - (m!)^2)\} = 0.$$

Taking into account that $R_n = nh(X_{1:n}) \sim \mathrm{Exp}(1)$, $n \geq 1$, we see that

$$T_n^{[m]} = \{(R_n^m - m!)^2 - a_m\}^2 \sim [(U^m - m!)^2 - a_m]^2$$

where

$$a_m = (2m)! - (m!)^2.$$

It follows that the statistics $T_n^{[m]}$ have for every $n \geq 1$ the distribution of $[(U^m - m!)^2 - a_m]^2$, and we reject $H : X \sim F$ if $T_n^{[m]}$ is large enough. Moreover, we can state the following result.

PROPOSITION 8. *The significance probability of the test using* $T_n^{[m]}$ *is*

$$(3.9) \qquad P_t^{[m]} := P[T_n^{[m]} > t] = \begin{cases} 1 - e^{-b_m^{(2)}(t)} + e^{-b_m^{(3)}(t)} & \text{if } 0 < t \leq t_m, \\ e^{-b_m^{(1)}(t)} - e^{-b_m^{(2)}(t)} + e^{-b_m^{(3)}(t)} & \text{if } t_m < t \leq t'_m, \\ e^{-b_m^{(3)}(t)} & \text{if } t > t'_m, \end{cases}$$

where

$$b_m^{(1)}(t) = (m! - \sqrt{a_m - \sqrt{t}}\,)^{1/m}, \qquad b_m^{(2)}(t) = (m! + \sqrt{a_m - \sqrt{t}}\,)^{1/m},$$

$$b_m^{(3)}(t) = (m! + \sqrt{a_m + \sqrt{t}}\,)^{1/m}, \qquad t_m = (a_m - (m!)^2)^2, \quad t'_m = a_m^2.$$

The proof of (3.9) is similar to the proof of Proposition 7.

COROLLARY. *$P_t^{[1]}$ is given by* (3.8) *and $P_t^{[2]}$ is given by the formula*

$$P_t^{[2]} = P[T_n^{[2]} > t]$$
$$= \begin{cases} 1 - e^{-\sqrt{2+\sqrt{20-\sqrt{t}}}} + e^{-\sqrt{2+\sqrt{20+\sqrt{t}}}} & \text{if } 0 < t \le 256, \\ e^{-\sqrt{2-\sqrt{20-\sqrt{t}}}} - e^{-\sqrt{2+\sqrt{20-\sqrt{t}}}} + e^{-\sqrt{2+\sqrt{20+\sqrt{t}}}} & \text{if } 256 < t \le 400, \\ e^{-\sqrt{2+\sqrt{20+\sqrt{t}}}} & \text{if } t > 400. \end{cases}$$

4. Tests for exponentiality. We consider corresponding tests for $X \sim \operatorname{Exp}(\alpha)$ when α is not specified. Note that in this case $h(x) = -\log(1 - F(x)) = \alpha x$. Using $T_n^{(1)} := T_n^{(1)}(\alpha)$, $T_n^{(2)} := T_n^{(2)}(\alpha)$ in (3.5) and (3.6) respectively, we replace α by the estimator $\widehat{\alpha}_n$. We have proved in [7] the following results.

PROPOSITION 9. *If $X_n \sim F$, $n \ge 1$, are independent then*

$$\widehat{T}_n^{(1)} := 2T_n^{(1)}(\widehat{\alpha}_n) = \frac{n}{4}(\overline{X_n^2}/(\overline{X_n})^2 - 2)^2 \xrightarrow{D} \chi^2(1),$$

where $\widehat{\alpha}_n = 1/\overline{X_n}$.

PROPOSITION 10. *If $X_n \sim F$, $n \ge 1$, are independent then*

$$\widehat{T}_n^{(2)} := \frac{4}{3}T_n^{(2)}(\widehat{\alpha}_n) = \frac{8}{3}\left(\overline{U_n^2} - \frac{1}{\widehat{\alpha}_n}\overline{U_n}\right)^2 = \frac{8n}{3}\left(\frac{\overline{U_n^2}}{(\overline{X_{2n}})^2} - \frac{\overline{U_n}}{\overline{X_{2n}}}\right)^2 \xrightarrow{D} \chi^2(1),$$

where $\widehat{\alpha}_n = 1/\overline{X_{2n}}$.

PROPOSITION 11. *Let $\widehat{T}_n := T(\widehat{\alpha}_n) = (\widehat{U}_n^2 - 2\widehat{U}_n)^2$ where $\widehat{U}_n = n\widehat{\alpha}_n U_n = nU_n/\overline{X_n}$, and let $\widehat{P}_t := P[\widehat{T}_n > t]$ stand for the associated significance probability. Then $\lim_{n\to\infty} \widehat{P}_t = P_t$, where P_t is given by Proposition* 7.

Now by Proposition 8 we have the following generalization of Proposition 11.

PROPOSITION 12. *Let*

$$\widehat{T}_n^{[m]} := T_n^{[m]}(\widehat{\alpha}_n) = \{[(n\widehat{\alpha}_n U_n)^m - m!]^2 - a_m\}^2$$

and let $\widehat{P}_t^{[m]} := P[\widehat{T}_n^{[m]} > t]$ stand for the associated significance probability. Then

$$\lim_{n\to\infty} \widehat{P}_t^{[m]} = P_t^{[m]}, \quad m \ge 1,$$

where $P_t^{[m]}$ is given by Proposition 8.

Proof. Since $\widehat{\alpha}_n \xrightarrow{P} \alpha$, from (3.7) we get $n\widehat{\alpha}_n U_n = (\widehat{\alpha}_n/\alpha)R_n \xrightarrow{D} U$ and so

$$\widehat{T}_n^{[m]} \xrightarrow{D} [(U^m - m!)^2 - a_m]^2,$$

which is distributed as $T_n^{[m]}$.

Acknowledgements. The authors are very grateful to the referee for a number of useful suggestions.

References

[1] W. Dziubdziela and B. Kopociński, *Limiting properties of the k-th record values*, Zastos. Mat. 15 (1976), 187–190.

[2] Z. Grudzień and D. Szynal, *On the expected values of k-th record values and associated characterizations of distributions*, in: Proc. 4th Pannonian Symp. on Math. Statist. (Bad Tatzmannsdorf, 1985), 119–127.

[3] —, —, *Characterization of continuous distributions in terms of moments of extremal statistics*, J. Math. Sci. 81 (1996), 2912–2936.

[4] —, —, *Characterizations of continuous distributions via moments of record values*, J. Appl. Statist. Soc. 9 (2000), 93–103.

[5] G. D. Lin, *Characterizations of continuous distributions via expected values of two functions of order statistics*, Sankhyā Ser. A 52 (1990), 84–90.

[6] K. Morris and D. Szynal, *A goodness-of-fit test for the uniform distribution based on a characterization*, J. Math. Sci., submitted.

[7] —, —, *Goodness-of-fit tests based on characterizations of continuous distributions*, Appl. Math. (Warsaw) 27 (2000), 475–488.

[8] D. A. Pierce, *The asymptotic effect of substituting estimators for parameters in certain types of statistics*, Ann. Statist. 10 (1982), 475–478.

[9] R. H. Randles, *On the asymptotic normality of statistics with estimated parameters*, ibid., 462–474.

[10] Y. H. Too and G. D. Lin, *Characterizations of uniform and exponential distributions*, Statist. Probab. Lett. 7 (1989), 357–359.

Department of Applied Mathematics
University of Adelaide
North Tce, Adelaide
South Australia, 5001
E-mail: kmorris@stats.adelaide.edu.au

Institute of Mathematics
Maria Curie-Skłodowska University
Pl. M. Curie-Skłodowskiej 1
20-031 Lublin, Poland
E-mail: szynal@golem.umcs.lublin.pl

Received on 1.9.2000;
revised version on 26.1.2001 (1549)