ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
21,4 (1993), pp. 511–519

F. CZEKAŁA (Wrocław)

THE ASYMPTOTIC DISTRIBUTIONS OF STATISTICS BASED ON LOGARITHMS OF SPACINGS

1. Introduction and preliminaries. Let $X_1, \ldots, X_n$ be a sample from a distribution on the interval $[0,1]$. We denote by $0 = X_{0:n} \le X_{1:n} \le \ldots \le X_{n:n} \le X_{n+1:n} = 1$ the order statistics from the sample. The purpose of this paper is to find the asymptotic distributions of the statistics $\sum_{i=0}^{n} \log(X_{i+1:n} - X_{i:n})$ and $\sum_{i=0}^{n}(X_{i+1:n} - X_{i:n}) \log(X_{i+1:n} - X_{i:n})$ in the case when the distribution of the sample has a density which is a step function on $[0,1]$. The asymptotic distributions of the above statistics in the case of the uniform distribution on $[0,1]$ were found by Darling [1] and Gebert and Kale [2], respectively. The problem of finding asymptotic distributions of some other statistics based on distributions with densities in the form of step functions was considered by Weiss [3], [4].

Let $k \ge 1$ be a fixed integer and let $0 = x_0 < x_1 < \ldots < x_k = 1$ be fixed real numbers. Let us introduce the following notation:

$$I_0 = [0,1], \quad I_1 = [x_0, x_1), \quad \ldots, \quad I_{k-1} = [x_{k-2}, x_{k-1}), \quad I_k = [x_{k-1}, x_k],$$
$$|I_i| = x_i - x_{i-1}, \quad i = 1, \ldots, k.$$

Let $f_i > 0$, $i = 1, \ldots, k$, be fixed numbers such that $\sum_{i=1}^{k} f_i |I_i| = 1$. These numbers together with the intervals I_i define a probability density f:

$$f(x) = \sum_{i=1}^{k} f_i \mathbf{1}_{I_i}(x). \tag{1}$$

Define $p_i = f_i |I_i|$. We have $\sum p_i = 1$ and hence there exist numbers $0 = x'_0 < x'_1 < \ldots < x'_k = 1$ such that the intervals I'_i, defined similarly to I_i, have lengths p_i. It is well known that there exists a probability space $(\Omega, \mathcal{F}, P)$, on which we can define a vector of random elements

1991 *Mathematics Subject Classification*: 62E20, 62G20, 62G30.

Key words and phrases: spacings, convergence in distribution, consistency of tests.

$$(\underline{U}, \underline{Y}_1, \ldots, \underline{Y}_k) : \Omega \to (\mathbb{R}^\infty)^{k+1} \tag{2}$$

with the following properties:

(a) the coordinates of this vector are stochastically independent,

(b) $\underline{U} = (U_1, U_2, \ldots)$ is a sequence of independent random variables uniformly distributed on I_0,

(c) for $i = 1, \ldots, k$, $\underline{Y}_i = (Y_{i,1}, Y_{i,2}, \ldots)$ are sequences of independent random variables with uniform distribution on I_i.

We can now define a sequence $\underline{Z} = (Z_1, Z_2, \ldots)$ of independent random variables with density f:

$$Z_n = \sum_{i=1}^{k} \mathbf{1}_{I_i'}(U_n) Y_{i,n}, \quad n \geq 1. \tag{3}$$

The independence of the variables Z_n results from the properties of the vector (2). We check that they have density f. Let $A \subset I_0$ be a Borel set and let λ be the Lebesgue measure. Then we have

$$\begin{aligned} P(Z_n \in A) &= \sum_{i=1}^{k} P(Z_n \in A \cap I_i) = \sum_{i=1}^{k} P(U_n \in I_i') P(Y_{i,n} \in A \cap I_i) \\ &= \sum_{i=1}^{k} f_i |I_i| \frac{\lambda(A \cap I_i)}{|I_i|} = \int_A f(x)\, d\lambda(x). \end{aligned}$$

Now denote by $N_{i,n}$ the number of random variables $Z_1, \ldots, Z_n$ taking values in the interval I_i, that is,

$$N_{i,n} = \sum_{j=1}^{n} \mathbf{1}_{I_i}(Z_j). \tag{4}$$

Since $Z_j \in I_i$ if and only if $U_j \in I_i'$, we have

$$N_{i,n} = \sum_{j=1}^{n} \mathbf{1}_{I_i'}(U_j). \tag{5}$$

It follows that the sequence $(N_{1,n}, \ldots, N_{k,n})$ is independent of $(\underline{Y}_1, \ldots, \underline{Y}_k)$. For simplicity of notation we now introduce two real functions defined on $\mathbb{N} \times I_0^\infty$ as follows:

$$\phi(n, \underline{y}) = \sum_{j=0}^{n} \log(y_{j+1:n} - y_{j:n}), \tag{6}$$

$$\phi^*(n, \underline{y}) = \sum_{j=0}^{n} (y_{j+1:n} - y_{j:n}) \log(y_{j+1:n} - y_{j:n}), \tag{7}$$

where $n \geq 1$, $\underline{y} = (y_1, y_2, \ldots) \in I_0^\infty$. The Darling and the Gebert–Kale statistics based on a sample from the uniform distribution on I_0 can now be written as $\phi(n, \underline{U})$ and $\phi^*(n, \underline{U})$, where $\underline{U}$ is the sequence (2).

2. Darling statistic. The asymptotic distribution of the Darling statistic in the case of the uniform distribution on I_0 is given by the following theorem:

LEMMA 1 (Darling [2]). *If $\underline{U} = (U_1, U_2, \ldots)$ is a sequence of independent random variables with uniform distribution on* $[0,1]$, *then*

$$\frac{\phi(n, \underline{U}) + n(\log n + \gamma)}{\sqrt{\pi^2/6 - 1}\sqrt{n}} \xrightarrow{d} N(0,1), \tag{8}$$

where $\gamma = 0.5772\ldots$ is Euler's constant.

Using Lemma 1 we find the limiting distribution of the statistic ϕ when the distribution of the sample variables is given by (1):

THEOREM 1. *If $\underline{Z} = (Z_1, Z_2, \ldots)$ is a sequence of independent random variables distributed according to* (1) *then*

$$\frac{\phi(n, \underline{Z}) + n\{E[\log(nf(Z_1))] + \gamma\}}{\sqrt{\pi^2/6 - 1 + \mathrm{Var}[\log f(Z_1)]}\sqrt{n}} \xrightarrow{d} N(0,1). \tag{9}$$

To prove Theorem 1 we need some lemmas.

LEMMA 2. *Let X, X_n, $n \geq 1$, be random variables such that $X_n \xrightarrow{P.1} 0$, $\sqrt{n} X_n \xrightarrow{d} X$. Then*

$$\sqrt{n}\,|\log(1 + X_n) - X_n| \xrightarrow{P} 0. \tag{10}$$

Proof. Using the Taylor expansion it is easy to show that $|\log(1+x) - x| \leq x^2$ for $|x| \leq 1/2$. Consider the set $B_n = \bigcap_{j=n}^{\infty}\{|X_j| \leq 1/2\}$. Since $X_n \xrightarrow{P.1} 0$, $\lim P(B_n) = 1$. The assumption that $X_n \xrightarrow{P.1} 0$, together with $\sqrt{n} X_n \xrightarrow{d} X$, gives $\sqrt{n} X_n^2 \xrightarrow{P} 0$. We conclude that for arbitrary $\varepsilon, \eta > 0$ there exists j_0 such that $P(\Omega \setminus B_{j_0}) < \eta$ and $P(\sqrt{n} X_n^2 \geq \varepsilon) \leq \eta$ for $n \geq j_0$. Hence

$$\begin{aligned} P(\sqrt{n}\,|\log(1 + X_n) - X_n| \geq \varepsilon) &\\ &\leq P(\Omega \setminus B_{j_0}) + P(B_{j_0}, \sqrt{n}\,|\log(1 + X_n) - X_n| \geq \varepsilon) \\ &\leq \eta + P(B_{j_0}, \sqrt{n} X_n^2 \geq \varepsilon) \leq 2\eta \quad \text{for } n \geq j_0. \end{aligned}$$

This proves (10).

LEMMA 3. *For arbitrary $i = 1, \ldots, k$ and $\varepsilon > 0$ we have*

$$\frac{1}{\sqrt{n}}\left|\log\left(\min_{1 \leq j \leq N_{i,n}} Y_{i,j} - x_{i-1}\right)\right| \xrightarrow{P} 0, \tag{11}$$

$$(12)\qquad \frac{1}{\sqrt{n}}|\log(x_1 - \max_{1\le j\le N_{i,n}} Y_{i,j})| \xrightarrow{P} 0\,.$$

Proof. The random variable $N_{i,n}$ is independent of the sequence $\underline{Y}_i$, and $N_{i,n} \xrightarrow{P.1} \infty$, so it suffices to show that

$$\frac{1}{\sqrt{n}}|\log(\min_{j\le j\le n} Y_{i,j} - x_{i-1})| \xrightarrow{P} 0\,,$$

and similarly in the case of (12). We have

$$\begin{aligned} P\bigg(\frac{1}{\sqrt{n}}|\log(\min_{1\le j\le n} Y_{i,j} - x_{i-1})| \ge \varepsilon\bigg) &= P(\min_{1\le j\le n} Y_{i,j} - x_{i,j} \le \exp(-\varepsilon\sqrt{n}))\\ &= 1-(1-|I_i|^{-1}\exp(-\varepsilon\sqrt{n}))^n\,. \end{aligned}$$

We now prove that $(1-|I_i|^{-1}\exp(-\varepsilon\sqrt{n}))^n \to 1$. Take arbitrary $\eta > 0$ and $\delta > 0$ such that $\exp(-\delta) \ge 1-\eta$. For sufficiently large n,

$$1 \ge (1-|I_i|^{-1}\exp(-\varepsilon\sqrt{n}))^n \ge (1-\delta/n)^n \ge \exp(-\delta) - \eta \ge 1-2\eta\,.$$

This ends the proof of Lemma 3.

LEMMA 4. *Let $X_{i,n}$, $i = 1,\ldots,k$, $n \ge 0$, be the random variables defined as follows*:

$$X_{i,n} = \begin{cases} 0 & \textit{for } n = 0\,,\\ \dfrac{1}{\sqrt{\pi^2/6-1}\sqrt{n}}\bigg[\phi(n,\underline{Y}_i) + n\bigg(\log\dfrac{n}{|I_i|} + \gamma\bigg)\bigg] & \textit{for } n \ge 1\,. \end{cases}$$

Then

$$(13)\qquad \bigg(X_{i,N_{i,n}}, i = 1,\ldots,k, \frac{N_{i,n} - np_i}{\sqrt{n}}, i = 1,\ldots,k\bigg) \xrightarrow{d} (X_1,\ldots,X_k,W_1,\ldots,W_k)\,,$$

where the X_i are independent and normally $N(0,1)$ distributed random variables, the vector $(W_1,\ldots,W_k)$ is independent of $(X_1,\ldots,X_k)$ and has the multivariable normal distribution $N(0,\Sigma)$, where $\Sigma = [\sigma_{i,j}]$ and

$$\sigma_{i,j} = \begin{cases} -p_ip_j & \textit{for } i \ne j\,,\\ p_i - p_i^2 & \textit{for } i = j\,, \end{cases}$$

$i, j = 1,\ldots,k$.

Proof. We first show that $(X_{1,n},\ldots,X_{k,n}) \xrightarrow{d} (X_1,\ldots,X_k)$. It is obvious that spacings from the uniform distribution on a certain interval divided by the length of this interval have the same distribution as spacings from the uniform distribution on $[0,1]$. Thus

$$\frac{1}{\sqrt{\pi^2/6-1}\sqrt{n}}[\phi(n,\underline{Y}_i/|I_i|) + n(\log n + \gamma)] \xrightarrow{d} X_i\,.$$

Moreover,

$$\phi(n, \underline{Y}_i/|I_i|) + n(\log n + \gamma) = \phi(n, \underline{Y}_i) + n\log(n/|I_i|) + \gamma) - \log|I_i|.$$

Since $\log(|I_i|/\sqrt{n}) \to 0$, we have $X_{i,n} \xrightarrow{d} X_i$. Because X_i are independent, we have also

$$(14) \qquad (X_{1,n}, \ldots, X_{k,n}) \xrightarrow{d} (X_1, \ldots, X_k).$$

Now we show that

$$(15) \qquad \left(\frac{N_{1,n} - np_1}{\sqrt{n}}, \ldots, \frac{N_{k,n} - np_k}{\sqrt{n}}\right) \xrightarrow{d} (W_1, \ldots, W_k).$$

This follows from the central limit theorem because $N_{i,n} = \sum_{j=1}^{n} \mathbf{1}_{I_1'}(U_j)$, $E[\mathbf{1}_{I_i'}(U_1)] = p_i$ and $\mathrm{Cov}[\mathbf{1}_{I_i'}(U_1), \mathbf{1}_{I_j'}(U_1)] = \sigma_{i,j}$. Since the sequence $N_{i,n}$ is independent of the sequence $\underline{Y}_i$, (14) and (15) together give (13).

Proof of Theorem 1. Consider the sequence $Z_{\alpha_{i,1}}, Z_{\alpha_{i,2}}, \ldots$ of those successive random variables Z_n whose values belong to the interval I_i. The sequence $\alpha_{i,n}$, $n \geq 1$, is determined by the sequence $\underline{U}$, so it is independent of $\underline{Y}_i$. Since $Z_{\alpha_{i,n}} = Y_{i,\alpha_{i,n}}$, the sequence $Z_{\alpha_{i,n}}$ has the same probability distribution as $\underline{Y}_i$. It can be shown similarly that the joint distribution of the vector $(Z_{\alpha_{i,n}}, n \geq 1, \ldots, Z_{\alpha_{k,n}}, n \geq 1)$ is the same as that of $(\underline{Y}_1, \ldots, \underline{Y}_k)$. It follows from the above remarks and from Lemma 3 that the statistic $\phi(n, \underline{Z})$ has asymptotically the same distribution as $\sum_{i=1}^{k} \phi(N_{i,n}, \underline{Y}_i)$.

To prove Theorem 1 it is now enough to show that

$$(16) \qquad \frac{\sum_{i=1}^{k} \phi(N_{i,n}, \underline{Y}_i) + n(E[\log(nf(Z_1))] + \gamma)}{\sqrt{n}\sqrt{\pi^2/6 - 1 + \mathrm{Var}[\log f(Z_1)]}} \xrightarrow{d} N(0,1).$$

Set $C = \sqrt{\pi^2/6 - 1}$. Since $\sum_{i=1}^{k} N_{i,n} = n$, after elementary calculations we obtain

$$\begin{aligned}
(17) \quad & \frac{1}{C\sqrt{n}}\Big[\sum_{i=1}^{k} \phi(N_{i,n}, \underline{Y}_i) + n(E[\log(nF(Z_1))] + \gamma)\Big] \\
&= \frac{1}{C\sqrt{n}}\left[\sum_{i=1}^{k} \phi(N_{i,n}, \underline{Y}_i) + N_{i,n}\left(\log\frac{N_{i,n}}{|I_i|} + \gamma\right)\right] \\
&\quad - \frac{1}{C\sqrt{n}}\sum_{i=1}^{k-1}(N_{i,n} - np_i)\log\frac{N_{i,n}}{N_{k,n}} + \frac{1}{C\sqrt{n}}\sum_{i=1}^{k}(N_{i,n} - np_i)\log|I_i| \\
&\quad - \frac{1}{C\sqrt{n}}\sum_{i=1}^{k} np_i \log\left(\frac{N_{i,n}}{np_i}\right).
\end{aligned}$$

We first show that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{k} np_i \log\left(\frac{N_{i,n}}{np_i}\right) \xrightarrow{P} 0\,. \tag{18}$$

Define $\alpha(x) = \log(1+x) - x$. We have

$$\begin{aligned}\sum_{i=1}^{k} p_i \log\left(\frac{N_{i,n}}{np_i}\right) &= \sum_{i=1}^{k} p_i \frac{N_{i,n}-np_i}{np_i} + \sum_{i=1}^{k} p_i \alpha\left(\frac{N_{i,n}-np_i}{np_i}\right)\\ &= \sum_{i=1}^{k} p_i \alpha\left(\frac{N_{i,n}-np_i}{np_i}\right).\end{aligned}$$

Lemma 2 shows that

$$\sqrt{n}\alpha\left(\frac{N_{i,n}-np_i}{np_i}\right) \xrightarrow{P} 0$$

and this implies (18).

It follows from (17), (18) and Lemma 4 that

$$\begin{aligned}\frac{1}{C\sqrt{n}}\Big[\sum_{i=1}^{k} \phi(N_{i,n}, \underline{Y}_i) + n(E[\log(nf(Z_1))] + \gamma)\Big] \\ \xrightarrow{d} \sum_{i=1}^{k} \sqrt{p_i} X_i + \frac{1}{C}\sum_{i=1}^{k} W_i \log|I_i| - \frac{1}{C}\sum_{i=1}^{k-1} W_i \log\frac{p_i}{p_k}\,,\end{aligned} \tag{19}$$

where X_i, W_i are the same as in Lemma 4. Since $\sum_{i=1}^{k} p_i = 1$ and $\sum_{i=1}^{k} W_i = 0$, we have

$$\sum_{i=1}^{k} W_i \log|I_i| - \sum_{i=1}^{k-1} W_i \log\frac{p_i}{p_k} = \sum_{i=1}^{k} W_i \log\frac{|I_i|}{p_i}\,.$$

The variance of the sum of the coordinates of an arbitrary random vector is equal to the sum of the elements of the covariance matrix of this vector. Hence

$$\begin{aligned}\operatorname{Var}\left[\sum_{i=1}^{k} W_i \log\frac{|I_i|}{p_i}\right] &= \sum_{i=1}^{k}\sum_{j=1}^{k}\left(\log\frac{|I_i|}{p_i}\log\frac{|I_j|}{p_j}\right)\sigma_{i,j}\\ &= \sum_{i=1}^{k} p_i \log^2\frac{|I_i|}{p_i} - \left(\sum_{i=1}^{k} p_i \log\frac{|I_i|}{p_i}\right)^2 = \operatorname{Var}[\log f(Z_1)]\,.\end{aligned}$$

Now (19) shows that

$$\frac{\sum_{i=1}^{k} \phi(N_{i,n}, \underline{Y}_i) + n(E[\log(nf(Z_1))] + \gamma)}{\sqrt{n}C\sqrt{1 + (1/C^2)\operatorname{Var}[\log f(Z_1)]}} \xrightarrow{d} N(0,1)\,.$$

This ends the proof of Theorem 1.

3. Gebert–Kale statistic. In this section we prove an analogue of Theorem 1 concerning the Gebert–Kale statistic. Our considerations are based on the following theorem which gives the asymptotic distribution of the Gebert–Kale statistic in the case of the uniform distribution on $[0,1]$.

LEMMA 5 (Gebert and Kale [2]). *If $\underline{U} = (U_1, U_2, \ldots)$ is a sequence of independent random variables uniformly distributed on $[0,1]$ then*

$$\frac{\sqrt{n}}{\sqrt{\pi^2/3-3}}(\phi^*(n,\underline{U}) + \log n - (1-\gamma)) \xrightarrow{d} N(0,1). \tag{20}$$

The asymptotic distribution of the Gebert–Kale statistic when the underlying distribution has density (1) is given by the following theorem.

THEOREM 2. *If $\underline{Z} = (Z_1, Z_2, \ldots)$ is a sequence of independent random variables with density* (1) *then*

$$\frac{\sqrt{n}\left(\phi^*(n,\underline{Z}) + E\left[\frac{\log(nf(Z_1))}{f(Z_1)}\right] - (1-\gamma)\right)}{\sqrt{\left(\frac{\pi^2}{3}-2\right)\operatorname{Var}\left[\frac{1}{f(Z_1)}\right] + \left(\frac{\pi^2}{3}-3\right)}} \xrightarrow{d} N(0,1). \tag{21}$$

We prove Theorem 2 analogously to Theorem 1. We need some counterparts of Lemmas 3 and 4.

LEMMA 6. *For arbitrary $i = 1, \ldots, k$ we have*

$$\sqrt{n}(\min_{1\le j\le N_{i,n}} Y_{i,j} - x_{i-1}) \log(\min_{1\le j\le N_{i,n}} Y_{i,j} - x_{i-1}) \xrightarrow{P} 0, \tag{22}$$

$$\sqrt{n}(x_i - \max_{1\le j\le N_{i,n}} Y_{i,j}) \log(x_i - \max_{1\le j\le N_{i,n}} Y_{i,j}) \xrightarrow{P} 0, \tag{23}$$

Proof. It is easy to show that $n(\min_{1\le j\le n} Y_{i,j} - x_{i-1})$ converges in distribution to an exponentially distributed random variable. In the proof of Lemma 3 we have shown that $(1/\sqrt{n})\log(\min_{1\le j\le n} \underline{Y}_j - x_{i-1}) \xrightarrow{P} 0$. Hence

$$n(\min_{1\le j\le n} Y_{i,j} - x_{i-1}) \frac{1}{\sqrt{n}} \log(\min_{1\le j\le n} Y_{i,j} - x_{i-1}) \xrightarrow{P} 0. \tag{24}$$

Since $N_{i,n}$ is independent of $\underline{Y}_i$ and $N_{i,n} \xrightarrow{P.1} \infty$, (22) follows from (24). The proof of (23) is quite similar and is omitted.

LEMMA 7. *Let $X^*_{i,n}$, $i = 1, \ldots, k$, $n \ge 0$, be the random variables defined as follows*:

$$X^*_{i,n} = \begin{cases} 0 & \textit{for } n = 0, \\ \frac{\sqrt{n}}{\sqrt{\pi^2/3-3}}\left[\phi^*(n,\underline{Y}_i) + |I_i|\log\frac{n}{|I_i|} - |I_i|(1-\gamma)\right] & \textit{for } n \ge 1. \end{cases}$$

Then

$$(25)\qquad \left(X^*_{1,N_{1,n}},\ldots,X^*_{k,N_{k,n}},\frac{N_{1,p}-np_1}{\sqrt{n}},\ldots,\frac{N_{k,n}-np_k}{\sqrt{n}}\right) \xrightarrow{d} (|I_1|X_1,\ldots,|I_k|X_k,W_1,\ldots,W_k),$$

where $(X_1,\ldots,X_k,W_1,\ldots,W_k)$ *is defined by* (13) *in Lemma* 4.

Proof. Taking into account the remark at the beginning of the proof of Lemma 4, we deduce from Lemma 5 that

$$\frac{\sqrt{n}}{\sqrt{\pi^2/3-3}}\left(\phi^*\left(n,\frac{\underline{Y}_i}{|I_i|}\right)+\log n-(1-\gamma)\right)\xrightarrow{d} X_i\,.$$

After some easy calculations we get

$$\frac{\sqrt{n}}{\sqrt{\pi^2/3-3}}\left(\phi^*(n,\underline{Y}_i)+|I_i|\log\frac{n}{|I_i|}-|I_i|(1-\gamma)\right)\xrightarrow{d}|I_i|X_i\,.$$

The rest of the proof is similar to the proof of Lemma 4 and is omitted.

Proof of Theorem 2. Similarly to the proof of Theorem 1 the statistic $\phi^*(n,\underline{Z})$ can be replaced by $\sum_{i=1}^k \phi^*(N_{i,n},\underline{Y}_i)$. Set $D=1/\sqrt{\pi^2/3-3}$. We have to show that

$$(26)\qquad \frac{D\sqrt{n}\left(S^*_n+E\left[\frac{\log(nf(Z_1))}{f(Z_1)}\right]-(1-\gamma)\right)}{\sqrt{(1+D^2)\operatorname{Var}[1/f[Z_1)]+1}}\xrightarrow{d} N(0,1)\,.$$

We have the following identity:

$$\begin{aligned}D\sqrt{n}&\left[\sum_{i=1}^k\phi^*(N_{i,n},\underline{Y}_i)+\sum_{i=1}^k|I_i|\log\frac{np_i}{|I_i|}-(1-\gamma)\right]\\&=D\sqrt{n}\left[\sum_{i=1}^k\phi^*(N_{i,n},\underline{Y}_i)+\sum_{i=1}^k|I_i|\log\frac{N_{i,n}}{|I_i|}-(1-\gamma)\right]\\&\quad-D\sum_{i=1}^k\frac{|I_i|}{p_i}\frac{N_{i,n}-np_i}{\sqrt{n}}-D\sum_{i=1}^k|I_i|\sqrt{n}\alpha\left(\frac{N_{i,n}-np_i}{np_i}\right),\end{aligned}$$

where $\alpha(x)=\log(1+x)-x$. Similarly to the proof of Theorem 1 we can show that

$$\sum_{i=1}^k|I_i|\sqrt{n}\alpha\left(\frac{N_{i,n}-np_i}{np_i}\right)\xrightarrow{P}0\,.$$

Hence Lemma 7 yields that

$$D\sqrt{n}\Big(\sum_{i=1}^{k}\phi^*(N_{i,n},\underline{Y}_i)+E\Big[\frac{\log(nf(Z_1))}{f(Z_1)}\Big]-(1-\gamma)\Big)$$
$$\xrightarrow{d}\sum_{i=1}^{k}\frac{|I_i|}{\sqrt{p_i}}X_i-D\sum_{i=1}^{k}\frac{|I_i|}{p_i}W_i\,.$$

Computing the variance of the right side above, we get the conclusion of Theorem 2.

4. The consistency of tests based on the Darling and Gebert–Kale statistics. Let us now consider the problem of testing the hypothesis that a given distribution F is equal to the uniform distribution on $[0,1]$. Using Theorems 1 and 2 we can show that tests based on the statistics ϕ and ϕ^* are consistent if the alternative distributions have a density of type (1). For ϕ we have

$$\frac{1}{\sqrt{n}}[\phi(n,\underline{Z})+n(\log n+\gamma)]$$
$$=\frac{1}{\sqrt{n}}[\phi(n,\underline{Z})+n(E[\log(nf(Z_1))]+\gamma)]-\sqrt{n}E[\log(f(Z_1))]\,.$$

The desired consistency follows from the known fact that $E[\log(f(Z_1))]>0$ if the density f is not identically 1. Similarly for ϕ^* the consistency is equivalent to $E[\log(f(Z_1))/f(Z_1)]<0$ for f not identically 1.

References

[1] D. A. Darling, *On a class of problems relating to the random division of an interval*, Ann. Math. Statist. 24 (1953), 239-253.

[2] J. R. Gebert and B. K. Kale, *Goodness of fit tests based on discriminatory information*, Statist. Hefte 3 (1969), 192–200.

[3] L. Weiss, *The asymptotic power of certain tests of fit based on sample spacings*, Ann. Math. Statist. 28 (1957), 783–786.

[4] —, *The limiting joint distribution of the largest and smallest sample spacings*, ibid. 30 (1959), 590–593.

FRANCISZEK CZEKAŁA
MATHEMATICAL INSTITUTE
UNIVERSITY OF WROCŁAW
PL. GRUNWALDZKI 2/4
50-384 WROCŁAW, POLAND

Received on 18.9.1991