DISTRIBUTION RESULTS FOR DISTANCE FUNCTIONS BASED ON THE MODIFIED EMPIRICAL DISTRIBUTION FUNCTION OF M. KAC*

MIKLÓS CSÖRGŐ

(Received Aug. 4, 1969)

1. Introduction

Let $N_\lambda, X_1, X_2, \cdots$ be independent random variables, N_λ having a Poisson distribution with mean λ and each X_i having the same continuous distribution function $F(y)$. Let $\psi_y(x)$ be 0 or 1 according as $x>y$ or $x\leqq y$. The modified empirical distribution function was defined by M. Kac [10] as

$$(1.1)\qquad F_\lambda^*(y)=\lambda^{-1}\sum_{j=1}^{N_\lambda}\psi_y(X_j)\ ,\qquad -\infty<y<\infty\ ,$$

where the sum is taken to be zero if $N_\lambda=0$. Analogous one and two sided Kac statistics of the original one and two sided Kolmogorov statistics are $\text{l.u.b.}_{-\infty<y<\infty}[F(y)-F_\lambda^*(y)]$ and $\text{l.u.b.}_{-\infty<y<\infty}|F(y)-F_\lambda^*(y)|$ respectively. The exact and limiting distribution of the first one of these random variables was studied recently by J. L. Allen and J. A. Beekman in [1], and they also studied the exact distribution of the two sided Kac statistic in [2] whose asymptotic distribution was found by M. Kac in [10] to be

$$(1.2)\qquad \lim_{\lambda\to\infty} P\{\text{l.u.b.}_{-\infty<y<\infty}|F(y)-F_\lambda^*(y)|\leqq\alpha/\lambda^{1/2}\}=P^*(\alpha)\ ,$$

where

$$P^*(\alpha)=\begin{cases}\dfrac{4}{\pi}\sum_{k=0}^{\infty}(-1)^k/(2k+1)\exp[-(2k+1)^2\pi^2/8\alpha^2]\ , & \alpha>0\ ,\\ 0\ , & \alpha\leqq 0\ .\end{cases}$$

Let y_b be a real number with $F(y_b)=b$. We, in [4], derived an explicit form for

$$(1.3)\qquad P_\lambda(\varepsilon, b)=P\{\text{l.u.b.}_{-\infty<y\leqq y_b}|F(y)-F_\lambda^*(y)|\leqq\varepsilon\}$$

* Work supported in part by The Canadian Mathematical Congress Summer Research Institute at Kingston, Summer 1969.

and for

$$M_\lambda(\varepsilon, b)=P\left\{\text{l.u.b.}_{-\infty<y\leq y_b}\frac{F(y)-F_\lambda^*(y)}{1-F(y)}\leq\varepsilon\right\}. \tag{1.4}$$

In Section 2 of this paper we derive various limiting distributions for some functions of the random variable involved in the probability statement of (1.3). Distribution results are proved for two sample versions of the one and two sided Kac statistics in Section 3.

Let n be a positive integer and $Y_1<Y_2<\cdots<Y_n$ be the order statistics corresponding to $X_1, X_2,\cdots, X_n$. Define

$$F_{n,\lambda}(y)=\begin{cases} 0\,, & y<Y_1 \\ K/\lambda\,, & Y_k\leq y<Y_{k+1},\ k=1,2,\cdots,n-1 \\ n/\lambda\,, & y\geq Y_n\,. \end{cases} \tag{1.5}$$

Thus $F_{n,\lambda}(y)=(n/\lambda)F_n(y)$, where $F_n(y)$ is the ordinary empirical distribution function. $F_{n,\lambda}(y)$ as defined in (1.5) will be used in the sequel.

As long as $F(y)$ is continuous, the distribution of the Kac statistics is independent of $F(y)$ and we can therefore confine our attention to the single case $F(x)=x$, $0\leq x\leq 1$.

2. The asymptotic distribution of some Kac statistics

THEOREM 1. *For $N_\lambda, X_1, X_2,\cdots$ subject to the previous conditions, and $0<b\leq 1$, with $F(y_b)=b$*

$$\lim_{\lambda\to\infty} P\{\text{l.u.b.}_{-\infty<y<y_b}(F(y)-F^*(y))\leq\alpha/\lambda^{1/2}\} = \begin{cases} (2/\pi)^{1/2}\displaystyle\int_0^{\alpha/b^{1/2}} e^{-u^2/2}du\,, & \alpha>0 \\ 0\,, & \alpha\leq 0\,. \end{cases} \tag{2.1}$$

THEOREM 2. *For $N_\lambda, X_1, X_2,\cdots$ subject to the previous conditions, and $0\leq a<b\leq 1$, with $F(y_a)=a$ and $F(y_b)=b$*

$$\lim_{\lambda\to\infty} P\{\text{l.u.b.}_{y_a\leq y\leq y_b}(F(y)-F^*(y))\leq\alpha/\lambda^{1/2}\}=R_\alpha(a, b)\,, \tag{2.2}$$

where

$$R_\alpha(a, b)=\frac{1}{\pi}\int_{-\infty}^{\alpha/a^{1/2}} e^{-u^2/2}\left(\int_0^{(\alpha/a^{1/2}-u)(a/(b-a))^{1/2}} e^{-t^2/2}dt\right)du\,, \quad -\infty<\alpha<+\infty\,.$$

THEOREM 3. *For $N_\lambda, X_1, X_2,\cdots$ subject to the previous conditions, and $0<b\leq 1$,*

(2.3) $$\lim_{\lambda\to\infty} P\{\text{l.u.b.}_{-\infty<y<y_b}|F(y)-F_\lambda^*(y)|\leqq\alpha/\lambda^{1/2}\}=P^*(\alpha, b)\ ,$$

where

$$P^*(\alpha, b)=\begin{cases} \dfrac{4}{\pi}\sum_{k=0}^{\infty}(-1)^k/(2k+1)\exp[-(2k+1)^2\pi^2 b/8\alpha^2]\ , & \alpha>0 \\ 0\ , & \alpha\leqq 0\ . \end{cases}$$

THEOREM 4. *For $N_\lambda, X_1, X_2, \cdots$ subject to the previous conditions, and $0\leqq a<b\leqq 1$,*

(2.4) $$\lim_{\lambda\to\infty} P\{\text{l.u.b.}_{y_a\leq y\leq y_b}|F(y)-F_\lambda^*(y)|<\alpha/\lambda^{1/2}\}=P^*(\alpha, b)G(\alpha, a)\ ,$$

where $P^*(\alpha, b)$ is as in Theorem 3 and

$$G(\alpha, a)=1-(2/\pi)^{1/2}\int_{\alpha/a^{1/2}}^{\infty}\exp(-x^2/2)dx + \frac{2a^{1/2}\exp(-\alpha^2/2a)}{(2\pi)^{1/2}\alpha}\int_0^{(2k+1)\pi/2}\exp(ax^2/2\alpha^2)\sin x dx\ .$$

From (2.1) it follows that

(2.5) $$\lim_{\lambda\to\infty} P\{\text{l.u.b.}_{-\infty<y<\infty}(F(y)-F_\lambda^*(y))\leqq 0\}=0\ ,$$

that is the probability of the event that the theoretical distribution function $F(y)$ does not exceed Kac's modified empirical distribution function $F_\lambda^*(y)$ all along the interval $-\infty<y<\infty$, tends to zero as $\lambda\to\infty$. It also follows from Theorem 1 that the same is true for the interval $-\infty<y<y_b$, with $F(y_b)=b$, that is we have

COROLLARY 1.

(2.6) $$\lim_{\lambda\to\infty} P\{\text{l.u.b.}_{-\infty<y<y_b}(F(y)-F_\lambda^*(y))\leqq 0\}=0\ .$$

On the other hand Theorem 2 implies the following interesting result.

COROLLARY 2.

(2.7) $$\begin{aligned}&\lim_{\lambda\to\infty} P\{\text{l.u.b.}_{y_a\leq y\leq y_b}(F(y)-F_\lambda^*(y))\leqq 0\}\\ &=\frac{1}{\pi}\int_0^{\infty}e^{-u^2/2}\int_0^{u(a/(b-a))^{1/2}}e^{-t^2/2}dtdu\\ &=\frac{2\arctan(a/(b-a))^{1/2}}{2\pi}\\ &=\frac{1}{\pi}\arcsin(a/b)^{1/2}\ .\end{aligned}$$

Thus, the probability of the event that the theoretical distribution function $F(y)$ does not exceed Kac's modified empirical distribution function $F_\lambda^*(y)$ all along the interval $y_a \leqq y \leqq y_b$ remains positive in the limit as $\lambda \to \infty$. A similar statement was proved by Rényi [12] and Gihman [8] concerning the asymptotic behaviour of an ordinary empirical and theoretical distribution function as follows:

$$\lim_{n\to\infty} P\{\text{l.u.b.}_{y_a \leqq y \leqq y_b}(F(y)-F_n(y))\leqq 0\} = \frac{1}{\pi} \text{ arc sin } \{a(1-b)/b(1-a)\}^{1/2} .$$

Comparing (2.7) and this result we see that on a finite interval the ordinary and Kac's randomized empirical distribution function behave essentially the same way in relation to a theoretical distribution function as n and $\lambda \to \infty$ respectively.

PROOF OF THEOREMS 1, 2, 3 AND 4. Theorems 1, 2, 3 and 4 can be proved very easily via the results of [5] and a slightly modified version of Kac's method [10] as follows. Using the earlier mentioned distribution free property of the Kac statistic let $F(y)=y$, $0\leqq y \leqq 1$, and consider the process

$$x_\lambda(y)=\lambda^{1/2}\left\{y-\lambda^{-1}\sum_{j=1}^{N_\lambda}\psi_y(X_j)\right\} , \qquad 0\leqq a \leqq y \leqq b \leqq 1 , \tag{2.8}$$

with independent increments. It follows then from the analysis of Kac [10] that

$$\begin{aligned}
&\lim_{\lambda\to\infty} P\{\text{l.u.b.}_{y_a\leqq y\leqq y_b}(F(y)-F_\lambda^*(y))\leqq \alpha/\lambda^{1/2}\} \\
&\quad = \lim_{\lambda\to\infty} P\left\{\lim_{r\to\infty} \text{l.u.b.}_{1\leqq k \leqq 2^r} x_\lambda\left(a+\frac{b-a}{2^r}k\right)\leqq \alpha\right\} \\
&\quad = \lim_{\lambda\to\infty}\lim_{r\to\infty} P\left\{\text{l.u.b.}_{1\leqq k \leqq 2^r} x_\lambda\left(a+\frac{b-a}{2^r}k\right)\leqq \alpha\right\} \\
&\quad = P\{\text{l.u.b.}_{a\leqq u\leqq b}\, x(u)\leqq \alpha\} ,
\end{aligned} \tag{2.9}$$

where $\{x(u), 0\leqq u \leqq 1\}$ is the Wiener process. Similarly, in case of the two sided Kac statistic one gets

$$\begin{aligned}
&\lim_{\lambda\to\infty} P\{\text{l.u.b.}_{y_a\leqq y\leqq y_b}|F(y)-F_\lambda^*(y)|\leqq \alpha/\lambda^{1/2}\} \\
&\quad = P\{\text{l.u.b.}_{a\leqq u\leqq b}|x(u)|\leqq \alpha\} ,
\end{aligned} \tag{2.10}$$

where, again, $\{x(u), 0\leqq u \leqq 1\}$ is the Wiener process. The actual statements of Theorems 1, 2, 3 and 4 follow from known results and some calculations. Specifically we refer to (2.3), (3.2), (2.5) and (3.4) of [5] in case of Theorems 1, 2, 3 and 4 respectively.

Now we define

$$K_\lambda^+(b)=\text{l.u.b.}_{-\infty<y\leqq y_b}(F(y)-F_\lambda^*(y))$$
(2.11)
$$K_\lambda^-(b)=-\text{g.l.b.}_{-\infty<y\leqq y_b}(F(y)-F_\lambda^*(y)) ,$$

where $0<b\leqq 1$ and $F(y_b)=b$ in both cases, and put

$$R_\lambda(b)=K_\lambda^+(b)+K_\lambda^-(b) , \tag{2.12}$$

called the range of the two random variables of (2.11).

The distribution of the range of the original Kolmogorov-Smirnov statistic was derived by Kuiper [11]. Here we prove two theorems concerning the random variables of (2.11) and (2.12).

THEOREM 5. *For $N_\lambda, X_1, X_2, \cdots$ subject to the previous conditions, and $0<b\leqq 1$,*

$$\begin{aligned}(2.13)\quad &\lim_{\lambda\to\infty} P\{\lambda^{1/2}K_\lambda^-(b)\leqq v,\ \lambda^{1/2}K_\lambda^+(b)\leqq z\}\\ &=\frac{4}{\pi}\sum_{k=0}^{\infty}\frac{1}{2k+1}\exp\left[-\frac{(2k+1)^2\pi^2 b}{2(v+z)^2}\right]\sin\left[\frac{(2k+1)\pi z}{(v+z)}\right] ,\\ &\qquad\qquad v>0,\ z>0 .\end{aligned}$$

THEOREM 6. *For $N_\lambda, X_1, X_2, \cdots$ subject to the previous conditions, and $0<b\leqq 1$,*

$$\lim_{\lambda\to\infty} P\{\lambda^{1/2}R_\lambda(b)<v\}=\int_0^v \delta(b;t)dt , \qquad v>0 , \tag{2.14}$$

where

$$\delta(b;t)=8\sum_{k=1}^{\infty}(-1)^{k-1}k^2\phi(ktb^{-1/2})$$

is the asymptotic density of the range $\lambda^{1/2}R_\lambda(b)$ as $\lambda\to\infty$, and where $\phi(\cdot)$ stands for the normal density function with zero mean and unit variance.

PROOF OF THEOREMS 5 AND 6. From the analysis of Kac [10] as used in (2.9) it follows that

$$\begin{aligned}(2.15)\quad &\lim_{\lambda\to\infty} P\{\lambda^{1/2}K_\lambda^-(b)\leqq v,\ \lambda^{1/2}K_\lambda^+(b)\leqq z\}\\ &=P\{-\text{g.l.b.}_{0\leqq u\leqq b}x(u)\leqq v,\ \text{l.u.b.}_{0\leqq u\leqq b}x(u)\leqq z\}\\ &=P\{-v\leqq x(u)\leqq z,\ 0\leqq u\leqq b\leqq 1\} , \qquad v>0,\ z>0 ,\end{aligned}$$

where $\{x(u), 0\leqq u\leqq 1\}$ is the Wiener process. For the evaluation of this last statement we refer to (2.5) of [5] and this completes the proof of Theorem 5. Having established (2.15), the statement of Theorem 6 follows from (4.1), (4.2), (4.3) and (4.4) of [5].

3. Two sample versions of the one and two sided Kac statistic

Let $N_\lambda, X_{11}, X_{12}, \cdots$ and $X_{21}, X_{22}, \cdots$ be mutually independent random variables, N_λ having a Poisson distribution with mean λ and assume that the random variables X_{1i} and X_{2j} have continuous distribution functions $F(y)$ and $G(y)$ respectively. Consider random samples of the form $N_\lambda, X_{11}, X_{12}, \cdots, X_{1N_\lambda}$; $X_{21}, X_{22}, \cdots, X_{2N_\lambda}$ and the statistical hypothesis $F(y)=G(y)$. Let $F_\lambda^*(y)$ and $G_\lambda^*(y)$ be the modified empirical distribution functions of $X_{11}, X_{12}, \cdots, X_{1N_\lambda}$ and $X_{21}, X_{22}, \cdots, X_{2N_\lambda}$ respectively.

THEOREM 7. *For $N_\lambda, X_{11}, X_{12}, \cdots$ and $X_{21}, X_{22}, \cdots$ subject to the above conditions and $F(y)=G(y)$ $0<\alpha$, $c=[\alpha(2\lambda)^{1/2}]+1$,*

$$B_\lambda(\alpha)=P\{(\lambda/2)^{1/2}\,\text{l.u.b.}_{-\infty<y<\infty}(F_\lambda^*(y)-G_\lambda^*(y))\leqq\alpha\}$$
$$=1-\sum_{n=c}^{\infty}(e^{-\lambda}\lambda^n/n!)\binom{2n}{n-c}\Big/\binom{2n}{n}\,. \tag{3.1}$$

PROOF. By the independence of $N_\lambda, X_{11}, X_{12}, \cdots$ and $X_{21}, X_{22}, \cdots$,

$$B_\lambda(\alpha)=\sum_{n=0}^{\infty}P(N_\lambda=n)P\{(\lambda/2)^{1/2}\,\text{l.u.b.}_{-\infty<y<\infty}(F_{n,\lambda}(y)-G_{n,\lambda}(Y))\leqq\alpha\}\,. \tag{3.2}$$

For $N_\lambda=n$ let us pool the random samples $X_{11},\cdots,X_{1n}$ and $X_{21},\cdots,X_{2n}$ and let $Y_1, Y_2, \cdots, Y_{2n}$ be the order statistics of this pooled sample. Let $\beta_k=+1$ if Y_k is one of the values of $X_{11},\cdots,X_{1n}$ and $\beta_k=-1$ if Y_k is one of the values of $X_{21},\cdots,X_{2n}$. Put $S_k=\beta_1+\beta_2+\cdots+\beta_k$. We have

$$\text{l.u.b.}_{-\infty<y<\infty}(F_{n,\lambda}(y)-G_{n,\lambda}(y))=(\max_{1\leqq k\leqq 2n}S_k)/\lambda\,. \tag{3.3}$$

This can be seen immediately, for

$$\text{l.u.b.}_{-\infty<y<\infty}(F_{n,\lambda}(y)-G_{n,\lambda}(y))=\text{l.u.b.}_{-\infty<y<\infty}(n/\lambda)(F_n(y)-G_n(y))\,,$$

where $F_n(y)$ and $G_n(y)$ are the ordinary sample distribution functions of $X_{11},\cdots,X_{1n}$ and $X_{21},\cdots,X_{2n}$ respectively. Gnedenko and Koroljuk showed in [9] that $\text{l.u.b.}_{-\infty<y<\infty}(F_n(y)-G_n(y))=(\max_{1\leqq k\leqq 2n}S_k)/n$ and (3.3) follows. Slightly changing the analysis in [9] one obtains

$$P\{\max_{1\leqq k\leqq 2n}S_k\leqq\alpha(2\lambda)^{1/2}\}=\begin{cases}0\,, & \alpha\leqq 0\\ 1-\binom{2n}{n-c}\Big/\binom{2n}{n}\,, & 0<\alpha\leqq n/(2\lambda)^{1/2}\\ 1\,, & n/(2\lambda)^{1/2}<\alpha\,,\end{cases} \tag{3.4}$$

where $c=[\alpha(2\lambda)^{1/2}]+1$, and this, through (3.3) and (3.2), implies (3.1).

THEOREM 8. *For $N_\lambda, X_{11}, X_{12}, \cdots$ and $X_{21}, X_{22}, \cdots$ subject to the* *above conditions and* $F(y)=G(y)$, $1/(2\lambda)^{1/2}<\alpha$, $c=[\alpha(2\lambda)^{1/2}]+1$,

$$D_\lambda(\alpha)=P\{(\lambda/2)^{1/2}\,\mathrm{l.u.b.}_{-\infty<y<\infty}|F_\lambda^*(y)-G_\lambda^*(y)|\leqq\alpha\}=\sum_{n=0}^{\infty}(e^{-\lambda}\lambda^n/n!)\binom{2n}{n}^{-1}\sum_{k=-[n/c]}^{[n/c]}(-1)^k\binom{2m}{n-kc}\ . \tag{3.5}$$

PROOF. By the independence of $N_\lambda, X_{11}, X_{12},\cdots$ and $X_{21}, X_{22},\cdots$,

$$D_\lambda(\alpha)=\sum_{n=0}^{\infty}P(N_\lambda=n)P\{(\lambda/2)^{1/2}\,\mathrm{l.u.b.}_{-\infty<y<\infty}|F_{n,\lambda}(y)-G_{n,\lambda}(y)|\leqq\alpha\}\ . \tag{3.6}$$

For $N_\lambda=n$ and S_k as in proof of Theorem 7 we have here

$$\mathrm{l.u.b.}_{-\infty<y<\infty}|F_{n,\lambda}(y)-G_{n,\lambda}(y)|=\max_{1\leq k\leq 2n}|S_k|/\lambda\ . \tag{3.7}$$

Slightly changing the analysis in [9] one obtains

$$P\{\max_{1\leq k\leq 2n}|S_k|\leqq\alpha(2\lambda)^{1/2}\}=\begin{cases}0\ , & \alpha\leqq 1/(2\lambda)^{1/2}\\ \binom{2n}{n}^{-1}\sum_{k=-[n/c]}^{[n/c]}(-1)^k\binom{2n}{n-kc}\ , & 1/(2\lambda)^{1/2}<\alpha\leqq n/(2\lambda)^{1/2}\\ 1\ , & n/(2\lambda)^{1/2}<\alpha\ ,\end{cases} \tag{3.8}$$

where $c=[\alpha(2\lambda)^{1/2}]+1$, and this, through (3.7) and (3.6), implies (3.5).

Our next problem is to derive asymptotic versions of Theorems 7 and 8. This can be done in a slightly more general context. Let N_λ, $X_{11}, X_{12},\cdots$ and $N_\mu, X_{21}, X_{22},\cdots$ be mutually independent random variables, N_λ and N_μ having Poisson distribution with mean λ and μ respectively and assume that the random variables X_{1i} and X_{2j} have continuous distribution functions $F(y)$ and $G(y)$ respectively. We consider random samples of the form $N_\lambda, X_{11}, X_{12},\cdots, X_{1N_\lambda}$; $N_\mu, X_{21}, X_{22},\cdots, X_{2N_\mu}$ and the statistical hypothesis $F(y)=G(y)$. Let $F^*(y)$ and $G^*(y)$ be the respective modified empirical distribution functions of these two random samples. Let $F(y)=y$ and define the following two processes

$$x_\lambda(y)=\lambda^{1/2}\left\{y-\lambda^{-1}\sum_{j=1}^{N_\lambda}\psi_y(X_{1j})\right\}\ ,\qquad 0\leqq a\leqq y\leqq b\leqq 1\ , \tag{3.9}$$

and

$$x_\mu(y)=\mu^{1/2}y-\mu^{-1}\sum_{j=1}^{N_\mu}\psi_y(X_{2j})\ ,\qquad 0\leqq a\leqq y\leqq b\leqq 1\ . \tag{3.10}$$

Let $\Gamma=\lambda\mu/(\lambda+\mu)$ and consider the following combination of the processes of (3.9) and (3.10)

$$\Gamma^{1/2}(F_\lambda^*(y)-G_\lambda^*(y))=(\lambda/(\lambda+\mu))^{1/2}x_\mu(y)-(\mu/(\lambda+\mu))^{1/2}x_\lambda(y)\ , \tag{3.11}$$

with $0\leqq a\leqq y\leqq b\leqq 1$. Let $\Gamma\to\infty$ mean that $\mu, \lambda\to\infty$ in such a way that

$\mu/\lambda \to \rho$, where ρ is a constant. Then, repeating the argument of (2.9) and (2.10) we get

$$\lim_{\Gamma\to\infty} P\{\text{l.u.b.}_{y_a \leq y \leq y_b}(F_\lambda^*(y)-G_\mu^*(y)) \leq \alpha/\Gamma^{1/2}\} = P\{\text{l.u.b.}_{a\leq u\leq b}\, x(u) \leq \alpha\} , \tag{3.12}$$

and

$$\lim_{\Gamma\to\infty} P\{\text{l.u.b.}_{y_a \leq y \leq y_b}|F_\lambda^*(y)-G_\mu^*(y)| \leq \alpha/\Gamma^{1/2}\} = P\{\text{l.u.b.}_{a\leq u\leq b}|x(u)| \leq \alpha\} , \tag{3.13}$$

where, in both cases, $\{x(u), 0\leq u\leq 1\}$ is the Wiener process and $F(y_a) = a$, $F(y_b)=b$.

If $a=0$ and $b=1$ then (3.13) implies the two sample versions of (1.2) with distribution function as given there. Similarly, with a and b as in Theorems 1, 2, 3 and 4, (3.12) and (3.13) imply two sample versions of these theorems with respective distribution functions as given in (2.1), (2.2), (2.3) and (2.4). Two sample versions of (2.6) and (2.7) follow then immediately. Specifically, from the two sample versions of Theorem 2 of Section 2 we get the following dual of (2.7).

COROLLARY 3.

$$\lim_{\Gamma\to\infty} P\{\text{l.u.b.}_{y_a \leq y \leq y_b}(F_\lambda^*(y)-G_\mu^*(y)) \leq 0\} = \frac{1}{\pi} \arcsin (a/b)^{1/2} . \tag{3.14}$$

That is, the probability of the event that one randomized sample distribution function of Kac does not exceed another one all along the interval $y_a \leq y \leq y_b$ remains positive in the limit as $\Gamma \to \infty$.

A similar statement was proved in [6] concerning the asymptotic behaviour of two ordinary sample distribution functions.

Now if we define

$$\begin{aligned} K_{\lambda,\mu}^+(b) &= \text{l.u.b.}_{-\infty<y\leq y_b}(F_\lambda^*(y)-G_\mu^*(y)) \\ K_{\lambda,\mu}^-(b) &= -\text{g.l.b.}_{-\infty<y\leq y_b}(F_\lambda^*(y)-G_\mu^*(y)) \end{aligned} \tag{3.15}$$

and

$$R_{\lambda,\mu}(b) = K_{\lambda,\mu}^+(b) + K_{\lambda,\mu}^-(b) , \tag{3.16}$$

then two sample versions of Theorems 5 and 6 can be written down very easily with respective distribution functions as given in (2.13) and (2.14).

Also, the asymptotic version of Theorems 7 and 8 follows from consideration similar to those of the paragraph of this section which preceeds Corollary 3. In fact we have

$$\lim_{\lambda\to\infty} B_\lambda(\alpha)=\begin{cases} [2/(2\pi)^{1/2}]\int_0^\alpha e^{-u^2/2}du\ , & \alpha>0 \\ 0\ , & \alpha\leqq 0\ , \end{cases} \tag{3.17}$$

and

$$\lim_{\lambda\to\infty} D_\lambda(\alpha)=P^*(\alpha)\ , \tag{3.18}$$

where $P^*(\alpha)$ is as given under (1.2).

We go back for a moment to the one sample version of Kac's statistic, that is to the process of (2.8)

$$x_\lambda(y)=\lambda^{1/2}\left\{y-\lambda^{-1}\sum_{j=1}^{N_\lambda}\psi_y(X_j)\right\}\ ,\qquad 0\leqq y\leqq 1\ ,$$

with independent increments and $E\{x_\lambda(y)\}=0$, $E\{x_\lambda^2(y)\}=y$. Since the variance of this process is y, the most natural weight function of this process, in the spirit of Anderson and Darling's paper [3], would be $1/\sqrt{y}$ and then, in terms of stochastic processes, we would have the following statement

$$\begin{aligned}\lim_{\lambda\to\infty} P\{\text{l.u.b.}_{y_a\leqq y\leqq y_b}(F(y)-F_\lambda^*(y))/\sqrt{F}(y)\leqq\alpha/\lambda^{1/2}\} \\ =P\{\text{l.u.b.}_{a\leqq u\leqq b}x(u)\leqq\alpha\sqrt{u}\}\ ,\end{aligned} \tag{3.19}$$

where $\{x(u), 0\leqq u\leqq 1\}$ is the Wiener process. A similar statement holds for the two sided Kac statistic with weight function $1/\sqrt{F(y)}$. Using Doob's transformation [7]

$$x(u)=\sqrt{u}\,z(\log u/2\beta)\ ,$$

the process $x(u)/\sqrt{u}$ can be transformed into the Uhlenbeck process $z(u)$ with correlation parameter β, that is to a process $z(u)$ which is stationary Gaussian and Markovian with $E\{z(u)z(t)\}=\exp(-\beta|t-u|)$ and the statement of (3.19) is further equal to

$$P\{\text{l.u.b.}_{(1/2\beta)\log a\leqq t\leqq(1/2\beta)\log b}\,z(u)\leqq\alpha\}\ . \tag{3.20}$$

A similar statement can be written down for the two sided Kac statistic with weight function $1/\sqrt{F(y)}$ in terms of the Uhlenbeck process $z(u)$. Unfortunately it is not easy to evaluate this statement. The Laplace transform of the two sided version of (3.20) is given in [3]. In theory, the two sample version of Kac's original statistic with weight function $1/\sqrt{F(y)}$ can be handled the same way but we would, of course, have the same problem of evaluation as above.

One has similar difficulties when trying to derive the asymptotic version of (1.4). In terms of Wiener process we have there

$$\lim_{\lambda\to\infty} M_\lambda(\alpha/\lambda^{1/2}, b) = P\{\text{l.u.b.}_{0\leqq u\leqq b}\, x(u) \leqq \alpha(1-u)\}\ , \tag{3.21}$$

where $M_\lambda(\cdot, b)$ is defined as in (1.4), and whose exact form was derived in [4].

Acknowledgement

Financial support in the form of operating grants of the National Research Council of Canada is gratefully acknowledged by the author.

McGill University

References

[1] Allen, J. L. and Beekman, J. A. (1966). A statistical test involving a random number of random variables, *Ann. Math. Statist.*, **37**, 1305-1309.

[2] Allen, J. L. and Beekman, J. A. (1967). Distribution of a M. Kac statistic, *Ann. Math. Statist.*, **38**, 1919-1923.

[3] Anderson, T. W. and Darling, D. A. (1952). Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes, *Ann. Math. Statist.*, **23**, 193-212.

[4] Csörgő, Miklós and Alvo, Mayer (1970). Distribution results and power functions for Kac statistics, *Ann. Inst. Statist. Math.*, **22**, 257-260.

[5] Csörgő, M. (1967). A new proof of some results of Rényi and the asymptotic distribution of the range of his Kolmogorov-Smirnov type random variables, *Canadian J. Math.*, **19**, 550-558.

[6] Csörgő, M. (1965). Some Smirnov type theorems of probability theory, *Ann. Math. Statist.*, **36**, 1113-1119.

[7] Doob, J. L. (1942). The Brownian motion and stochastic equations, *Ann. of Math.*, **2**, 351-369.

[8] Gihman, T. (1952). On the empirical distribution function in the case of grouping of data (in Russian), *Dokl. Akad. Nauk SSSR*, **82**, 837-840.

[9] Gnedenko, B. V. and Koroljuk, V. S. (1951). On the maximum discrepancy between two empirical distributions (in Russian), *Dokl. Akad. Nauk SSSR*, **80**, 525-528.

[10] Kac, M. (1949). On deviations between theoretical and empirical distributions, *Proc. Nat. Acad. Sci. U.S.A.*, **35**, 252-257.

[11] Kuiper, N. H. (1960). Tests concerning random points on a circle, *Nederl. Akad. Wetensch. Proc.*, *Ser. A*, **63**, 38-47.

[12] Rényi, Alfréd (1953). On the theory of order statistics, *Acta Math. Acad. Sci. Hungar.*, **4**, 191-231.