Journal of the Korean
Statistical Society
Vol. 25, No. 3, 1996

Asymptotic Relative Efficiency of Chi-squared Type Tests Based on the Empirical Process †

Sangyeol Lee[1]

Abstract

The chi-squared type statistic generated from the empirical process can be used for testing the goodness of fit hypothesis on iid random sample. Lee (1995) showed that under some conditions, the chi-squared type statistic is asymptotically maximin in the sense of Strasser (1985). Since the chi-squared type statistic depends on the choice of points in the unit interval, it is worth investigating the points yielding more efficient tests. Motivated by this viewpoint, we are led to study the asymptotic relative efficiency of chi-squared type tests in the same setting of Lee (1995). Some examples are given for illustration.

Key Words : Chi-squared type statistic; Goodness of fit hypothesis; Empirical process; Asymptotic relative efficiency.

†This work is supported by Sookmyung Women's University Research Fund in 1996.
[1]Department of Statistics, Sookmyung Women's University, Seoul, 140-742, Korea.

1. INTRODUCTION

Suppose that one wishes to test a goodness-of-fit hypothesis on random sample. Koutrouvelis and Kellermeier (1981) and Csörgo (1986) have suggested the empirical characteristic function as a test procedure and proposed a chi-squared type test statistic. Similarly, it is possible to use the chi-squared type statistic generated by the empirical process. Lee (1995) investigated its statistical meaning by studying the asymptotic behavior of the empirical process under a sequence of contiguous alternatives. He showed that under certain conditions, the chi-squared type statistic is asymptotically maximin in the sense of Strasser (1985, P. 136). In this paper, we study the asymptotic relative efficiencey of the chi-squared type statistic in the same setting of Lee.

Let $X_1, X_2, \ldots,$ be iid random variables with distribution $F(x : \underline{\theta})$, where $\underline{\theta}$ is an unknown parameter. The estimated empirical process based on $X_1, \ldots, X_n$ is defined by

$$\hat{Y}_n(t) = n^{-1/2} \sum_{j=1}^{n} \left[I(F(X_j : \hat{\underline{\theta}}_n) \leq t) - t \right], \; t \in [0,1], \tag{1.1}$$

where $I(\cdot)$ denotes the indicator function and $\hat{\underline{\theta}}_n$ is an estimate of $\underline{\theta}$. Under regularity conditions, $\hat{Y}_n$ converges weakly to a mean zero Gaussian process Y (cf. Shorak and Wellner (1986, Ch. 5.5)). Assume that $\hat{Y}_n$ converges weakly to Y. Then the chi-squared type test statistic at the points $t_1, \ldots, t_k$ in $[0,1]$ is defined by

$$\hat{\underline{y}}_n' \Sigma^{-1} \hat{\underline{y}}_n, \tag{1.2}$$

where

$$\hat{\underline{y}}_n = \hat{\underline{y}}_n(t_1, \ldots, t_k)' = \left(\hat{Y}_n(t_1), \ldots, \hat{Y}_n(t_k) \right)', \tag{1.3}$$

and $\Sigma = (\sigma_{ij})_{i,j=1}^{k}$ is the positive definite matrix whose (i,j)-th entry is $EY(t_i)Y(t_j)$. Since $\hat{Y}_n \xrightarrow{\mathcal{D}} Y$, we have that

$$\hat{\underline{y}}_n' \Sigma^{-1} \hat{\underline{y}}_n \xrightarrow{\mathcal{D}} \chi^2(k), \tag{1.4}$$

where $\chi^2(k)$ denotes a chi-square random variable with the degrees of freedom k.

On the other hand, under some sequence of contiguous alternatives, the empirical process $\hat{Y}_n$ in (1.1) converges weakly to another Gaussian process $Y^*(t) = \mu(t) + Y(t)$, where $\mu(t)$ is a drift (cf. Durbin (1973)). As a consequence, under the alternatives,

$$\underline{\hat{y}}_n' \Sigma^{-1} \underline{\hat{y}}_n \xrightarrow{\mathcal{D}} \chi^2(k, \underline{\mu}' \Sigma^{-1} \underline{\mu}), \tag{1.5}$$

where $\underline{\mu} = (\mu(t_1), \ldots, \mu(t_k))'$ and $\chi^2(k, \nu)$ denotes a noncentral chi-square random variable with the degrees of freedom k and noncentrality parameter ν.

Lee (1995) showed that under certain conditions, (1.2) has the maximin property asymptotically as sample size increases. For showing the asymptotic maximin property, he introduced a family of alternative hypotheses $\mathcal{H} = \{\{H_n(\gamma, G)\} : \gamma > 0, G \in \mathcal{G}\}$, where $\mathcal{G}$ is an arbitrary index set. A typical example is

$$H_n : X_1, \ldots, X_n \sim (1 - \gamma/\sqrt{n})F + (\gamma/\sqrt{n})G, \ \gamma > 0,$$

where F is a true underlying distribution under the null hypothesis and G is a distribution function. Then one can write $\mathcal{H} = \{\{H_n(\gamma, G)\} : \gamma > 0, G \in \mathcal{G}\}$ where $\mathcal{G}$ is the family of all distributions on the real line.

In this paper, we denote by H_0 the null hypothesis and assume that $\{H_n(\gamma, G)\}$ at $\gamma = 0$ becomes the null $\{H_0\}$. Furthermore, it is assumed that $\hat{Y}_n$ in (1.1) converges weakly to a Gaussian process with a drift which is proportional to $\gamma \geq 0$, say, $\gamma\mu_G$ under $\{H_n(\gamma, G)\}$, whereas the covariance structure remains the same as under the null H_0. These assumptions are natural, for instance, in view of the results of Durbin (1973).

Note that the statistic defined by (1.2) depends on the choice of the points $t_1, \ldots, t_k$. Hence, it is natural to ask for the optimal point to yield more efficient test. In Section 2, we answer this question by computing the asymptotic relative efficiency for fixed $G \in \mathcal{G}$. Some examples are given in Section 3.

2. ASYMPTOTICALLY RELATIVE EFFICIENCY

We start Section 2 with a lemma. Since the proof is rather standard, we omit the proof.

Lemma 1. Let $\underline{Z}$ be a $k \times 1$ normal random vector. For testing $H_0 : \underline{Z} \sim \mathcal{N}(0, \Sigma)$ vs. $H_1 : \underline{Z} \sim \mathcal{N}(\gamma\underline{e}, \Sigma)$, $\gamma > 0$, where $\underline{e}$ is a vector with $\underline{e}'\Sigma^{-1}\underline{e} = 1$, the test

$$\phi(\underline{x}) = \begin{cases} 1 & \text{if } (\underline{e}, \underline{x}) > z_\alpha \\ 0 & \text{if } (\underline{e}, \underline{x}) < z_\alpha \end{cases}$$

is a UMP test, where $(\underline{a}, \underline{b}) = \underline{a}'\Sigma^{-1}\underline{b}$. For a significance level α, the critical value z_α is the $100(1-\alpha)$-th percentile of the standard normal distribution.

Generally speaking, if a sequence of tests $\{S_n\}$ can attain the same power with a smaller sample size than another tests $\{T_n\}$, we say that $\{S_n\}$ is more efficient than $\{T_n\}$. Based on this, we consider the problem of choosing k points $t_1, \ldots, t_k \in [0,1]$ which provides a more efficient chi-squared type test.

Let $\delta \in (0,1]$ and $[n\delta]$ denote the largest integer less than or equal to $n\delta$. Suppose that $X_1, \ldots, X_{[n\delta]}$ are available observations and $\hat{Y}_{[n\delta]}(t) \xrightarrow{\mathcal{D}} \gamma\sqrt{\delta}\mu_G(t) + Y(t)$ under $\{H_n(\gamma, G)\}$. (Examples can be found in Section 3). Here, $\hat{Y}_n$ is the empirical process defined by (1.1) and $Y(t)$ is a mean zero Gaussian process. For each $\underline{t} = (t_1, \ldots, t_k)'$, $t_i \in [0,1]$, $\hat{\underline{y}}_n(\underline{t}) = (\hat{Y}_n(t_1), \ldots, \hat{Y}_n(t_k))'\ \underline{\mu}_G(\underline{t}) = (\mu_G(t_1), \ldots, \mu_G(t_k))'$, and $\Sigma(\underline{t})$ denotes the positive defintie matrix whose (i,j)-th entry is $EY(t_i)Y(t_j)$. Since

$$\hat{\underline{y}}_n(\underline{t}) \xrightarrow{\mathcal{D}} \underline{X} \sim \mathcal{N}(\underline{0}, \Sigma(\underline{t})) \quad \text{under } H_0$$

and

$$\hat{\underline{y}}_n(\underline{t}) \xrightarrow{\mathcal{D}} \underline{X} \sim \mathcal{N}(\gamma\underline{\mu}_G(\underline{t}), \Sigma(\underline{t})) \quad \text{under } \{H_n(\gamma, G)\},$$

the original sequence of problem of testing H_0 vs. $\{H_n(\gamma, G)\}$ is asymptotically reduced to the problem of testing K_0 vs. K_1 such that

$$K_0 : \underline{X} \sim \mathcal{N}(0, \Sigma(\underline{t}))$$

and

$$K_1 : \underline{X} \sim \mathcal{N}(\gamma\underline{\mu}_G(\underline{t}), \Sigma(\underline{t})).$$

Let $(\underline{a}, \underline{b})_{\underline{t}} = \underline{a}'\Sigma(\underline{t})^{-1}\underline{b}$ and $||\underline{a}||_{\underline{t}} = (\underline{a}, \underline{a})_{\underline{t}}^{1/2}$. By Lemma 1,

$$\phi_{\underline{t}}(\underline{x}) = \begin{cases} 1 & \text{if } \left(\frac{\underline{\mu}_G(\underline{t})}{||\underline{\mu}_G(\underline{t})||_{\underline{t}}}, \underline{x}\right) > z_\alpha \\ 0 & \text{otherwise} \end{cases} \tag{2.1}$$

is an α-level UMP test for K_0 vs. K_1. On the other hand, if we use the sample size $[n\delta]$, it can be shown that for each $\gamma > 0$,

$$\left(\frac{\underline{\mu}_G(\underline{t})}{||\underline{\mu}_G(\underline{t})||_{\underline{t}}}, \underline{X}\right) \sim \mathcal{N}(\gamma\sqrt{\delta}||\underline{\mu}_G(\underline{t})||_{\underline{t}}, 1) \quad \text{under} \ K_1. \tag{2.2}$$

Now, let $\underline{s} = (s_1, \ldots, s_k)'$, $s_i \in [0,1]$. Then testing H_0 vs. $\{H_n(\gamma, G)\}$ via using $\hat{\underline{y}}_n(\underline{s})$ is reduced to testing K_0' vs. K_1' such that

$$K_0' : \underline{X} \sim \mathcal{N}(\underline{0}, \Sigma(\underline{s}))$$

and

$$K_1' : \underline{X} \sim \mathcal{N}(\gamma \frac{\underline{\mu}_G(\underline{s})}{||\underline{\mu}_G(\underline{s})||_{\underline{s}}}, \Sigma(\underline{s})).$$

Then

$$\phi_{\underline{s}}(\underline{x}) = \begin{cases} 1 & \text{if} \ \left(\frac{\underline{\mu}_G(\underline{s})}{||\underline{\mu}_G(\underline{s})||_{\underline{s}}}, \underline{x}\right)_{\underline{s}} > z_\alpha \\ 0 & \text{otherwise} \end{cases} \tag{2.3}$$

is a UMP test of K_0' vs. K_1' and

$$\left(\frac{\underline{\mu}_G(\underline{s})}{||\underline{\mu}_G(\underline{s})||_{\underline{s}}}, \underline{X}\right) \sim \mathcal{N}(\gamma||\underline{\mu}_G(\underline{s})||_{\underline{s}}, 1) \quad \text{under} \ K_1'. \tag{2.4}$$

From (2.2) and (2.4), one can observe that

$$\gamma\sqrt{\delta}||\underline{\mu}_G(\underline{t})||_{\underline{t}} = \gamma||\underline{\mu}_G(\underline{s})||_{\underline{s}} \ \text{ or } \ \delta = ||\underline{\mu}_G(\underline{s})||_{\underline{s}}^2/||\underline{\mu}_G(\underline{t})||_{\underline{t}}^2$$

yields the same power. Since we are comparing best tests in view of (2.1), (2.3) and Lemma 1, we are able to say that: if $\delta < 1, \hat{y}_n(\underline{t})$ yields a more efficient test than $\hat{y}_n(\underline{s})$, and otherwise $\hat{y}_n(\underline{s})$ would be preferred. This allows us to define the following.

Definition 1. For fixed $G, R = \underline{\mu}_G'(\underline{s})\Sigma^{-1}(\underline{s})\underline{\mu}_G(\underline{s})/\underline{\mu}_G'(\underline{t})\Sigma^{-1}(\underline{t})\underline{\mu}_G(\underline{t})$ is the asymptotic relative efficiency of the point $\underline{s}$ with resprect to $\underline{t}$.

For fixed G, we can say that $\underline{t}$ is 'optimal' if

$$\underline{\mu}_G'(\underline{t})\Sigma^{-1}(\underline{t})\underline{\mu}_G(\underline{t}) = \sup_{\underline{s}} \underline{\mu}_G'(\underline{s})\Sigma^{-1}(\underline{s})\underline{\mu}_G(\underline{s}).$$

In Definition 1, we can see that the test procedure based on the points which yields a larger drift and smaller variance is more efficient. This is obvious in one dimensional case and agrees to our intuition.

3. EXAMPLES

In this section we give some examples where the optimal point in the sense of Definition 1 can be derived via the weak convergence result of empirical processes. For any distribution function F, we set

$$F^{-1}(t) = \inf\{x : F(x) \geq t\}, \quad t \in [0,1].$$

3.1. Continuous Distribution Case

Suppose that $X_1, X_2, \ldots,$ are i.i.d. random variables with a continuous distribution F. Then it is well known that the empirical process

$$Y_n(t) = n^{-1/2} \sum_{j=1}^{n} \left[I(F(X_j) \leq t) - t \right], \quad t \in [0,1] \tag{3.1}$$

converges to a standard Brownian bridge $\overset{o}{W}(t)$ (cf. Billingsley (1968)).

Now let us consider a sequence of a contiguous alternatives

$$H_n : X_1, \ldots, X_n \sim (1 - \gamma/\sqrt{n})F + (\gamma/\sqrt{n})G, \quad \gamma > 0, \tag{3.2}$$

where $G \in \mathcal{G} = \{G : G \text{ is a continuous distribution}\}$.

Theorem 1. Let $\delta \in (0,1]$. Under the alternatives $\{H_n\}$ in (3.2), we have that

$$\hat{Y}_{[n\delta]}(t) \xrightarrow{\mathcal{D}} -\gamma\sqrt{\delta}(t - G \circ F^{-1}(t)) + \overset{o}{W}(t).$$

Proof. Let $U_1, U_2, \ldots,$ be i.i.d. uniformly distributed random variables over $[0,1]$ and let $F_n = (1 - \gamma/\sqrt{n})F + (\gamma/\sqrt{n})G$. Observe that

$$\begin{aligned} Y_{[n\delta]}(t) &= [n\delta]^{-1/2} \sum_{j=1}^{[n\delta]} \Big[I(F_n(X_j) \leq F_n \circ F^{-1}(t)) - t \Big] \quad \text{a.s.} \\ &\overset{\mathcal{D}}{=} [n\delta]^{-1/2} \sum_{j=1}^{[n\delta]} \left[I(U_j \leq t) - t \right] \\ &\quad + [n\delta]^{1/2} \Big(F_n \circ F^{-1}(t) - t \Big) + R_n(t), \end{aligned}$$

where

$$R_n(t) = [n\delta]^{-1/2} \sum_{j=1}^{[n\delta]} \left[I(U_j \le F_n \circ F^{-1}(t)) - F_n \circ F^{-1}(t) + t - I(U_j \le t) \right] \quad (3.3)$$

Since for $x \in R$,

$$[n\delta]^{1/2}(F_n(x) - F(x)) = -\gamma\left([n\delta]/n\right)^{1/2}(F(x) - G(x)), \quad (3.4)$$

we have

$$\sup_{0 \le t \le 1} \left| F_n \circ F^{-1}(t) - t \right| \le 2\gamma/\sqrt{n} \to 0 \quad \text{as } n \to \infty.$$

This together with the arguments of Billingsley(1968, P. 106) yields

$$\sup_t |R_n(t)| \xrightarrow{P} 0.$$

In view of (3.4), the theorem is established. □

Let $T = \{(t_1, \ldots, t_k) : 0 \le t_i \le 1, t_i \ne t_j \text{ for } i \ne j\}$. For each $\underline{t} \in T$, let $\hat{\underline{y}}_n(\underline{t}) = (\hat{Y}_n(t_1), \ldots, \hat{Y}_n(t_k))', \Sigma(\underline{t}) = (\sigma_{ij}(\underline{t}))_{i,j=1}^k$, $\sigma_{ij}(\underline{t}) = t_i \wedge t_j - t_i t_j$, and $\underline{\mu}(\underline{t}) = \underline{\mu}_G(\underline{t}) = (\mu(t_1), \ldots, \mu(t_k))'$. Here, $\mu(t) = \mu_G(t) = t - G \circ F^{-1}(t)$.

Now, let $G \in \mathcal{G}$ and $\underline{s}, \underline{t} \in T$. In view of Definition 1, the asymptotic relative efficiency is given by

$$\underline{\mu}'(\underline{s})\Sigma^{-1}(\underline{s})\underline{\mu}(\underline{s}) / \underline{\mu}'(\underline{t})\Sigma^{-1}(\underline{t})\underline{\mu}(\underline{t}).$$

and an optimal point $\underline{t}$ can be obtained by the way that

$$\underline{\mu}'(\underline{t})\Sigma^{-1}(\underline{t})\underline{\mu}(\underline{t}) = \sup_{\underline{s} \in T} \underline{\mu}'(\underline{s})\Sigma^{-1}(\underline{s})\underline{\mu}(\underline{s}).$$

3.2. Normal Distribution with Unknown Mean

Suppose that $X_1, \ldots, X_n$ are i.i.d. random variables whose distribution is $\Phi(x - \theta)$, where $\Phi(\cdot)$ denotes the distribution function of $\mathcal{N}(0,1)$ and θ is unknown location parameter. In the sequel, we denote $\phi = \Phi'$. Let H_0 and $\{H_n\}$ be the null and alternative hypotheses, respectively, such that for $1 \le i \le n$,

$$H_0 \ : \ X_i \sim \Phi(x - \theta), \quad (3.5)$$

$$H_n \ : \ X_i \sim \Phi_n(x - \theta), \quad (3.6)$$

where $\Phi_n(x) = (1-\gamma/\sqrt{n})\Phi + (\gamma/\sqrt{n})G, \gamma > 0, G \in \mathcal{G}$. Here, $\mathcal{G}$ denotes the family of all symmetric distributions G such that

$$\int x^2 dG(x) < \infty, \ \sup_x |G'(x)| < \infty \quad \text{and} \quad \sup_x |G''(x)| < \infty. \tag{3.7}$$

Now, we consider the empirical process defined by

$$\hat{Y}_n(t) = n^{-1/2} \sum_{j=1}^{n} \left[I(\Phi(X_j - \hat{\theta}_n) \le t) - t \right], \ t \in [0,1], \tag{3.8}$$

where $\hat{\theta}_n = n^{-1} \sum_{j=1}^{n} X_j$.

Theorem 2. Let $\delta \in (0,1]$. Under $\{H_n\}$ in (3.6),

$$\hat{Y}_{[n\delta]}(t) \xrightarrow{\mathcal{D}} -\gamma\sqrt{\delta}(t - G \circ \Phi^{-1}(t)) + Z(t), \tag{3.9}$$

where Z is a mean zero Gaussian process such that

$$EZ(s)Z(t) = s \wedge t - st - \phi(\Phi^{-1}(s))\phi(\Phi^{-1}(t)) \quad \text{for} \quad s,t \in [0,1]. \tag{3.10}$$

Proof. Let $\varepsilon_j = X_j - \theta$. Then,

$$\begin{aligned} \hat{Y}_{[n\delta]}(t) &= [n\delta]^{-1/2} \sum_{j=1}^{[n\delta]} \left[I(\varepsilon_j \le \Phi^{-1}(t) + (\hat{\theta}_{[n\delta]} - \theta)) - t \right] \\ &= [n\delta]^{-1/2} \sum_{j=1}^{[n\delta]} \left[I(\Phi_n(\varepsilon_j) \le t) - t \right] \\ &+ [n\delta]^{1/2} \Phi_n(\alpha_n(t)) + R_n(t), \end{aligned}$$

where

$$\alpha_n(t) = \Phi_n(\Phi^{-1}(t) + (\hat{\theta}_{[n\delta]} - \theta)) - t$$

and

$$R_n(t) = [n\delta]^{-1/2} \sum_{j=1}^{n} \left[I(\Phi_n(\varepsilon_j) \le \alpha_n(t)) - \alpha_n(t) + t - I(\Phi_n(\varepsilon_j) \le t) \right].$$

By Taylor's series expansion,

$$\begin{aligned}\alpha_n(t) &= \left[\Phi_n(\Phi^{-1}(t)+\hat{\theta}_{[n\delta]}-\theta)-\Phi_n(\Phi^{-1}(t))\right]+\left[\Phi_n(\Phi^{-1}(t))-t\right] \\ &= (\hat{\theta}_{[n\delta]}-\theta)\Phi_n'(\Phi^{-1}(t))+(\hat{\theta}_{[n\delta]}-\theta)^2\Phi_n''(\zeta)/2 \\ &+ \gamma(t-G\circ\Phi^{-1}(t))/\sqrt{n},\end{aligned}$$

where ζ lies between $\Phi^{-1}(t)$ and $\Phi^{-1}(t)+(\hat{\theta}_{[n\delta]}-\theta)$. Thus it follows from the fact $\hat{\theta}_{[n\delta]}-\theta=[n\delta]^{-1}\sum_{j=1}^{[n\delta]}\varepsilon_j$ that

$$[n\delta]^{1/2}\alpha_n(t)=[n\delta]^{-1/2}\sum_{j=1}^{[n\delta]}\varepsilon_j\phi(\Phi^{-1}(t))-\gamma\sqrt{\delta}(t-G\circ\Phi^{-1}(t))+\Delta_n(t),$$

where $\sup_t|\Delta_n(t)|\xrightarrow{P}0$ under $\{H_n\}$. This together with the arguments of Billingsley (1968, P. 106) yields $\sup_t|R_n(t)|\xrightarrow{P}0$. Therefore, we can write

$$\hat{Y}_{[n\delta]}(t)=-\gamma\sqrt{\delta}(t-G\circ\Phi^{-1}(t))+Z_{[n\delta]}(t)+\eta_n(t), \tag{3.11}$$

where $\sup_t|\eta_n(t)|\xrightarrow{P}0$ and

$$Z_n(t)=n^{-1/2}\sum_{j=1}^{n}\left[I(\Phi_n(\varepsilon_j)\le t)-t+\phi(\Phi^{-1}(t))\varepsilon_j\right].$$

Since

$$n^{-1/2}\sum_{j=1}^{n}[I(\Phi_n(\varepsilon_j)\le t)-t]\overset{\mathcal{D}}{=}n^{-1/2}\sum_{j=1}^{n}[I(U_j\le t)-t],$$

where U_j are iid $U[0,1]$ and since $\sup_t|\phi(\Phi^{-1}(t))|<\infty$, $\{Z_n(t):n\ge 1\}$ is tight. On the other hand, by central limit theorem,

$$(Z_n(t_1),\ldots,Z_n(t_k))\xrightarrow{\mathcal{D}}\mathcal{N}(\underline{0},\Sigma),$$

where (i,j)-th entry of Σ is equal to the right hand side of (3.10). As a consequence, Z_n converges weakly to a Gaussian process Z with mean zero and the covariance structure as in (3.10). In view of (3.11), we obtain (3.9).

Let $T=\{(t_1,\ldots,t_k):0<t_i<1/2,\ t_i\ne t_j \text{ for } i\ne j\}$ and let for each $\underline{t} \in T, \hat{\underline{y}}_n(\underline{t}) = (\hat{Y}_n(t_1), \ldots, \hat{Y}_n(t_k))', \Sigma(\underline{t}) = (\sigma_{ij}(\underline{t}))_{i,j=1}^k, \sigma_{ij}(\underline{t}) = t_i \wedge t_j - t_i t_j - \phi(\Phi^{-1}(t_i))\phi(\Phi^{-1}(t_j))$ and $\underline{\mu}(\underline{t}) = \underline{\mu}_G(\underline{t}) = (\mu(t_1), ..., \mu(t_k))'$, where $\mu(t) = \mu_G(t) = t - G \circ \Phi^{-1}(t)$.

The asymptotic relative efficiency and the optimal point can be obtained in the same way as Example (1).

REFERENCES

(1) Billingsley, P. (1968). *Convergence of Probability Measures*, Wiley, New York.

(2) Csörgo, S. (1986). Testing for normality in arbitrary dimension. *Annals of Statististics*, **14**, 708-723.

(3) Durbin, J. (1973). Weak vonvergence of the sample distribution function when parameters are estimated. *Annals of Statistics*, **1**, 279-290.

(4) Koutrouvelis, I. A. and Kellermeire, J. (1981). A goodness of fit test based on the empirical characteristic function when parameters must be estimated. *Journal of Royal Statistical Society*, B, **43**, 173-176.

(5) Lee, S. (1995). Asymptotic maximin property of chi-squared type tests based on the empirical process. *To appear in Statistics and Probability Letters.*

(6) Shorack, G. and Wellner, J. (1986). *Empirical Processes with Applications to Statistics.* Wiley, New York.

(7) Strasser, H. (1985). *Mathematical Theory of Statistics.* de Gruyter, Amsterdam.