KŌDAI MATH. SEM. REP.
19 (1967), 138–146

RELATIVE EFFICIENCY OF THE WALD SPRT AND THE CHERNOFF INFORMATION NUMBER

By Minoru Sakaguchi

Summary: Relative efficiencies measured in terms of ratio of ASN to fixed sample size of the Wald SPRT to the best competing fixed sample procedure with fixed error probabilities are given for the exponential family of densities $f_\theta(x)$. The limiting relative efficiency when the error probabilities approach zero in a particular manner is computed, and it is shown that this value is $+\infty$ at $\theta=\theta_{01}$, some parameter value determined by the two hypothetical densities (Section 1). In Section 2 we show that the parameter value θ_{01} is closely connected with the Chernoff information number discriminating between the two densities.

§1. Relative efficiency of the Wald SPRT for the exponential family of densities.

Let us consider the classical problem of testing H_0: $f(x)=f_0(x)$ versus H_1: $f(x)=f_1(x)$ when X is distributed with the generalized pdf $f(x)$. The test is to have the specified probabilities of error, α_0 (the probability of rejecting H_0 when H_0 is true) and α_1 (the probability of accepting H_0 when H_1 is true).

Writing the logarithm of likelihood ratio as

$$\log \prod_{j=1}^{N} \frac{f_1(x_j)}{f_0(x_j)} = \sum_{j=1}^{N} \log \frac{f_1(x_j)}{f_0(x_j)} \equiv \sum_{j=1}^{N} Z_j$$

the likelihood-ratio test accepts $H_1(H_0)$, if $\sum_{j=1}^{N} Z_j > (\leqq) k$. If N is not very small $\sum_{j=1}^{N} Z_j$ will be approximately normally distributed with mean $NE(z)$ and variance $N\sigma^2(z)$, under any distribution with density $f(x)$. Thus if a sample of fixed size is to be used to discriminate between two simple hypotheses H_0 and H_1, then the condition that the test is to have the specified error probabilities requires

$$\Phi\left(\frac{k+NI(0:1)}{N^{1/2}\sigma(z|H_0)}\right) \doteqdot \alpha_0, \qquad 1-\Phi\left(\frac{k-NI(1:0)}{N^{1/2}\sigma(z|H_1)}\right) \doteqdot \alpha_1$$

where

$$I(i:1-i) = \int f_i(x) \log \frac{f_i(x)}{f_{1-i}(x)}\, d\nu(x) \qquad (i=0, 1)$$

is the Kullback-Leibler information number [4] and

Received December 15, 1965.

$$\Phi(u)\equiv\int_u^\infty \frac{1}{\sqrt{2\pi}}e^{-t^2/2}dt.$$

With $\Phi(\lambda_\alpha)=\alpha$ and eliminating k we find

$$N\doteqdot(\lambda_{\alpha_0}\sigma(z|H_0)+\lambda_{\alpha_1}\sigma(z|H_1))^2/J^2=(\lambda_{\alpha_0}+\lambda_{\alpha_1})^2\bar{\sigma}^2/J^2 \tag{1.1}$$

where $J\equiv I(0:1)+I(1:0)$ and

$$\bar{\sigma}\equiv\frac{\lambda_{\alpha_0}\sigma(z|H_0)+\lambda_{\alpha_1}\sigma(z|H_1)}{\lambda_{\alpha_0}+\lambda_{\alpha_1}}. \tag{1.2}$$

Now, if we use the Wald sequential probability ratio test (SPRT) with the strength (α_0, α_1), (i.e., Pr. {Accept $H_{1-i}|H_i\}\leqq\alpha_i$, $i=0,1$), the Wald approximations "of neglecting excess" yields the operating characteristic (OC) function and average sample number (ASN) as follows:

$$L(H)\equiv\text{Pr}\,\{\text{Accept } H_0|H\}\doteqdot\frac{A^h-1}{A^h-B^h}, \tag{1.3}$$

$$E(n|H)\doteqdot\frac{L(H)\log B+(1-L(H))\log A}{E(Z|H)} \tag{1.4}$$

under any hypothesis H, where $A=(1-\alpha_1)/\alpha_0$, $B=\alpha_1/(1-\alpha_0)$ and h is the unique non-zero value satisfying

$$E\left[\left(\frac{f_1(X)}{f_0(X)}\right)^h\middle|H\right]=1 \qquad \text{(Wald [6]).}$$

Let us compute the relative efficiency of the Wald SPRT for the exponential family of pdf's (relative to the dominating measure ν)

$$f_\theta(x)=e^{\theta x-\omega(\theta)} \qquad \left(\int e^{\theta x}d\nu(x)=e^{\omega(\theta)}\right).$$

Existence of the moment generating function implies existence of all moments. In particular we have $E_\theta(X)=\omega'(\theta)$, $\text{Var}_\theta(X)=\omega''(\theta)$. For any parameter values θ_0, θ_1 with $\theta_0<\theta_1$ and θ in the domain of definition, we find

$$\begin{aligned}
I(\theta_i:\theta_{1-i})&=E_{\theta_i}[\log(f_{\theta_i}(X)/f_{\theta_{1-i}}(X))]=(\theta_i-\theta_{1-i})\omega'(\theta_i)-(\omega(\theta_i)-\omega(\theta_{1-i})),\\
J(\theta_0,\theta_1)&=I(\theta_0:\theta_1)+I(\theta_1:\theta_0)=(\theta_1-\theta_0)(\omega'(\theta_1)-\omega'(\theta_0)),\\
\sigma^2(Z|H_i)&=\text{Var}_{\theta_i}[\log(f_{\theta_1}(X)/f_{\theta_0}(X))]=(\theta_1-\theta_0)^2\omega''(\theta_i),\\
E(Z|H)&=E_\theta[\log(f_{\theta_1}(X)/f_{\theta_0}(X))]=(\theta_1-\theta_0)\omega'(\theta)-(\omega(\theta_1)-\omega(\theta_0)),
\end{aligned}$$

and

$$L(\theta)\equiv\text{Pr}\,\{\text{Accept } H_0:\theta=\theta_0|\theta\}=\frac{A^{h(\theta)}-1}{A^{h(\theta)}-B^{h(\theta)}}$$

where $h(\theta)$ is, for each θ, the unique non-zero value satisfying

$$\omega(\theta+(\theta_1-\theta_0)h(\theta))-\omega(\theta)=(\omega(\theta_1)-\omega(\theta_0))h(\theta). \tag{1.5}$$

Substituting these expressions into (1. 1)$\sim$(1. 4), the relative efficiency (measured in terms of ratio of ASN to fixed sample size) of the Wald SPRT to the best competing fixed sample procedure with fixed error probabilities α_0, α_1 is given by

$$\text{(1. 6)}\qquad \begin{aligned} \mathrm{RE}\,(\alpha_0, \alpha_1; \theta) &= \frac{E_\theta(n)}{N} \\ &\doteqdot \frac{L(\theta)\log B+(1-L(\theta))\log A}{(\theta_1-\theta_0)\omega'(\theta)-(\omega(\theta_1)-\omega(\theta_0))}\cdot\frac{(\omega'(\theta_1)-\omega'(\theta_0))^2}{\left(\lambda_{\alpha_0}\sqrt{\omega''(\theta_0)}+\lambda_{\alpha_1}\sqrt{\omega''(\theta_1)}\right)^2}. \end{aligned}$$

Next we shall obtain the limit of (1. 6) when $\alpha_1=a\alpha_0^b$ $(a, b>0)$ and $\alpha_0\to 0$. Since

$$L(\theta)\doteqdot\frac{(1-\alpha_0)^h}{1-a^h\alpha_0^{(1+b)h}} \xrightarrow[(\alpha_0\to 0)]{} \begin{cases} 1, & \text{if } h>0, \\ 0, & \text{if } h<0 \end{cases}$$

and since the well-known approximation

$$\alpha=\Phi(\lambda_\alpha)\doteqdot\left(\frac{1}{\sqrt{2\pi}}e^{-\lambda_\alpha^2/2}\right)\Big/\lambda_\alpha$$

gives

$$\log\alpha\doteqdot\log(2\pi)^{-1/2}-\frac{1}{2}\lambda_\alpha^2-\log\lambda_\alpha=-\frac{1}{2}\lambda_\alpha^2$$

we find that

$$\text{(1. 7)}\qquad \begin{aligned} &\frac{(\theta_1-\theta_0)\omega'(\theta)-(\omega(\theta_1)-\omega(\theta_0))}{(\omega'(\theta_1)-\omega'(\theta_0))^2}\cdot\lim_{\alpha_0\to 0}\mathrm{RE}(\alpha_0, a\alpha_0^b; \theta) \\ &=\begin{cases} -\dfrac{b}{2\left(\sqrt{\omega''(\theta_0)}+\sqrt{b\omega''(\theta_1)}\right)^2}, & \text{if } h(\theta)>0, \\[2ex] \dfrac{1}{2\left(\sqrt{\omega''(\theta_0)}+\sqrt{b\omega''(\theta_1)}\right)^2}, & \text{if } h(\theta)<0. \end{cases} \end{aligned}$$

This shows that the limiting relative efficiency depends on the .particular choice of θ and the relative rate at which α_0 and α_1 approach zero in a particular manner. Define as a function of y, for each fixed θ,

$$\text{(1. 8)}\qquad K_\theta(y)\equiv\begin{cases} (\omega(\theta+y)-\omega(\theta))/y, & \text{if } y\neq 0, \\ w'(\theta), & \text{if } y=0. \end{cases}$$

By the strict convexity of $\omega(\theta)$, this is strictly increasing both in θ and y and continuous at $y=0$. From (1. 5) $h(\theta)$ is explicitly given by

$$\text{(1. 9)}\qquad h(\theta)=\frac{1}{\theta_1-\theta_0}K_\theta^{-1}\left(\frac{\omega(\theta_1)-\omega(\theta_0)}{\theta_1-\theta_0}\right),$$

where $K_\theta^{-1}(y)$ is the inverse function of $K_\theta(y)$, for each fixed θ. Since $K_\theta^{-1}(\omega'(\theta))=0$ we have $h(\theta)=0$, if and only if

$$\text{(1. 10)}\qquad \omega'(\theta)=\frac{\omega(\theta_1)-\omega(\theta_0)}{\theta_1-\theta_0}.$$

Let us denote this value of θ by θ_{01}. This value exists uniquely between θ_0 and θ_1, and by the increasing property of $K_\theta(y)$ both in θ and y we have

$$h(\theta)\begin{Bmatrix}>\\=\\<\end{Bmatrix}0, \qquad \text{according as} \quad \theta\begin{Bmatrix}>\\=\\<\end{Bmatrix}\theta_{01}. \tag{1.11}$$

Thus if $b=1$ (that is if $\alpha_1=a\alpha_0$ and $\alpha_0\to 0$) we have from (1.7) and (1.11)

$$\lim_{\alpha_0\to 0}\mathrm{RE}\,(\alpha_0, a\alpha_0; \theta)=\frac{1}{2|(\theta_1-\theta_0)\omega'(\theta)-(\omega(\theta_1)-\omega(\theta_0))|}\cdot\left(\frac{\omega'(\theta_1)-\omega'(\theta_0)}{\sqrt{\omega''(\theta_0)}+\sqrt{\omega''(\theta_1)}}\right)^2 \tag{1.12}$$

which is ∞ if $\theta=\theta_{01}$, and >1 in the intervals of θ in which

$$\left|\omega'(\theta)-\frac{\omega(\theta_1)-\omega(\theta_0)}{\theta_1-\theta_0}\right|<\frac{1}{2(\theta_1-\theta_0)}\left(\frac{\omega'(\theta_1)-\omega'(\theta_0)}{\sqrt{\omega''(\theta_0)}+\sqrt{\omega''(\theta_1)}}\right)^2.$$

Another limiting value of the relative efficiency is obtained if we fix α_0 and α_1 and let $\theta\to\theta_{01}$. By applying l'Hospital's rule we can find

$$\frac{\partial L_h}{\partial h}\equiv\frac{\partial}{\partial h}\left(\frac{A^h-1}{A^h-B^h}\right)\xrightarrow[(h\to 0)]{}\frac{(-\log B)(\log A)}{2(\log A-\log B)},$$

and from (1.8) and (1.9) we have

$$h'(\theta_{01})=\frac{-2}{(\theta_1-\theta_0)}.$$

Since the first factor in the right hand side of (1.6) approaches to

$$\frac{h'(\theta_{01})(\log B-\log A)}{(\theta_1-\theta_0)\omega''(\theta_{01})}\left[\frac{\partial}{\partial h}L_h\right]_{h=0}$$

as $\theta\to\theta_{01}$, we thus have

$$\mathrm{RE}\,(\alpha_0, \alpha_1; \theta_{01})=\left(\frac{\omega'(\theta_1)-\omega'(\theta_0)}{(\theta_1-\theta_0)\sqrt{\omega''(\theta_{01})}}\right)^2\cdot\frac{(-\log B)(\log A)}{(\lambda_{\alpha_0}\sqrt{\omega''(\theta_0)}+\lambda_{\alpha_1}\sqrt{\omega''(\theta_1)})^2}, \tag{1.13}$$

where $A=(1-\alpha_1)/\alpha_0$ and $B=\alpha_1/(1-\alpha_0)$.

If $\theta_1=\theta_0+\Delta$ and $\Delta\to 0$, then, from (1.10), θ_{01} also tends to θ_0. Therefore from (1.13)

$$\lim_{\theta_1\to\theta_0}\mathrm{RE}\,(\alpha_0, \alpha_1; \theta_{01})=\frac{(-\log B)(\log A)}{(\lambda_{\alpha_0}+\lambda_{\alpha_1})^2}, \tag{1.14}$$

independently of both $\omega(\theta)$ and θ_0.

Equations (6), (7) and (13) all for normal densities, were derived by Bechhofer [1]. Paulson [5] showed that, for any densities $f_i(x)$ $(i=0, 1)$, and θ_0,

$$\lim_{\theta_1\to\theta_0}\mathrm{RE}\,(\alpha_0, \alpha_1; H_i)=\frac{\alpha_i\log\dfrac{\alpha_i}{1-\alpha_{1-i}}+(1-\alpha_i)\log\dfrac{1-\alpha_i}{\alpha_{1-i}}}{(1/2)(\lambda_{\alpha_0}+\lambda_{\alpha_1})^2} \qquad (i=0, 1)$$

independently of $f_i(x)$'s and θ_0, provided that conditions are assumed under which the central limit theorm for

$$\sum_{j=1}^{N} \log \frac{f_1(x_j)}{f_0(x_j)}$$

is applicable. These limits will not exceed 1, as are suggested by the optimality of the Wald SPRT (Wald and Wolfowitz [7]). This is proved as follows: let

$$f(t)=\frac{1}{\sqrt{2\pi}}e^{-(t-u)^2/2} \qquad \text{and} \qquad g(t)=\frac{1}{\sqrt{2\pi}}e^{-(t+v)^2/2}.$$

Then

$$I(f:g)\equiv\int_{-\infty}^{\infty} f(t)\log\frac{f(t)}{g(t)}dt=\frac{1}{2}(u+v)^2,$$

which combined with the convex property

$$\begin{aligned} I(f:g)&=\left(\int_{-\infty}^{0}+\int_{0}^{\infty}\right)f(t)\log\frac{f(t)}{g(t)}dt \\ &\geqq\left(\int_{-\infty}^{0}f(t)dt\right)\log\frac{\int_{-\infty}^{0}f(t)dt}{\int_{-\infty}^{0}g(t)dt}+\left(\int_{0}^{\infty}f(t)dt\right)\log\frac{\int_{0}^{\infty}f(t)dt}{\int_{0}^{\infty}g(t)dt}. \end{aligned}$$

gives

$$\Phi(u)\log\frac{\Phi(u)}{1-\Phi(v)}+(1-\Phi(u))\log\frac{1-\Phi(u)}{\Phi(v)}\leqq\frac{1}{2}(u+v)^2.$$

Replacing u and v with λ_{α_0} and λ_{α_1}, respectively, we get the inequality

$$\alpha_0\log\frac{\alpha_0}{1-\alpha_1}+(1-\alpha_0)\log\frac{1-\alpha_0}{\alpha_1}\leqq\frac{1}{2}(\lambda_{\alpha_0}+\lambda_{\alpha_1})^2.$$

EXAMPLES:

In the following examples

(a)$=\theta_{01}$,

(b)=limiting relative efficiency (1. 12), and

(c)=RE $(\alpha_0, \alpha_1; \theta_{01})$ divided by $(-\log B)(\log A)$.

(i) Normal distribution with known σ^2:

$$f_\theta(x)=(2\pi\sigma^2)^{-1/2}\exp\left[-\frac{(x-\mu)^2}{2}\right], \qquad \theta=\frac{\mu}{\sigma^2}, \qquad \omega(\theta)=\frac{\sigma^2\theta^2}{2}$$

(a)$=(\theta_0+\theta_1)/2$,

(b)$=\dfrac{(\theta_1-\theta_0)/8}{|\theta-(\theta_0+\theta_1)/2|}$, (c)$=(\lambda_{\alpha_0}+\lambda_{\alpha_1})^{-2}$.

(ii) Bernoulli distribution:

$$f_\theta(x)=p^x(1-p)^{1-x}, \qquad \theta=\log\frac{p}{1-p}, \qquad \omega(\theta)=\log(1+e^\theta),$$

(a)=the unique root of the equation

$$\frac{e^\theta}{1+e^\theta}=\frac{1}{\theta_1-\theta_0}\log\frac{1+e^{\theta_1}}{1+e^{\theta_0}},$$

or equivalently,

$$P_{01}=\left(\log\frac{1-P_0}{1-P_1}\right)\Big/\left(\log\frac{P_1}{P_0}+\log\frac{1-P_0}{1-P_1}\right),$$

where

$$\theta_i=\log\frac{P_i}{1-P_i} \qquad (i=0,1,01).$$

$$\text{(b)}=\frac{1}{2}\left(\frac{P_1-P_0}{\sqrt{P_0(1-P_0)}+\sqrt{P_1(1-P_1)}}\right)^2\Big/\left|P\log\frac{P_1}{P_0}-(1-P)\log\frac{1-P_0}{1-P_1}\right|,$$

$$\text{(c)}=\frac{(P_1-P_0)^2}{\left(\log\frac{P_1}{P_0}\right)\left(\log\frac{1-P_0}{1-P_1}\right)}\left(\sqrt{P_0(1-P_0)}\,\lambda_{\alpha_0}+\sqrt{P_1(1-P_1)}\,\lambda_{\alpha_1}\right)^{-2}.$$

(iii) Poisson distribution:

$$f_\theta(x)=m^x e^{-\lambda}/(x!), \qquad \theta=\log m, \qquad \omega(\theta)=e^\theta,$$

$$\text{(a)}=\log\left(\frac{e^{\theta_1}-e^{\theta_0}}{\theta_1-\theta_0}\right),$$

$$\text{(b)}=\frac{1}{2}(\sqrt{m_1}-\sqrt{m_0})^2\Big/\left|m\log\frac{m_1}{m_0}-(m_1-m_0)\right|,$$

$$\text{(c)}=\frac{m_1-m_0}{\log(m_1/m_0)}\cdot\left(\sqrt{m_0}\,\lambda_{\alpha_0}+\sqrt{m_1}\,\lambda_{\alpha_1}\right)^{-2}.$$

(iv) Geometrical distribution:

$$f_\theta(x)=\left(\frac{\mu}{1+\mu}\right)^x\cdot\frac{1}{1+\mu}, \qquad \theta=\log\frac{\mu}{1+\mu}, \qquad \omega(\theta)=-\log(1-e^\theta),$$

(a)=the unique root of the equation

$$\frac{e^\theta}{1-e^\theta}=\frac{1}{\theta_1-\theta_0}\log\frac{1-e^{\theta_0}}{1-e^{\theta_1}},$$

or equivalently,

$$\mu_{01}=\left(\log\frac{1+\mu_1}{1+\mu_0}\right)\Big/\left(\log\frac{\mu_1}{\mu_0}-\log\frac{1+\mu_1}{1+\mu_0}\right),$$

where

$$\theta_i=\log\frac{\mu_i}{1+\mu_i}\qquad(i=0,1,01)$$

$$(\mathrm{b})=\frac{1}{2}\left(\frac{(\mu_1-\mu_0)^2}{\sqrt{\mu_0(1+\mu_0)}+\sqrt{\mu_1(1+\mu_1)}}\right)^2\Big/\left|\mu\log\frac{\mu_1}{\mu_0}-(1+\mu)\log\frac{1+\mu_1}{1+\mu_0}\right|,$$

$$(\mathrm{c})=\frac{(\mu_1-\mu_0)^2}{\left(\log\frac{\mu_1}{\mu_0}\right)\left(\log\frac{1+\mu_1}{1+\mu_0}\right)}\left(\sqrt{\mu_0(1+\mu_0)}\,\lambda_{\alpha_0}+\sqrt{\mu_1(1+\mu_1)}\,\lambda_{\alpha_1}\right)^{-2}.$$

(v) Exponential distribution:

$$f_\theta(x)=\mu^{-1}e^{-x/\mu},\qquad \theta=-\mu^{-1},\qquad \omega(\theta)=-\log(-\theta),$$

$$(\mathrm{a})=\left\{(\theta_1-\theta_0)\log\frac{\theta_1}{\theta_0}\right\}^{-1},$$

$$(\mathrm{b})=\frac{1}{2}\left(\frac{\mu_1-\mu_0}{\mu_0+\mu_1}\right)^2\Big/\left|\left(\frac{1}{\mu_0}-\frac{1}{\mu_1}\right)\mu-\log\frac{\mu_1}{\mu_0}\right|,$$

$$(\mathrm{c})=\left\{(\mu_1-\mu_0)\log\frac{\mu_1}{\mu_0}\right\}^{-2}(\mu_0\lambda_{\alpha_0}+\mu_1\lambda_{\alpha_1})^{-2}.$$

§ 2. Relative efficiencies and the Chernoff information number.

Equations (1. 7) and (1. 11) show that

$$\lim_{\alpha_0\to 0}\mathrm{RE}\,(\alpha_0,a\alpha_0^b;\theta)$$

is greater than 1 in an interval of θ around θ_{01} and tends to ∞ as $\theta\to\theta_{01}$. The Wald SPRT for discriminating between $f_{\theta_0}(x)$ and $f_{\theta_1}(x)$, both in the exponential family of densities with the same $\omega(\theta)$, is extremely inefficient, in the above sense, if the unknown true parameter-value is $\theta=\theta_{01}$. This value θ_{01} is closely connected with the information numbers of Kullback-Leibler (Kullback [4]), and of Chernoff [2]. The Chernoff information number for descriminating between two densities $f_0(x)$ and $f_1(x)$ (relative to the dominating measure ν) is defined by

$$-\log\left(\inf_{0<\tau<1}\int[f_1(x)]^\tau[f_0(x)]^{1-\tau}d\nu\right),\tag{2. 1}$$

and measures how difficult it is to decide between f_0 and f_1 with the Bayes test (Chernoff [2], Joshi [3]). Now Theorem 2. 1 in Kullback's book [4] gives, as its restatement, THEOREM. *Let $f_0(x)$ and $f_1(x)$ be given generalized probability densities. Let*

$$G \equiv \left\{ g(x) \,\middle|\, g(x) \geqq 0[\nu], \ \int g(x) d\nu = 1, \quad and \quad I(g:f_1) = I(g:f_0) \right\}.$$

Then we have

$$\min_{g \in G} I(g:f_1) = -\log F(\tau^*), \tag{2. 2}$$

where

$$F(\tau) \equiv \int [f_1(x)]^{\tau} [f_0(x)]^{1-\tau} d\nu$$

and τ^* is determined by the equation

$$F'(\tau^*) = \int f_1^{\tau^*} f_0^{1-\tau^*} \log \frac{f_1}{f_0} d\nu = 0.$$

The minimizing $g(x)$ is given by

$$g(x) = g^*(x) = [f_1(x)]^{\tau^*} [f_0(x)]^{1-\tau^*} / F(\tau^*). \tag{2. 3}$$

We see that the minimum pseudo-distance $-\log F(\tau^*)$ is the Chernoff information number (2. 1).

If $f_i(x) = f_{\theta_i}(x) \equiv e^{\theta_i x - \omega(\theta_i)}$ $(i=0, 1)$, i.e., belong to the exponential family of densities, then straight-forward calculation gives

$$g^*(x) = f_{\theta_{01}}(x) = e^{\theta_{01} x - \omega(\theta_{01})} \tag{2. 4}$$

where θ_{01} is uniquely determined by the equation (1. 10). Moreover, for the τ^* in (2. 3), we have

$$\theta_{01} = \tau^* \theta_1 + (1 - \tau^*) \theta_0, \qquad 0 < \tau^* < 1.$$

Thus for the densities of exponential family, (2. 2), together with (2. 4), gives the Chernoff inforffiation number equal to

$$\frac{\theta_{01} - \theta_0}{\theta_1 - \theta_0} \omega(\theta_1) + \frac{\theta_1 - \theta_{01}}{\theta_1 - \theta_0} \omega(\theta_0) - \omega(_{01}).$$

This number measures how convex the $\omega(\theta)$ is at $\theta = \theta_{01}$, just like the factor $(\omega''(\theta_{01}))^{-1}$ in the relative efficiency (1. 13).

REFERENCES

[1] BECHHOFER, R., A note on the relative efficiency of the Wald sequential probability ratio test. Journ. Amer. Stat. Assoc., **55** (1960), 660–663.

[2] CHERNOFF, H., A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. Ann. Math. Stat. **23** (1952), 492–505.

[3] JOSHI, D. D., L'information en statistique mathèmatique et dans la théorie des

communications. Thèse, Faculté des Sciences de l'Université de Paris (June 1957).

[4] KULLBACK, S., Information Theory and Statistics. John Wiley, New York (1959).

[5] PAULSON, E., A note on the efficiency of the Wald sequential test. Ann. Math. Stat. **18** (1947), 447–450.

[6] WALD, A., Sequential Analysis. John Wiley, New York (1947).

[7] WALD, A., AND J. WOLFOWITZ, Optimum character of the sequential probability ratio test. Ann. Math. Stat. **19** (1948), 326–339.

DEPARTMENT OF APPLIED MATHEMATICS,
OSAKA UNIVERSITY.