韓國數學教育學會誌〈數學教育〉
1987. 5, Vol. XXV, No. 2

Journal of the Korea Society of Mathematical Education
May. 1987, Vol. XXV, No. 2

An Extended Large Deviation Theorem for Empirical Distributions

by Yong Kab Choi

Gyeongsang National University, Korea

I. Introduction

Fu(1985) states as follows : if F_n is an empirical distribution resulting from a sample of independent observations from a common population F_o and if A is an well-defined subset of probability distributions which does not contain F_o, then

$$lim_{n\to\infty} \frac{1}{n} \log P(F_n \in A \mid F_o) = -K(A, F_o),$$

where

$$K(A, F_o) = \begin{cases} inf\{K(F, F_o) : F \in A\}, & if\ A \neq \phi \\ \infty, & if\ A = \phi, \end{cases}$$

and $K(F, F_o)$ is a Kullback-Leibler information number of F with respect to F_o. He proves simply this result by a new method of K-regular technique.

In this paper, we shall generalize this Fu's result by using some theorems in [1] - [6].

II. Preliminaries

Let X be a sample space of points X, and (X, β) be a measurable space. Let p and q be probability measures on β. If $q \ll p$ on β, that is, q is absolutely continuous with respect to p on β, let $r(x)$ be a density, $0 \leqq r(x) < \infty$, i.e. $dp = r(x)dp$ on β. We define a Kullback-Leibler information number of q with respect to p by

$$K(q, p) = \begin{cases} \int \log r(x)\ dq & if\ q \ll p \\ \infty & otherwise. \end{cases}$$

If $q \ll p$ and $p \ll q$, then $dq = \frac{1}{r(x)} dq$ on β, where $\frac{1}{r(x)}$ is a density. Hence we define a Kullback-Leibler information number of p with respect to q by

$$K(p, q) = \begin{cases} \int \log \frac{1}{r(x)} dp & if\ p \ll q \\ \infty & otherwise. \end{cases}$$

We note that $K(p, q)$ and $K(q, p)$ are well-defined.

Lemma 2.1 [1], [6] *Let* $q \ll p$ *and* $p \ll q$. *For real* $t \in [0, 1]$, *let* $f(t) = \int exp(t$

log $r(x))$ dp. *Then we have*

(i) $f(0)=f(1)=1$,

(ii) $f(t)$ *is a continuous and strictly convex function* $[0,1]$,

(iii) $f'(0^+)=-K(p,q)$ *and* $f'(1^-)=K(q,p)$, *where* $f'(0^+)$ *and* $f'(1^-)$ *are right and left hand limits of* $f'(t)$ *at* $t=0$, $t=1$, *respectively.*

(iv) *For all* $t\in[0,1]$, $f(t)<\infty$ *iff* $K(q,p)<\infty$ *and* $K(p,q)<\infty$.

Now, suppose that $Y=Y(x)$ is a real-valued measurable function on (X,β). Let the moment generating function of $Y(x)$ be

$$\phi(t)=\int exp(t\,Y(x))\ dp,\quad t\in[0,1], \tag{2,1}$$

and let for real a

$$I(a)=inf\{\phi(t)\,exp(-ta):\ t\in[0,1]\}.$$

Define a set A by

$$A=\{q:\ \int Y(x)\ dq\ \text{exists and}\ \geqq a\},$$

and let

$$K(A,p)=\begin{cases} inf\{K(q,p):\ q\in A\} & \text{if}\ A=\phi \\ \infty & \text{if}\ A=\phi,\end{cases}$$

where $K(q,p)$ is a Kullback-Leibler information number of q with respect to p.

Lemma 2.2 [3], [4], [5] *For real* a, $I(a)=exp(-K(A,p))$.

Ⅲ. Main Results

Let Λ be a family of all probability distributions defined on the real line X, and $X_1,\ldots,X_n$ be a sequence of identically independent distributed (i.i.d.) random variables from the distribution function $F\in\Lambda$. Let F_n be an empirical distribution generated by these observations.

Definition 3.1 [2] Let A be a subset of Λ and $F_o\notin A$. The subset A is said to be K-regular with respect to F_o if there exists an $F^*\in A$ such that $F^*\ll F_o$ and $A\subset H^{\oplus}(F^*,F_o)$,

where

$$H^{\oplus}(F^*,F_o)=\{F\in\Lambda:\text{and}\ \int log\frac{dF^*}{dF_o}\ dF\geqq K(F^*,F_o)\},$$

$$K(F^*,F_o)=\int log\frac{dF^*}{dF_o}\ dF^*$$

and $\dfrac{dF^*}{dF_o}$ is the Radon-Nikodym derivative of F^* with respect to F_o.

Theorem 3.2 [2] *Let* $F_o\in\Lambda$. *If* A *is a subset with nonempty interior which does not contain* F_o *and satisfies conditions:*

(i) A *is* K*-regular with respect to* F_o,

(ii) $K(A,F_o)=K(int(A),F_o)$, where $int(A)$ *stands for the interior of* A, *then*

$$lim_{n\to\infty}\ \frac{1}{n}\ log\ P(F_n\in A\,|\,F_o)=-K(A,F_o).$$

Let $T:\ \Lambda\to R$ be a functional and F_n be an empirical distribution on (X,β).

Let $U(\varepsilon)=\{F\in\Lambda: T(F)$ exists and $\geqq\varepsilon\}$ for given real ε.

Then we can obtain a generalized theorem (Theorem 3.3).

Theorem 3.3 *If $U(\varepsilon)$ is a subset with nonempty interior which does not contain F_o and $K(U(\varepsilon),F_o)=K(int(U(\varepsilon)),F_o)$, then for given real ε*

$$lim_{n\to\infty}\ \frac{1}{n}\ log\ P(T(F_n)\geqq\varepsilon\ |F_o)=-K(U(\varepsilon),F_o).$$

Proof For $F^*,F_o\in\Lambda$, let $F^*\ll F_o$ and put $Y(x)=log\ \dfrac{dF^*(x)}{dF_o(x)}$.

To prove that $U(\varepsilon)$ is $K-$regular $w.r.t.$ F_o, we take for $F\in U(\varepsilon)$

$$T(F)=\int Y(x)\ dF.$$

Then $F\in U(\varepsilon)$ iff

$$T(F)=\int log\ \frac{dF^*}{dF_o}\ dF\geqq\varepsilon.$$

From Lemma 2.2, we have

$$K(U(\varepsilon),F_o)=-log\ I(\varepsilon). \tag{3.1}$$

According to the definition of (2.1), let

$$\phi(t)=\int exp(t\ log\frac{dF^*}{dF_o})\ dF_o,\ \ 0\leqq t\leqq 1.$$

Then by Lemma 2.1, $\phi(t)<\infty$ for all $t\in[0,1]$. Since

$$\int exp(t(log\frac{dF^*}{dF_o}-\varepsilon))\ dF_o$$

is a strictly convex function of t on $[0,1]$, there exists $h>0$ such that

$$\begin{aligned}I(\varepsilon)&=inf_{0\leq t\leq 1}\phi(t)\,exp(-t\ \varepsilon)\\&=\phi(h)\,exp(-h\ \varepsilon),\ \text{say}.\end{aligned} \tag{3.2}$$

This $h>0$ is a unique number so that

$$\phi'(h)\,exp(-h\ \varepsilon)-\varepsilon\phi(h)\,exp(-h\ \varepsilon)=0,$$

that is,

$$\phi'(h)/\phi(h)=\varepsilon. \tag{3.3}$$

Thus $I(\varepsilon)$ satisfies $\phi'(h)/\phi(h)=\varepsilon$ for some $h>0$. Put

$$dF^*=\frac{exp(h\ log\dfrac{dF^*}{dF_o})}{\phi(h)}\ dF_o. \tag{3.4}$$

Then the quantity $exp(hlog\ \dfrac{dF^*}{dF_o})/\phi(h)$ represents a density $w.r.t.$ F_o.

By our definition of $T(F)$, and from (3.3) and (3.4), we have for $F^*\in\Lambda$

$$\begin{aligned}T(F^*)&=\int log\frac{dF^*}{dF_o}dF^*\\&=\int(log\ \frac{dF^*}{dF_o})\ \frac{exp(h\ log\dfrac{dF^*}{dF_o})}{\phi(h)}dF_o\end{aligned}$$

$$= \frac{\phi^1(h)}{\phi(h)} = \varepsilon. \tag{3.5}$$

Thus $F^* \in U(\varepsilon)$ and $\int log \frac{dF^*}{dF_o} dF \geqq K(F^*, F_o)$.

Therefore, $U(\varepsilon)$ is K-regular *w. r. t.* F_o. From Theorem 3.2, we have

$$lim_{n\to\infty} \frac{1}{n} \; log \; P(T(F_n) \geqq \varepsilon \mid F_o)$$

$$= lim_{n\to\infty} \frac{1}{n} \; log \; P(F_n \in U(\varepsilon) \mid F_o)$$

$$= -K(U(\varepsilon), \; F_o).$$

Thus, the proof is complete.

Remark 3.4 *From* (3.2), (3.4) *and* (3.5),

$$K(F^*, F_o) = \int log \frac{dF^*}{dF_o} dF^*$$

$$= \int log \frac{exp(h \; log \frac{dF^*}{dF_o})}{\phi(h)} dF^*$$

$$= \int (h \; log \frac{dF^*}{dF_o} - log \; \phi(h)) \; dF^*$$

$$= h \; \varepsilon - log \; \phi(h)$$

$$= - log \; I(\varepsilon).$$

Thus from (3.1)

$$K(U(\varepsilon), \; F_o) = - log \; I(\varepsilon) = K(F^*, \; F_o).$$

References

1. Bahadur, R. R.: Some limit theorems in statistics. Regional conference series in applied mathematics, *Society for industrial and applied mathematics,* Philadelphia (1971).
2. Fu, J. C.: On probabilities of large deviations for empirical distributions. *Statistic & Decisions 3,* 115-142 (1985).
3. Groeneboom, P., Oosterhoff, J., Ruymgaart, F. H.: Large deviation theorems for empirical probability measures. *Ann. Prob. Vol. 7,* 553-586, (1979).
4. Hoadley, A. B.: On the probability of large deviations of functions of several empiric cdf's. *Ann. Math. Statist.* Vol. 38, 360-381, (1967).
5. Hoeffding, W.: On probabilities of large deviations, Proc. Fifth Berkeley Symp. *Math. Statist. Prob. 1,* 203-219, (1965).
6. Steinebach, J.: Large deviation probabilities and some related topics, *Carleton Mathematical Lecture Notes No. 28,* (1980).