Journal of the Korean
Statistical Society
Vol. 19, No. 2, 1990

On Large Deviation of the Sample Medians

Chong－Sun Hong*

ABSTRACT

Consider the following problem in the large deviation theory. For constants $a_1, \cdots, a_p$ the tail probability $P(M_1 > a_1, \cdots, M_p > a_p)$ of the sample medians $(M_1, \cdots, M_p)$ is supposed to converge to zero as sample size increases. This paper shows that this probability converges to zero exponentially fast and estimates the convergence rates of the above tail probability of the sample medians. Also compare with the rates about the sample means.

1. Introduction

Suppose that $\{\underline{X}n\} = \{(\underline{X}_{n1}, \cdots, \underline{X}_{np})'\}$ is a sequence of p-variate iid random vectors. Let $\underline{a} = (a_1, \cdots, a_p)'$ be fixed in $\mathbf{R}^p$. For each $n = 1, 2, \cdots$ let

$$P_n = P\left(\frac{\sum_{i=1}^{n} X_{i1}}{n} > a_1, \cdots, \frac{\sum_{i=1}^{n} X_{ip}}{n} \geq a_p\right)$$
$$= P(\overline{X}_1 > a_1, \cdots, \overline{X}_p > a_p). \qquad (1)$$

It is assumed that the given $\underline{a}$ satisfies the condition that $P_n > 0$ for each n, and $P_n \to 0$ as $n \to \infty$. And assume that the random vector has the moment generating function $\phi(\underline{t})$ of $\underline{X}_n$, i.e.,

$$\phi(\underline{t}) = E[exp(\underline{t}'\underline{X}_n)].$$

Define

$$\psi(\underline{t}) = exp\ (-\underline{a}'\underline{t})\ \phi(\underline{t}). \qquad (2)$$

Let T denote the set of all values $\underline{t}$ for which $\phi(\underline{t}) < \infty$. We suppose that the T is non-degenerate, and that there exists positive $\underline{\tau} = (\tau_1, \cdots, \tau_p)'$, $\tau_j > 0$ for any j, in the interior of T such that

$$\psi(\underline{\tau}) = \inf_{\underline{t}} \{\psi(\underline{t})\} = \rho(say). \qquad (3)$$

Then we have

$$n^{-1} \log P_n = \log \rho + o(1). \qquad (4)$$

In univariate case($p=1$), estimates ρ in (4) were obtained by Cramér(1938), Blackwell and Hodges (1959), Chernoff(1952), Bahadur and Rao(1960) and Bahadur(1960). The estimating method given by Bahadur and Rao(1960) and Bahadur(1960) is essentially a variant or extention of Cramér's work.

* Assistant Professor, Department of Statistics, Sung Kyun Kwan University, 3-53 Myung Ryun-dong, Chongro-ku

Blackwell and Hodges(1959) considered the case that X is a lattice variable. And Chernoff(1952) considered the case that X is a discrete random variable. Bahadur and Rao(1960) proved in the general case, that is, discrete and continuous as well, and Bahadur(1960) refined the Bahadur and Rao(1960)'s work. Sievers(1969, 1975) generalized the previous large deviations and considered large deviation of the sample median. Steinebach(1978) extended to the multi-dimensional case. And Bahadur and Zabell(1979) extended large deviations of the sample mean to general vector space. Hong and David (1988) obtained, among other matters, an alternative of the estimate of large deviation of the sample median. Moreover, Hong(1989) extended the case of weighted averages to the regression model, and those which are not necessarily independent and identically distributed.

Now this paper extends Sievers(1969) and Hong and David(1988)'s works to p sample medians, and obtains the convergence rates of large deviations $P(M_1>a_1, \cdots, M_p>a_p)$ of the sample medians via the theory of Bahadur(1960, 1971) and Bahadur and Rao(1960). In Section 3, the rates of the sample medians are compared with those of the sample means over Normal distributions. The standard bivariate Normal distribution is used to calculate the convergence rates of the deviations.

2. Deviation

It is assumed without any loss of the generality through out the paper that $E(\underline{X}_n)=\underline{0}$, and $\underline{a}>\underline{0}$, where $\underline{0}$ is the null vector. This satisfies the condition that $P_n>0$.

Let M be the sample median and a be a fixed constant. Large deviation $P(M>a)$ of the sample median was considered by Sievers(1969), and Hong and David (1988) gave a version of it on the way to consider the asymptotic variance of the rounded sample median.

We here consider the case $P(M_1\geq a_1, \cdots, M_p\geq a_p)$, where $\underline{M}=(M_1, \cdots, M_p)'$ is a vector of the sample medians from $\{\underline{X}_n=(\underline{X}_{n1}, \cdots, \underline{X}_{np})'\}$. Without loss of generality the sample size n is an odd number, $n=2\upsilon+1$.

Let $\underline{Y}_n=(\underline{Y}_{n1}, \cdots, \underline{Y}_{np})'$ be a sequence of the indicator vector, in which $\underline{Y}_{nj}$ denotes the indicator vector of the event $X_{nj}\geq a_j$, then with the argument similar to that in Sievers(1969) and Hong and David(1988), we express as the followings:

$$P(M_1\geq a_1, \cdots, M_p\geq a_p) \tag{5}$$
$$=P\ (\Sigma^{2\upsilon+1}{}_{i=1}\ Y_{i1}>\upsilon+1, \cdots,\ \Sigma^{2\upsilon+1}{}_{i=1}\ Y_{ip}>\upsilon+1)$$
$$=P(\overline{Y}_1\geq 1/2,\ \cdots,\ \overline{Y}_p\geq 1/2), \tag{6}$$

where the second equality is due to the fact that $\Sigma^{2\upsilon+1}{}_{i=1}\ Y_1$ must be an integer. Note that it satisfies the underlying condition that the above probability is positive and converges to zero as n increases.

The previous works about large deviations reviewed in Section 1 were carried into execution for a sequence of random variables which have moment generating functions. The large deviation (5) of the sample medians, however, does not require the existence condition of moment generating function, because any sample medians can be expressed in terms of Bernoulli random variables.

Then the moment generating function of $\underline{Y}_n$ can be obtained

$$\phi(\underline{t})=P^0+\sum_{i=1}^{p} P^1{}_i\ exp(t_i)+\sum_{i=1}^{p}\sum_{j=1(j>i)}^{p} P^2{}_{ij}\ exp(t_i+t_j)+\cdots$$
$$+P^p\ exp(\Sigma^p{}_{i=1}\ t_i),$$

where

$$P^0=P(X_{ni}<a_i \text{ for all } i),$$
$$P^1{}_i=P(X_{ni}\leq a_i,\ X_{nj}<a_j \text{ for all } j(j\neq i)),$$
$$P^2{}_{ij}=P(X_{ni}\geq a_i,\ X_{nj}\geq a_j,\ X_{nk}<a_k \text{ for all } k(k\neq i,\ k\neq j)),$$
$$\vdots$$

And $\psi(\underline{t})$ defined in (2) turns out to equal

$$\psi(\underline{t})=exp(-1/2\ \underline{1}'\underline{t})\ \phi(\underline{t}),$$

where $\underline{1}$ is the $p\times1$ unit vector. Therefore we state

Assertion

$$\lim_{n\to\infty} n^{-1} \log P(M_1\underline{>}a_1,\cdots,\ M_p\underline{>}a_p)=\log\ \rho,$$

where

$$\rho=\inf_{\underline{t}} exp(-1/2\ \underline{1}'\underline{t})(P^0+\sum_{i=1}^{n} P^1{}_i\ exp(t_i)+\sum_{i=1}^{p}\sum_{j=1(j>i)}^{p} P^2{}_{ij}\ exp(t_i+t_j)$$
$$+\cdots P^p\ exp(\sum_{i=1}^{p} t_i)).$$

The estimate ρ in the bivariate case($p=2$) will be examined. We have

$$\psi(t_1,\ t_2)=exp(-(t_1+t_2)/2)\ (P^0+P^1{}_1\ exp(t_1)$$
$$+P^1{}_2\ exp(t_2)+P^2\ exp(t_1+t_2)).$$

Then solve

$$\frac{d}{dt_1}\ \psi(t_1,\ t_2)=0 \quad \text{and} \quad \frac{d}{dt_2}\ \psi(t_1,\ t_2)=0,$$

so that one gets

$$\tau_1=\frac{1}{2}\log\frac{P^0\ P^1{}_2}{P^1{}_1P^2} \quad \text{and} \quad \tau_2=\frac{1}{2}\log\frac{P^0\ P^1{}_1}{P^1{}_2\ P^2}$$

Thus,

$$\psi(t_1,\ t_2)=2\{[P^0P^2]^{1/2}+[P^1{}_1\ P^1{}_2]^{1/2}\}.$$

Therefore, we obtain in the form (4)

$$\lim_{n\to\infty}\quad n^{-1}\log P(M_1\underline{>}a_1,\ M_2\underline{>}a_2)$$
$$=\log\ 2\{(P^0P^2)^{1/2}+(P^1{}_1\ P^1{}_2)^{1/2}\}$$
$$=\log\ 2\{[P(X_1<a_1,\ X_2<a_2)\ P(X_1\underline{>}a_1,\ X_2\underline{>}a_2)]^{1/2}$$
$$+[P(X_1<a_1,\ X_2\underline{>}a_2)\ P(X_1\underline{>}a_1,\ X_2<a_2)]^{1/2}\}. \tag{7}$$

For a general case($p\underline{>}3$), the estimate ρ defined in Assertion does exist, but it is not easy to formulate it. Futher works will be remained to readers.

3. Comparison

Let $\underline{X}_n$ be the p-variate Normal random vector with mean vector $\underline{0}$ and variance-covariance matrix $V=\begin{bmatrix}1 & r\\ r & 1\end{bmatrix}$. Define $\Phi(\cdot)$ is the standard Normal distribution function and $\Phi(\cdot,\ \cdot)$ is the bivariate standard Normal distribution function described above. Since

$$P^0=\Phi(a_1,\ a_2),\qquad P^1{}_2=\Phi(a_1)-\Phi(a_1,\ a_2),$$
$$P^1{}_1=\Phi(a_2)-\Phi(a_1,\ a_2),\ \text{and}\quad P^2=1-\Phi(a_1)-\Phi(a_2)+\Phi(a_1,\ a_2).$$

$$\lim_{n\to\infty} n^{-1} \log P(M_1 \geq a_1,\ M_2 \geq a_2)$$
$$= \log 2\{(\Phi(a_1,\ a_2)\,[1-\Phi(a_1)-\Phi(a_2)+\Phi(a_1,\ a_2)])^{1/2}$$
$$+([\Phi(a_1)-\Phi(a_1,\ a_2)]\,[\Phi(a_2)-\Phi(a_1,\ a_2)])^{1/2}\}. \qquad (8)$$

Now consider large deviation of the sample means for the standard bivariate Normal distribution. Then it is well known that

$$\lim_{n\to\infty} n^{-1} \log P(\bar{X}_1 \geq a_1,\ \bar{X}_2 \geq a_2) = -1/2\ \underline{a}'\,\underline{V}^{-1}\underline{a}$$
$$= -\frac{a_1\ a_2}{1+r}. \qquad (9)$$

Table 1 below shows that the deviation rates of the sample means and the sample medians over the variation of $a_1 = a_2 = a$ and the correlation r. In such case, formulation (8) and (9) will be

$$\log 2\{[\Phi(a,\ a)\ (1-2\Phi(a)+\Phi(a,\ a))]^{1/2}+(\Phi(a)-\Phi(a,\ a))\},$$

and $-a^2(1+r)$, respectively. The values of $\Phi(\cdot)$ *and* $\Phi(\cdot,\ \cdot)$ were extracted from Pearson(1931).

Table 1. Convergence Rates

A : negative correlation

a \ r		−0.8	−0.6	−0.4	−0.2	−0.1
0.5	$\bar{X}$	−1.250	−0.625	−0.417	−0.313	−0.278
	M	−0.351	−0.261	−0.213	−0.181	−0.169
1.0	$\bar{X}$	−5.000	−2.500	−1.667	−1.250	−1.111
	M	−1.110	−0.961	−0.822	−0.713	−0.668
2.0	$\bar{X}$	−20.000	−10.000	−6.667	−5.000	−4.444
	M	−3.090	−3.066	−2.931	−2.691	−2.555
3.0	$\bar{X}$	−45.000	−22.500	−15.000	−11.250	−10.000
	M	−5.915	−5.914	−5.883	−5.677	−5.475

B : positive correlation

a \ r		0.1	0.2	0.4	0.6	0.8
0.5	$\bar{X}$	−0.227	−0.208	−0.179	−0.156	−0.139
	M	−0.149	−0.141	−0.126	−0.113	−0.100
1.0	$\bar{X}$	−0.909	−0.833	−0.714	−0.625	−0.556
	M	−0.591	−0.558	−0.499	−0.446	−0.394
2.0	$\bar{X}$	−3.636	−3.333	−2.857	−2.500	−2.222
	M	−2.288	−2.161	−1.928	−1.716	−1.512
3.0	$\bar{X}$	−8.182	−7.500	−6.429	−5.625	−5.000
	M	−4.945	−4.663	−4.128	−3.649	−3.207

From the Table 1, it is seen that the deviation rates of the sample means are smaller than those of the sample medians. This reflects the fact that the tail probability of the sample medians is greater than that of the sample means in bivariate Normal distribution.

The ratios of deviation rates of the sample means versus the sample medians are obtained in Table converges to zero faster than that of the sample medians.

Table 2. Convergence Ratio

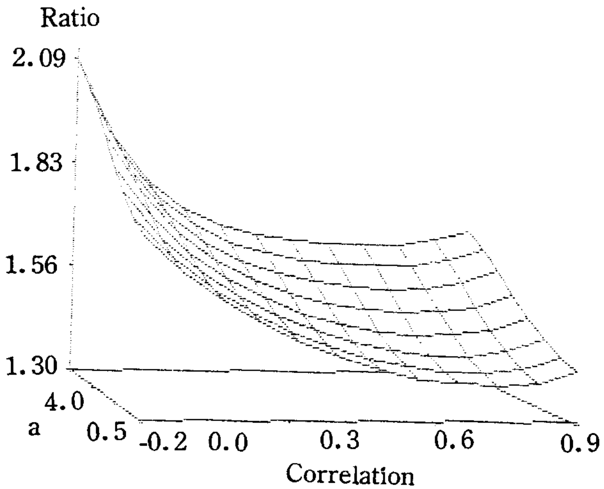
A : negative correlation

a \ r		−0.8	−0.6	−0.4	−0.2	−0.1
0.5	RATIO	3.561	2.391	1.957	1.723	1.642
1.0	RATIO	4.505	2.602	2.028	1.753	1.664
1.5	RATIO	5.598	2.917	2.139	1.800	1.698
2.0	RATIO	6.473	3.261	2.274	1.858	1.739
2.5	RATIO	7.121	3.564	2.417	1.920	1.784
3.0	RATIO	7.608	3.805	2.550	1.982	1.827
3.5	RATIO	7.983	3.991	2.665	2.039	1.866
4.0	RATIO	8.276	4.138	2.759	2.092	1.901

B : positive correlation

a \ r		0.1	0.2	0.4	0.6	0.8
0.5	RATIO	1.514	1.482	1.420	1.389	1.393
1.0	RATIO	1.538	1.494	1.433	1.402	1.409
1.5	RATIO	1.561	1.515	1.453	1.425	1.435
2.0	RATIO	1.590	1.542	1.482	1.457	1.470
2.5	RATIO	1.622	1.574	1.517	1.497	1.513
3.0	RATIO	1.654	1.608	1.557	1.542	1.559
3.5	RATIO	1.686	1.642	1.598	1.587	1.605
4.0	RATIO	1.714	1.674	1.637	1.628	1.644

And Table 2 tells us that, when the correlation r is fixed, the ratio increases as a increases. But for a given constant a, the ratio decreases as the correlation r moves from negative to positive values. However the ratio stops decreasing after $r=0.6$ and starts to increase slowly. It is an interesting point we might say that the function of the ratio is convex with respect to the correlation. The ratio function is plotted in Fig. 1 below.



In the case $r=0$, that is, $\underline{X}_{n1}$ and $\underline{X}_{n2}$ are independent, the convergence rates of the sample means and medians are summations of the convergence rates of the sample mean and median which are well known large deviation results. And the ratio in above case ($a_1=a_2=0$ and $r=0$) converges to $\pi/2=1,571$ as a decreases.

References

1. Bahadur, R.R.(1960). "On the Asymptotic Efficiency of Tests and Estimates," Sankhyá, 22, 229-252.
2. Bahadur, R.R.(1971). "Some Limit Theorems in Statistics," Regional Conference Series in Applied Mathematics Report 4, Society for Industrial and Applied Mathematics Philadelphia.
3. Bahadur, R.R. and Rao, R.R.(1960). "On Deviations of the Sample Mean," Annals of Mathematical Statistics, Vol. 31, 1015-1027.
4. Bahadur, R.R. and Zabell, S.L.(1979). "Large Deviation of the Sample Mean in General Vector Space," Annals of Probability, Vol. 7, 587-621.
5. Blackwell, D. and Hodges, J.L.(1959). "The Probability in the Extreme Tail of a Convolution," Annals of Mathematical Statistics, Vol. 30, 1113-1120.
6. Cramér, H.(1938). "Sur un Nouveau Théoréme—Limite de la Théorie des Probabilitiés, Actualités Scientifiques et Industrielles, No. 736, Hermann C^ie, Paris.
7. Chernoff, H.(1952). "A Measure of Asymptotic Efficiency for Tests of a Hypothesis Based on the Sum of Observation," Annals of Mathematical Statistics, Vol. 23, 493-507.
8. Gupta, S.S.(1963). "Bibliography on the Multivariate Normal Integrals and Related Topics," Annals of Mathematical Statistics, Vol. 34, 829-838.
9. Hong, C.S.(1989). "On Estimating Restricted Regression Parameters," The Korean Economic Review, Vol. 17, No. 1, 305-323.
10. Hong, C.S. and David, H.T.(1988). "Granularity and Efficiency," Journal of the American Statistical Association, Vol. 83, No. 403, 880-885.
11. Pearson, K.(1931). Tables for Statistics and Biometricians 2, Cambridge University Press.
12. Sievers, G.L.(1969). "On the Probability of Large Deviations and Exact Slpes," Annals of Mathematical Statistics, Vol. 40, 1908-1921.
13. Sievers, G.L.(1975). "Multivariate Probabilities of Large Deviations," Annals of Statistics, Vol. 3, 897-905.
14. Steinebach, J.(1978). "Convergence Rates of Large Deviation Probabilities in the Multidimensional Case," Annals of Probability, Vol. 6, 751-759.