The Korean Communications
in Statistics Vol. 5, No. 3, 1998
pp. 855-868

Distribution of the Estimator for Peak of a Regression Function Using the Concomitants of Extreme Order Statistics[1]

S. H. Kim[2] and T. S. Kim[3]

Abstract

For a random sample of size n from general linear model, $Y_i = \theta(X_i) + \varepsilon_i$, let $Y_{i:n}$ denote the ith order statistics of the Y sample values. The X-value associated with $Y_{i:n}$ is denoted by $X_{[i:n]}$ and is called the concomitant of ith order statistics. The estimator of the location of a maximum of a regression function, $\theta(x)$, was proposed by $\hat{x}(r) = \sum_{i=1}^{r} X_{[n-i+1]}/r$ and was found the convergence rate of it under certain weak assumptions on θ. We will discuss the asymptotic distributions of both $\theta(X_{[n-r+1]})$ and $\hat{x}(r)$ when r is fixed as $n \to \infty$(i.e. extreme case) on the basis of the theorem of the concomitants of order statistics. And we will investigate the asymptotic behavior of Max{ $\theta(X_{[n-r+1:n]})$, . , $\theta(X_{[n:n]})$} as an estimator for the peak of a regression function.

1. Introduction

Let $\theta(x_0) > \theta(x)$ for any $x \neq x_0$ in C, where θ is a real bounded continuous function defined on a bounded interval $C \subseteq R$. The aim is to find x_0 based on n samples (X_1, Y_1), $\cdots$, (X_n, Y_n) with $Y_i = \theta(X_i) + \varepsilon_i$. Here ε_1, . . ε_n are independent and identically distributed random variables with zero expectation.

In this paper, we study the distribution of the so-called best-r-points-average estimation for the location of a maximum of a regression function, used in Changchien(1990). This method has been used to search for the optimum range of burden distribution indices of blast furnace to extract iron from large quantities of iron-bearing materials. For given n samples,

1) The authors wihses to acknowledge the financial support of the Korea Research Foundation made in the program year of 1997, Project No. 1997-001-D00067
2) Associate Professor, Department of Statistics and Information, Anyang University, Korea.
3) Professor, Department of Statistics, Won-Kwang University, Korea.

$(X_1, Y_1), \cdots, (X_n, Y_n)$, where $X_1, \cdots, X_n$ are over C according to a certain distribution, let $X_{[i:n]}$ be the concomitant corresponding to the Y order statistics. Then the best-r-points-average estimator, $\hat{x}(r)$ of x_0 is defined by $\hat{x}(r) = \sum_{i=1}^{r} X_{[n-i+1]}/r$. This is the mean of the concomitants of extreme order statistics.

The research of concomitants of order statistics were initiated by David(1973) and Bhattacharya(1974). The extended work were obtained by David and Galambos(1974), and David et al.(1977). Yang(1977) has investigated the general distribution theory, exact moment and applications. And Bhattacharya(1984) summarized the previous results. The recent work was Nagaraja and David(1994)'s.

Hann(1981) used a similar approach with uniformly distributed points over C, although it was assumed that the regression function $\theta(\cdot)$ is observed without error. Chen et al.(1996) found the convergence rates of $\widehat{x_0}(r) = \sum_{i=1}^{r} X_{[n-i+1]}/r$ on some restricted families of error distributions, which is based on the theory of extreme value distribution. We will study here the asymptotic distribution of this best-r-points estimator extensively on the basis of the theory of the concomitants of order statistics. Other non-sequential estimate of x_0 was Muller(1989)'s, which is based on kernel estimate $\theta(\cdot)$ over C with data-driven bandwidth.

Many sequential approaches have been treated in the literature such as Kiefer and Wolfowitz(1952)'s recursive stochastic approximation method, Hotelling(1941)'s stagewise approach, the response surface method (Box and Wilson(1951)). Fabian(1967) considered a modified KW procedure. Following Box and Wilson(1951), Nadaraya(1964) and Devroye(1978) treated the response surface methodology. And Chen(1988) described a two-stage estimator. These methods have been found to be useful in many applications, but there are still many cases that these methods are not appropriate. For example, FSH (a hypophyseal hormone) curve is not practically feasible when selecting the design points sequentially for identifying a peak in the curve. So our method will be fitted well in this case.

To facilitate the discussions in sections 3, 4 and 5, we briefly review the aspects of the extreme-value distribution theory in section2. In section 3, using the theorem of concomitants of order statistics, we obtain the finite and asymptotic distributions of $\theta(X_{[n-r+1:n]})$ under the regression model. The asymptotic distribution of $X_{[n-r+1:n]}$ will be investigated in section 4 extensively. In section 5, we will investigate the distributive behavior of an estimator of the peak of a regression function such as Max{ $\theta(X_{[n-r+1:n]}), \ldots, \theta(X_{[n:n]})$ }.

The symbols $\xrightarrow{d}$ and $\xrightarrow{P}$ represent convergence in distribution and convergence in probability respectively. Let α_T (or $\alpha(T)$) and ω_T (or $\omega(T)$)denote the left and right endpoints of F_T, respectively, which are defined as $\alpha_T = \inf\{t: F_T(t) > 0\}$ and $\omega_T = \sup\{t: F_T(t) \le 1\}$.

2. Extreme-Value Distribution

For the discussions in section 3, 4 and 5, we briefly review the theorem of the extreme-value distribution which can be found in chapter 5 of Reiss(1989) on the name of the domain of convergence. Let $Y_1, \cdots, Y_n$ be independent random variables with common distribution $F_Y(y)$. Let $Y_{n:n}$ = Max $\{Y_1, \cdots, Y_n\}$. The distribution $F_Y(y)$ is said to be in the domain of attraction of a distribution G (written $F_Y \in D(G)$ or $Y \in D(G)$ if there are $a_n(>0)$, b_n so that

$$\lim_{n\to\infty} F^n{}_Y(a_n x + b_n) = G(x)$$

at every continuity point of G. If $F_Y(y)$ is such that a limiting distribution exists after suitable standardization, then this limiting distribution must be one of just three types, namely,

$$\begin{aligned} \Phi_\alpha(x) &= 0 & & x \le 0, \\ &= \exp(-x^{-\alpha}) & & x > 0, \alpha > 0; \end{aligned}$$

$$\begin{aligned} \Psi_\alpha(x) &= \exp[-(-x)^\alpha] & & x \le 0, \\ &= 1 & & x > 0, \alpha > 0; \end{aligned}$$

$$\Lambda(x) = \exp(-e^{-x}), \qquad -\infty < x < \infty.$$

Let's consider how to choose the norming constants a_n and b_n. If $\omega(Y) = \infty$, and $\lim_{y\to\infty} \frac{y f_Y(y)}{1 - F_Y(y)} = \alpha$, with some $\alpha > 0$, then there are constants $a_n > 0$ and b_n such that $(Y_{n:n} - b_n)/a_n$ converges to Φ_α. The constants can be chosen as $a_n = F^{-1}{}_Y(1 - 1/n)$ and $b_n = 0$.

If $\omega(Y) < \infty$ and $\lim_{y \to \omega(Y)} (\omega(Y) - y) f_Y(y) / [1 - F_Y(y)] = \alpha$, then $(Y_{n:n} - b_n)/a_n$ converges to Ψ_α, where the constants can be chosen as $a_n = F^{-1}{}_Y(1 - 1/n)$ and $b_n = \omega(Y)$.

If $\int_{-\infty}^{\omega(Y)} (1 - F_Y(y))\, dy < \infty$ and $\lim_{y \to \omega(Y)} \frac{f_Y(y)}{[1 - F_Y(y)]^2} \int_y^{\omega(Y)} [1 - F_Y(u)]\, du = 1$, then $(Y_{n:n} - b_n)/a_n$ converges to Λ, where the constants can be chosen as $a_n = [n f_Y(b_n)]^{-1}$ and $b_n = F_Y{}^{-1}(1 - 1/n)$.

3. Distribution of $Z_{[n-r+1:n]}$

Suppose (Y_i, Z_i) $(i = 1, 2, \cdots, n)$ is a random sample of n observations of a bivariate random variable (Y, Z) with $Y = Z + \varepsilon$, where Z and ε are independent variables. Note that only the Y 's are observed and the corresponding Z_i is $\theta(X_i)$. Let $Y_{i:n}$ denote the ith order statistics of the Y sample values. The Z-values associated with $Y_{i:n}$ is denoted by $Z_{[i:n]}$ and is called the concomitant of the ith order statistic. We can formulate this problem as follows.

$$
\begin{aligned}
Y_i &= \theta(X_i) + \varepsilon_i, \quad i = 1, \cdots, n \\
Y_{n-r+1:n} &= \theta(X_{[n-r+1:n]}) + \varepsilon_{[n-r+1]} \\
Y_{n-r+1:n} &= Z_{[n-r+1:n]} + \varepsilon_{[n-r+1]}
\end{aligned} \tag{3.1}
$$

From (3.1), we note that $\varepsilon_{[n-r+1]}$ and $Z_{[n-r+1:n]}$ are independent and $\varepsilon_{[n-r+1]}$ has a same distribution as ε. Hence the distribution of both $Z_{[n-r+1:n]}$ and $Z_{[n-r+1:n]} - \varepsilon$ has a same distribution. Assume that r is fixed as n approaches infinity (extreme case). Firstly, we consider the finite distribution of $Z_{[n-r+1:n]}$.

$$
\begin{aligned}
&P\{ Z_{[n-r+1:n]} \le z \} \\
&= \int_{y_0 < y_1} P[Z_{[n-r+1:n]} \le z \mid Y_{n-r:n} = y_0,\ Y_{n-r+1:n} = y_1] f_{Y_{n-r:n}}(y_0) \frac{f(y_1)}{1 - F_Y(y_0)}\, dy_1\, dy_0 \\
&= \int_{y_0 < y_1} P[Z \le z \mid Y = y_1] f_{Y_{n-r:n}}(y_0) \frac{f(y_1)}{1 - F_Y(y_0)}\, dy_1\, dy_0 \\
&= \int_{y_0 = -\infty}^{\infty} \int_{y_1 = y_0}^{\infty} P[\varepsilon \ge y_1 - z \mid Y = y_1] \frac{f(y_1)}{1 - F_Y(y_0)}\, dy_1\, f_{Y_{n-r:n}}(y_0)\, dy_0 .
\end{aligned} \tag{3.2}
$$

Now we investigate the asymptotic distribution of $Z_{[n-r+1:n]}$ as $n \to \infty$. Considering on

(3.1), we have to divide two cases such as $\omega(Y) < \infty$ and $\omega(Y) = \infty$. That implies $\omega(\varepsilon) < \infty$ and $\omega(\varepsilon) = \infty$ individually.

(1) $\omega(Y) < \infty$ (i.e. $\omega(\varepsilon) < \infty$)

When the distribution of Y is limited on the right (i.e. $\omega(\varepsilon) < \infty$) , $Y_{n-r+1:n}$ converges in probability to $\omega(Y)$. Then from (3.1), we have

$$Z_{[n-r+1:n]} \xrightarrow{d} \omega(Y) - \varepsilon. \tag{3.3}$$

Hence if $Y \in D(\Psi_\alpha)$, then $\omega(Y) < \infty$ and (3.3) is satisfied. And if $Y \in D(\Lambda)$ and $\omega(Y)$ is finite, then (3.3) is also holds. In this case, the distribution of $Z_{[n-r+1]}$ depends on both $\omega(Y)$ and ε. If we consider ε is a continuous random variable and under the independence assumption between Z and ε, (3.3) may be expressed as follows.

$$\theta(X_{[n-r+1:n]}) - \theta(x_0) \xrightarrow{d} \omega(\varepsilon) - \varepsilon.$$

(3.4)

(2) $\omega(Y) = \infty$ (i.e. $\omega(\varepsilon) = \infty$)

Firstly, suppose that $F_Y(y) < 1$ for every finite y. Then there exists (Gnedenko, 1943) a sequence of constants $\{A_n\}$ such that $\{Y_{n-r+1:n} - A_n\} \xrightarrow{P} 0$

$$\text{iff } \lim_{y\to\infty} \frac{1-F_Y(y+\varepsilon)}{1-F_Y(y)} = 0 \text{ for every } \varepsilon > 0 \tag{3.5}$$

$$\text{or } \lim_{y\to\infty} \frac{f_Y(y+\varepsilon)}{f_Y(y)} = 0 .$$

If Y is satisfied on (3.5), then we have

$$Z_{[n-r+1:n]} - A_n \xrightarrow{d} -\varepsilon.$$

Example 1. The condition (3.5) is satisfied, in particular, when Y is standard normal, with $A_n = (2\log n)^{\frac{1}{2}}$. Hence we then have

$$Z_{[n-r+1:n]}-(2\log n)^{\frac{1}{2}} \xrightarrow{d} -\varepsilon.$$

It can be known that $\omega(Y)=\infty$ imply $\omega(\varepsilon)=\infty$ since $\theta(x_0)<\infty$. Suppose $(Y_{n:n}-b_n)/a_n$ has a limiting distribution $G(y)$ for suitable choices of $a_n(>0)$ and b_n. Then the two possible limiting distributions unbounded on the right are Φ_α and Λ. The corresponding limiting distributions for $(Y_{n-k+1:n}-b_n)/a_n$ ($k=1,\cdots,r$) are

$$G^{(k)}(x) = G(x)\sum_{j=0}^{k-1}[-\log G(x)]^j/j! \ . \tag{3.6}$$

$G^{(k)}(y)$ is the c.d.f. of kth lower record value from the c.d.f. G. (Nagaraja and David(1994) Lemma 1). With $U_{r,n}=(Y_{n-r+1:n}-b_n)/a_n$, (3.1) can be written,

$$Z_{[n-r+1:n]} = a_n U_{r,n}+ b_n - \varepsilon_{[n-r+1]} \tag{3.7}$$

We consider the two families of extreme-value distributions in turn, using results essentially given by Gnedenko(1943).

$F_Y(y)$ belongs to the domain of attraction of Φ_α (i.e. $Y\in D(\Phi_\alpha)$),

$$\text{iff} \ \lim_{t\to\infty}\frac{1-F_Y(t)}{1-F_Y(kt)} = k^\alpha \text{ for every } k>0.$$

$$(\text{or} \ \lim_{t\to\infty}\frac{tf_Y(t)}{1-F_Y(t)} = \alpha \text{ for every } \alpha.)$$

Correspondingly, $b_n=0$, and a_n may be taken as $F_Y^{-1}(1-\frac{1}{n})$. If $F_Y^{-1}(1-\frac{1}{n})$ approaches infinity, by (3.7), we have the following result,

$$\frac{Z_{[n-r+1:n]}}{a_n} \xrightarrow{d} \lim_{n\to\infty} U_{r,n} = U_r \tag{3.8}$$

$$(\text{i.e. } \lim_{n\to\infty} P[\frac{Z_{[n-r+1:n]}}{a_n} \le x] = \Phi_\alpha^{(r)}(x) \ ),$$

where $\Phi_\alpha^{(r)}(x)$ is the $c.d.f.$ of rth lower record value from the $c.d.f.$ Φ_α.

By one of the very useful sufficient conditions of von Mises(1936), $F_Y(y)$ belongs to the domain of attraction of Λ (i.e. $Y\in D(\Lambda)$), if $F_Y(y)$ is less than 1 for every finite y, is twice differentiable at least for all y greater than some value y', and is such that

$$\lim_{t\to\infty} \frac{d}{dt} [\frac{1-F_Y(t)}{f_Y(t)}] = 0.$$

Correspondingly, a_n and b_n may be taken as

$$a_n = \frac{1}{nf(b_n)} \text{과} \quad b_n = F_Y^{-1}(1-\frac{1}{n}).$$

In this case, $\lim_{n\to\infty} a_n = a$ may assume any nonnegative value. From (3.7), we can have three limiting distributions depending on a values as follows.

$$\begin{aligned} &Z_{[n-r+1:n]} - F_Y^{-1}(1-\frac{1}{n}) \longrightarrow^d -\varepsilon &, \quad a=0 \\ &Z_{[n-r+1:n]} - F_Y^{-1}(1-\frac{1}{n}) \longrightarrow^d aU_r-\varepsilon &, \quad 0 < a < \infty \\ &\frac{Z_{[n-r+1:n]}}{a_n} - \frac{b_n}{a_n} \longrightarrow^d U_r &, \quad a=\infty. \end{aligned} \tag{3.9}$$

Example 2. The family of Weibull distributions

$$\begin{aligned} F(y) &= 0\,, & y \le 0 \\ &= 1-e^{-y^{\alpha}} &, \quad y>0, \alpha>0 \end{aligned}$$

satisfies $\lim_{n\to\infty} F^n(a_n y + b_n) = e^{-e^{-y}}$, where $a_n = \frac{1}{\alpha}(\log n)^{\frac{(1-\alpha)}{\alpha}}$, $b_n = (\log n)^{\frac{1}{\alpha}}$.

Thus

$$\lim_{n\to\infty} a_n = \begin{cases} \infty & 0<\alpha<1 \\ 1 & \alpha=1 \\ 0 & \alpha>1 \end{cases}.$$

and from (3.9), we have

$$\begin{aligned} &\frac{Z_{[n-r+1:n]}}{a_n} - \alpha(\log n) \longrightarrow^d U_r &, \quad 0<\alpha<1 \\ &Z_{[n-r+1:n]} - (\log n)^{\frac{1}{\alpha}} \longrightarrow^d U_r - \varepsilon &, \quad \alpha=1 \\ &Z_{[n-r+1:n]} - (\log n)^{\frac{1}{\alpha}} \longrightarrow^d -\varepsilon &, \quad \alpha>1. \end{aligned}$$

4. Distribution of $\widehat{x_0}(r)$.

Chen et al. (1996) found the convergence rates of $\widehat{x_0}(r) = \sum_{i=1}^{r} X_{[n-i+1]}/r$ on the conditions of both $\varepsilon \in D(\Psi_a)$ and $\varepsilon \in D(\Lambda)$. On the nonparametric point of view, the asymptotic distribution of $\widehat{x_0}(r) = \sum_{i=1}^{r} X_{[n-i+1]}/r$ will be available only on $\omega(Y) < \infty$ case. It can be seen that the limiting distributions of $Z_{[n-k+1]}$'s, $k = 1, \cdots, r$ on (3.8) and (3.9) depends on Y 's limiting distributions. Hence using (3.4), we can consider the asymptotic distribution of the best- r-points-average estimator for the peak of regression function.

Condition R. $\omega_Z < \infty$ and F_Z satisfies $c_1 z^{\tau} \le 1 - F_Z(\omega_z - z) \le c_2 z^{\tau}$ as $z \to 0^+$. Here, $0 < \tau \le 1$ and c_1 and c_2 are positive constants.

Let us have an an example which illustrates when Condition R will hold.

Example 3. Assume $f_X(x)$ is supported on C and is bounded away from zero and infinity over C.
Note that

$$P(\omega_Z - z \le Z \le \omega_Z) = P(\omega_Z - z \le \theta(X) \le \omega_Z)$$

$$= P(\theta(x_0) - z \le \theta(X)) = P(\theta(x_0) - \theta(X) \le z).$$

Now suppose there exist some positive constants c_3, c_4, and ρ, $\rho \ge 1$, such that

$$c_3 |x - x_0|^{\rho} \ge |\theta(x) - \theta(x_0)| \ge c_4 |x - x_0|^{\rho} \text{ for all } x \in C. \tag{4.1}$$

Since $P(|X - x_0| \le (\frac{z}{c_3})^{1/\rho}) \le P(\theta(x_0) - \theta(X) \le z) \le P(|X - x_0| \le (\frac{z}{c_4})^{1/\rho})$,
Z satisfied Condition R with $\tau = 1/\rho$.

Condition E. $\omega_\varepsilon < \infty$ and for some $k \ge 0$, f_ε and F_ε satisfy $(-1)^k f_\varepsilon^{(k)}(\omega_\varepsilon) > 0$, $f_\varepsilon^{(j)}(\omega_\varepsilon) = 0$ for every $0 \le j \le k-1$, and $\lim_{t \to \omega_\varepsilon} (\omega_\varepsilon - t) f_\varepsilon(t) / [1 - F_\varepsilon(t)] = k+1$.

Lemma 1. (Chen et al. (1996)) Suppose $Y = Z + \varepsilon$, where Z and ε are independent. Let Z satisfy Condition R. Then $Z_{[n-i+1:n]} - \omega_Z = O_P((\log n/n)^{1/[1+(k+1)]/\tau}$ for $i = 1, \cdots, r$, under Condition E.

Lemma 2. (Chen et al. (1996))

Assume there exist some positive constants c_3, c_4 such (4.1) holds. Then $\widehat{x_0} - x_0 = O_P((\log n/n)^{1/[1+(k+1)]/\tau})$, where $\widehat{x_0}(r) = \sum_{i=1}^{r} X_{[n-i+1]}/r$.

Theorem. Let r_n be a positive real sequence such that $r_n \to \infty$ and $\frac{r_n}{n} \to 0$. Assume that $\theta(x)$ is a bounded function with 3 continuous derivatives at $x = x_0$. Under Condition R and E,

$$\left(\frac{n}{\log n}\right)^{1/2(1+(k+1)/\tau)} (X_{[n-i+1:n]} - x_0) \longrightarrow^d (\omega(\varepsilon) - \varepsilon)^{\frac{1}{2}} \text{ and}$$

$$\frac{r_n^{1/2}}{(Var(\omega_\varepsilon - \varepsilon)^{\frac{1}{2}})^{1/2}} \left[\left(\frac{n}{\log n}\right)^{1/2(1+(k+1)/\tau)} (\widehat{x_0}(r_n) - x_0) - E(\omega_\varepsilon - \varepsilon)^{\frac{1}{2}} \right] \longrightarrow^d N(0,1).$$

Proof. Let $g_n = (\log n/n)^{1/[1+(k+1)]/\tau}$.

Since $X_{[n-r+1:n]} = x_0 + O_P(g_n)$ by Lemma 2 and $g_n \to 0$ as $n \to \infty$, we can expand as follows:

$$\theta(X_{[n-r+1:]}) = \theta(x_0) + \theta'(x_0)(X_{[n-r+1:n]} - x_0) + \frac{\theta''(x_0)}{2}(X_{[n-r+1:n]} - x_0)^2 + O_P(g_n{}^6).$$

But $\theta'(x_0) = 0$ and by Lemma 1 and 2, it can be shown that

$$[\theta(X_{[n-r+1:]}) - \theta(x_0)] - [g_n^{-2}(X_{[n-r+1:n]} - x_0)^2] = O_P(g_n{}^6).$$

Hence $(\theta(X_{[n-r+1:n]}) - \theta(x_0)) - g_n{}^{-2}(X_{[n-r+1:n]} - x_0)^2$ converges to 0 in probability, and that means $(\theta(X_{[n-r+1:n]}) - \theta(x_0))$ and $g_n{}^{-1}(X_{[n-r+1:n]} - x_0)^2$ have a same limiting distribution.

From (3.4),

$$(\theta(X_{[n-r+1:n]}) - \theta(x_0)) \longrightarrow^d \omega(\varepsilon) - \varepsilon.$$

This implies $g_n{}^{-2}(X_{[n-r+1:n]} - x_0)^2 \longrightarrow^d \omega(\varepsilon) - \varepsilon$.

By a continuous mapping theorem of limiting distribution, we have

$$g_n^{-1} |X_{[n-r+1:n]} - x_0| \longrightarrow^d (\omega(\varepsilon) - \varepsilon)^{1/2}.$$

Since Z and ε are independent and $Y_{n-i+1:n} \rightarrow^P \omega(Y)$ for $i = 1, \cdots, r$ it can be shown that $\theta(X_{[n-r+1:n]}), \ldots, \theta(X_{[n:n]})$ are asymptotically independent. Hence $X_{[n-r+1:n]}, \ldots, X_{[n:n]}$ are asymptotically independent. Now by the central limit theorem, we have,

$$\frac{r_n^{1/2}}{(Var(\omega_\varepsilon - \varepsilon)^{\frac{1}{2}})^{1/2}} [(\frac{n}{\log n})^{1/[1+(k+1)/\tau]} |\widehat{x}_0(r_n) - x_0| - E(\omega_\varepsilon - \varepsilon)^{\frac{1}{2}}] \longrightarrow^d N(0,1).$$

Example 4. In order to have some ideas on how well our asymptotic distribution of the best-r-points-average method distributed, we consider $Y = \theta(X) + \varepsilon$, where $\theta(x) = 1 + 3\exp(-(x-0.5)^2/0.01)$ (symmetric peak at (0.5, 4.0)= $(x_0, \theta(x_0))$) and ε is distributed by Uniform $[-\sigma\sqrt{3}, \sigma\sqrt{3}]$. Then k=0 and $\tau = 1/2$.

$$(\frac{n}{\log n})^{1/3} |X_{[n-i+1:n]} - x_0| \longrightarrow^d (\sigma\sqrt{3} - \varepsilon)^{\frac{1}{2}} \text{ and}$$

$$\frac{r_n^{1/2}}{(\frac{1}{3\sqrt{3}\sigma})^{1/2}} [(\frac{n}{\log n})^{1/3} |\widehat{x}_0(r_n) - x_0| - \frac{2^{3/2}}{3^{3/4}} \sigma^{1/2}] \longrightarrow^d N(0,1).$$

5. Distribution of $V_{r,n}$.

Let $V_{r,n}$ be max { $Z_{[n-r+1]}, \ldots, Z_{[n:n]}$ }. $V_{r,n}$ is a maximum of the concomitants of extreme order statistics and a good estimator for the maximum of regression function, $\theta(x_0)$. Using the theory of the concomitants of order statistics, we can investigate the behavior of $V_{r,n}$ as an estimator for the peak of regression function. Firstly, we consider the finite distribution of $V_{r,n}$.

$$P\{V_{r,n} \le v\}$$
$$= P(Z_{[n-r+1:n]} \le v, \ldots, Z_{[n:n]} \le v) \quad (5.1)$$

$$= \int_{y_0 < y_1 < \cdots < y_r} P(Z_{[n-r+1:n]} \le v, \cdots,$$
$$Z_{[n:n]} \le v \mid Y_{n-r:n} = y_0, Y_{n-r+1:n} = y_1, \cdots, Y_{n:n} = y_r)$$
$$f_{Y_{n-r:n}}(y_0)\, r! \left\{ \prod_{i=1}^{r} \frac{f_Y(y_i)}{1-F_Y(y_0)} \right\} dy_0$$
$$= \int_{y}^{\infty} \left(r! \int_{y_0 < y_1 \cdots < y_r} \left\{ \prod_{i=1}^{r} P[Z_{[n-i+1:n]} \le v \mid Y_{n-r+i:n} = y_i] \frac{f_Y(y_i)}{1-F_Y(y_0)}\, dy_i \right\} \right)$$
$$f_{Y_{n-r:n}}(y_0)\, dy_0$$

The integrand of the multiple integral in brackets above is a symmetric function of the variables $y_1, \ldots, y_r$. Thus it can be expressed as

$$\left(\int_{y=y_0}^{\infty} F_Z(v \mid y) \frac{f_Y(y)}{1-F_Y(y_0)}\, dy \right)^r.$$

This simplification leads to the following compact representation for the c.d.f. of $V_{r,n}$:

$$F_{r,n}(v) = \int_{y_{\cdots}}^{\infty} \left(\int_{y=y_0}^{\infty} \frac{P[Y-\varepsilon \le v \mid Y=y] f_Y(y)}{1-F_Y(y_0)}\, dy \right)^r f_{Y_{n-r:n}}(y_0)\, dy_0 \qquad (5.2)$$
$$= \int_{y_{\cdots}}^{\infty} \left(\int_{y=y_0}^{\infty} \frac{P[\varepsilon \ge y-v \mid Y=y] f_Y(y)}{1-F_Y(y_0)}\, dy \right)^r f_{Y_{n-r:n}}(y_0)\, dy_0.$$

Now we investigate the asymptotic distribution of $V_{r,n}$ as $n \to \infty$ when r is fixed. As section 3, we divide two cases such as $\omega(Y) < \infty$ and $\omega(Y) = \infty$. That implies $\omega(\varepsilon) < \infty$ and $\omega(\varepsilon) = \infty$ individually.

(1) $\omega(Y) < \infty$ (i.e. $\omega(\varepsilon) < \infty$)

From (3.3) of section 3, $Z_{[n-r+1:n]} \xrightarrow{d} \omega(Y) - \varepsilon$ implies that

$$V_{r,n} \xrightarrow{d} \omega(Y) - \min(\varepsilon_{[n-r+1]}, \ldots, \varepsilon_{[n:n]}).$$

Because $(\varepsilon_{[n-r+1]}, \ldots, \varepsilon_{[n:n]})$ are independent random variables, we have,

$$V_{r,n} \xrightarrow{d} \omega(Y) - \min(\varepsilon_1, \ldots, \varepsilon_r).$$

Since we assume ε is a continuous random variable and the independence assumption between Z and ε, we can express as follows.

$$V_{r,n} - \theta(x_0) \longrightarrow^{d} \omega(\varepsilon) - \varepsilon_{1:r}.$$

(2) $\omega(Y) = \infty$ (i.e. $\omega(\varepsilon) = \infty$)

If (3.5) is satisfied, then $Z_{[n-r+1:n]} - A_n \longrightarrow^{d} -\varepsilon$. Hence $V_{r,n}$ can be expressed,

$$V_{r,n} - A_n \longrightarrow^{d} -\varepsilon_{1:n}$$

(3.1) can be expressed as follows.

$$\frac{Z_{[i:n]} - b_n}{a_n} = \frac{Y_{i:n} - b_n}{a_n} - \frac{\varepsilon_{[i]}}{a_n} \quad n-r+1 \le i \le n. \tag{5.3}$$

Let W_1 be a random variable distributed by G, which is a limiting distribution of $\frac{Z_{[n:n]} - b_n}{a_n}$. And let W_i ($i = 2, \cdots, r$) be the successive lower record value distributed by $G^{(i)}$ from (3.6).

When $G = \Phi_a$, a_n can be taken to be $F_Y^{-1}(1-1/n)$ and b_n to be 0. Since a_n approaches infinity, $\frac{\varepsilon_{[i]}}{a_n} \longrightarrow^{p} 0$. Hence from (5.3), we conclude that

$$\left(\frac{Z_{[n:n]}}{a_n}, \ldots, \frac{Z_{[n-r+1:n]}}{a_n} \right) \longrightarrow^{d} (W_1, \ldots, W_r).$$

Thus we have, with $b_n = 0$,

$$\left\{ \frac{(V_{r,n} - b_n)}{a_n} \right\} \longrightarrow^{d} W_1. \tag{5.4}$$

When $G = \Lambda$, we can choose a_n = E[$Y - b_n$ | $Y > b_n$] and $b_n = F_Y^{-1}(1-1/n)$. Suppose $a_n \rightarrow a$. If a is zero, then $Y_{n-i+1:n} - b_n \longrightarrow^{p} 0$ and hence (3.5) is satisfied with $A_n = a_n$. If a is infinite, (5.4) holds. If a is finite and positive, using (5.3),

$$\left(\frac{Z_{[n:n]} - b_n}{a_n}, \ldots, \frac{Z_{[n-r+1:n]} - b_n}{a_n} \right) \longrightarrow^{d} \left(W_1 + \frac{\varepsilon_1}{a}, \ldots, W_r + \frac{\varepsilon_r}{a} \right)$$

This means that if $\frac{(V_{r,n} - b_n)}{a_n} \longrightarrow^d V$, then

$$V =^d \max (W_1 + \frac{\varepsilon_1}{a}, \ldots, W_r + \frac{\varepsilon_r}{a}), \tag{5.5}$$

where and ε's are i.i.d. random variables. Since W's are dependent, it is difficult to find explicit form of the c.d.f. of V in (5.5).

References

[1] Bhattacharya, P. K. (1974). Convergence of sample paths of normalized sums of induced order statistics, *Annals of Statistics,* Vol. 2, 1034-1039.

[2] Bhattacharya, P. K. (1984). *Induced order statistics: Theory and Applications, Handbook of Statistics* (P. R. Krishnaiah and P. K. Sen edited).

[3] Box, G. E. P. and Wilson, K. B. (1951). On the experimental attainment of optimum conditions. *Journal of Royal Statistical Society Series B,* Vol. 13, 1-45.

[4] Changchien, G. M. (1990). *Optimization of blast furnace burden distribution.* In Proceedings of the 1990 Taipei Symposium in Statistics, June 28-30, 1990 (M. T. Chao and P. E. Cheng, Eds.), 63-78.

[5] Chen, H. (1988). Lower rate of convergence for locating a maximum of a function. *Annals of Statistics,* Vol. 3, 1330-1334.

[6] Chen, H. , Huang, M. L. and Huang, W. J. (1996) Estimation of the location of the maximum of a regression function using extreme order statistics. *Journal of Multivariate Analysis,* Vol. 57, 191-214.

[7] David, H. A. (1973) Concomitants of order statistics. *Bulletin of International Institute,* Vol. 4, 295-300.

[8] David, H. A. and Galambos, J. (1974) The asymptotic theory of concomitants of order statistics. *Journal Applied Probability,* Vol. 11, 762-770.

[9] David, H. A. , O'Connell, M. J. and Yang, S. S. (1977) Distribution and expected value of the ranks of a concomitant of an order statistics. *Annals of Statistics,* Vol. 5, 996-1002.

[10] Devroye, L. P. (9184). The uniform convergence of nearest neighbor regression function estimators and their application in optimization. *IEEE Transaction Information Theory, IT-24, 142-151.*

[11] Fabian, V. (9167). Stochastic approximation of minima with improved asymptotic speed. *Annals of Mathematical Statistics,* Vol. 38, 91-200.

[12] Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une serie aleatoire. *Annals of Mathematics,* Vol. 44, 423-453.

[13] Haan, L. DE (1981). Estimation of the minimum of a function using order statistics.

Journal of American Statistical Association, Vol. 76, 467-469.

[14] Hotelling, H. (1941). Experimental determination of the maximum of a function. *Annals of Mathematical Statistics,* Vol. 12, 20-45.

[15] Kieffer, J., and Wolfowitz, J. (1952). Stochastic estimation of the maximum of a regression function. *Annals of Mathematical Statistics,* Vol. 23, 462-466.

[16] Muller, H. G. (1989) Adaptive nonparametric peak estimation. *Annals of Statistics,* Vol. 17, 1053-1069.

[17] Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability Applied,* Vol. 9, 141-142.

[18] Nagaraja, H. N. and David, H. A. (1994) Distribution of the maximum of concomitants of selected order statistics. *Annals of Statistics,* Vol. 22, 478-494.

[19] Reiss, R. D. (1989). *Approximate Distributions of Order Statistics: With Applications to Nonparametric Statistics.* Springer-Verag, New York.

[20] von Mises, R. (1936). La distribution de la plus grande den valeurs. Rev. Math. Union Interbalkanique 1, 141-160. [Reproduced in Selected Papers of Richard von Mises, *2, American Mathematical Society, Providence.*

[21] Yang, S. S. (1977) General distribution theory of the concomitants of order statistics. *Annals of Statistics, 5996-1002.*