Estimation of Variance and Its Properties in Measurement Error Model

Chen Qiuhua[1] and Cui Hengjian[2]

[1]*Department of Mathematics and Physics, North China Electric Power University(Beijing), Beijing, 102206, China*

[2]*Department of Statistics and Financial Mathematics, School of Mathematical Sciences, Beijing Normal University, Beijing, 100875, China*

Email: hjcui@bnu.eud.cn

Abstract

This paper introduces variance estimators of Y in measurement error model $X = Y + E$ with known $Var(E)$. An unbiased estimator and an adjust non-negative estimator of $Var(Y)$ are given. Meanwhile, some properties of expectation and variance for both estimators are discussed, and some relationships of expectation (or variance) with sample size and $Var(Y)/Var(E)$ are also shown.

Key words: Measurement error model, independence, unbiased estimator, non-negative estimator.

1 Introduction

A general measurement error model is as following $X = Y + E$, where Y, E are independent, error E has known variance σ_0^2, see Fuller (1987). It is obvious that the variance of X is $\sigma^2 = \sigma_y^2 + \sigma_0^2$, where σ_y^2 is the variance of Y. We are interested in estimating the variance of Y. Let $X_1, X_2, \cdots, X_n$ be a sample from X. Of course, an unbiased estimator of σ^2 is the sample variance $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})^2$ with $\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i$, and

$$\hat{\sigma}_y^2 = S^2 - \sigma_0^2$$

[0]Corresponding author: Cui Hengjian, email: hjcui@bnu.eud.cn

is naturally an unbiased estimator of σ_y^2. If $Y \sim N(\mu, \sigma_y^2)$ and $E \sim N(0, \sigma_0^2)$, then it is well known that S^2 is a minimum uniformly variance and unbiased estimator of σ^2, and $\hat{\sigma}_y^2$ is a minimum uniformly variance and unbiased estimator of σ_y^2 since S^2 is a complete and sufficient statistic of σ^2.

Unfortunately, $\hat{\sigma}_y^2$ is not always non-negative and unreasonable as an estimator of the variance σ_y^2. For example, when $Y \sim N(0, \sigma_y^2)$ and $E \sim N(0, 1)$, then $X \sim N(0, \sigma_y^2 + 1)$, and the probability of $\hat{\sigma}_y^2$ is negative for various n and σ_y^2 is listed in the following table.

Table. $P\{\hat{\sigma}_y^2 < 0\}$ for various n and σ_y^2

n	3	5	8	10
$\sigma_y^2 = 0.5$	0.487	0.385	0.299	0.260
$\sigma_y^2 = 1$	0.394	0.260	0.164	0.124
$\sigma_y^2 = 2$	0.283	0.144	0.061	0.034

From the above table, $P\{\hat{\sigma}_y^2 < 0\}$ is always positive, if $n = 5$ and $\sigma_y^2 = 1$, then the probability of $\hat{\sigma}_y^2$ is negative is larger than $1/4$. Note that S^2 is a complete and sufficient statistic of σ^2 in normal case, thus $S^2 - \sigma_0^2$ is just unique a.s. unbiased estimator of σ_y^2 based on the S^2. This means the unbiased and non-negative estimator of σ_y^2 based on S^2 *doesn't* exist for normal case.

This paper would give a non-negative estimator of σ_y^2 without unbias requirement and some properties are shown.

2. Non-negative Estimator and Lemmas

Denote that $\xi^+ = \max\{\xi, 0\}$ and $\xi^- = \max\{-\xi, 0\}$ are the positive and negative part of a random variable ξ respectively. Intuitively, if $\hat{\sigma}_y^2$ is negative, we would take "0" as an estimator of σ_y^2. Therefore, we take the positive part of $\hat{\sigma}_y^2$, i.e.

$$\hat{\sigma}_y^{2+} =: (S^2 - \sigma_0^2)^+$$

as a naive non-negative estimator of σ_y^2 based on S^2. To obtain main results, the following two lemmas are needed.

Lemma 1. If $E(X^2) < +\infty$, then $\hat{\sigma}_y^2$ is an unbiased estimator of σ_y^2 and $\hat{\sigma}_y^2 \xrightarrow{p.} \sigma_y^2$ as $n \to +\infty$. If $E(X^4) < +\infty$, then

$$V_n =: Var(\hat{\sigma}_y^2) = \frac{\mu_4}{n} - \frac{(n-3)\sigma^4}{n(n-1)}. \tag{2.1}$$

Moreover, $\sqrt{n}(\hat{\sigma}_y^2 - \sigma_y^2) \xrightarrow{d.} N(0, Var((X-\mu)^2))$ as $n \to +\infty$ provided $Var((X-\mu)^2) > 0$, where " $\xrightarrow{p.}$ " and " $\xrightarrow{d.}$ " stand for the convergence in probability and in distribution respectively, $\mu_4 = E(X-\mu)^4$, $\mu = E(X)$.

Lemma 2. Let ξ be a random variable with $Var(\xi) < +\infty$, then

$$Var(\xi) = Var(\xi^+) + Var(\xi^-) + 2E(\xi^+)E(\xi^-) \geq Var(\xi^+), \tag{2.2}$$

and

$$P\{\xi - E(\xi) \geq x\} \leq \frac{Var(\xi)}{Var(\xi) + x^2} \tag{2.3}$$

for any $x > 0$.

3. Main Results

Denote $F_{n-1}(t)$, $f_{n-1}(t)$ are the cumulative distribution function and density of $(n-1)S^2/\sigma^2$ respectively. The main theorems are listed below.

Theorem 1.

(i). Suppose $E(X^2) < +\infty$, we have $\lim_{n\to+\infty} E(\hat{\sigma}_y^{2+}) = \sigma_y^2$, and if $F_{n-1}(t)$ is continuous at "0" uniformly for σ_y, then $\lim_{\sigma_y^2\to+\infty}[E(\hat{\sigma}_y^{2+}) - \sigma_y^2] = 0$; if $f_{n-1}(t)$ is independent of σ^2, then $E(\hat{\sigma}_y^{2+})$ is a strictly increasing function of σ_y^2 and $E(\hat{\sigma}_y^{2+}) - \sigma_y^2$ is a strictly decreasing function of σ_y^2.

(ii). If $E(X^4) < +\infty$, then $|E(\hat{\sigma}_y^{2+}) - \sigma_y^2| \leq \sigma_0^2 V_n/(V_n + \sigma_y^4)$, where V_n is in (2.1).

Theorem 2. If $E(X^4) < +\infty$, then $Var(\hat{\sigma}_y^{2+}) \leq Var(\hat{\sigma}_y^2) = V_n$ and $MSE(\hat{\sigma}_y^{2+}) \leq MSE(\hat{\sigma}_y^2)$, where MSE stands for the mean square error. Meanwhile, $Var(\hat{\sigma}_y^{2+})$ is an increasing function of σ_y^2 provided $f_{n-1}(t)$ is independent of σ^2.

Theorems 1 and 2 say that although $\hat{\sigma}_y^{2+}$ is a biased estimator, but its variance and MSE are smaller than that of $\hat{\sigma}_y^2$.

Theorem 3. If $E(X^2) < +\infty$, then $\hat{\sigma}_y^{2+} \xrightarrow{p.} \sigma_y^2$ as $n \to +\infty$, and if $0 < Var((X-\mu)^2) < +\infty$, then $\sqrt{n}(\hat{\sigma}_y^{2+} - \sigma_y^2) \xrightarrow{d.} N(0, Var((X-\mu)^2))$ as $n \to +\infty$.

Theorem 3 says that $\hat{\sigma}_y^{2+}$ and $\hat{\sigma}_y^2$ have the same properties of consistency and asymptotic normality.

4 Comparisons and Figures for Normal Case

In this section, we give the comparisons of $Bias = E(\hat{\sigma}_y^{2+}) - \sigma_y^2$ for various σ_y, n and σ_0 in normal case. Note that $S^2 \sim \frac{\sigma_0^2+\sigma_y^2}{n-1}\chi_{n-1}^2$, it means $f_{n-1}(t)$ is the density of χ_{n-1}^2 which is independent of σ^2 in this sense, and

$$\int_z^\infty x^i dH_{n-1}(x) = (n^i - 1)(1 - H_{n-1+2i}(z))$$

for $i = 1, 2$ and any $z \geq 0$, where χ_{n-1}^2 and $H_{n-1}(x)$ are the random variable and the cumulative chi-square distribution function of with degree of freedom $n-1$ respectively. Therefore,

$$\begin{aligned} Bias &= E[(\frac{\sigma_0^2+\sigma_y^2}{n-1}\chi_{n-1}^2 - \sigma_0^2)^+] - \sigma_y^2 = \int_A^{+\infty}(\frac{\sigma_0^2+\sigma_y^2}{n-1}x - \sigma_0^2)dH_{n-1}(x) - \sigma_y^2 \\ &= \sigma_0^2 H_{n-1}(A) - (\sigma_0^2+\sigma_y^2)H_{n+1}(A), \end{aligned}$$

where $A = \frac{(n-1)\sigma_0^2}{\sigma_0^2+\sigma_y^2}$, and

$$\begin{aligned} DMSE &=: MSE(\hat{\sigma}_y^2) - MSE(\hat{\sigma}_y^{2+}) \\ &= MSE(\hat{\sigma}_y^2) - E[(\frac{\sigma_0^2+\sigma_y^2}{n-1}\chi_{n-1}^2 - \sigma_0^2)^+ - \sigma_y^2]^2 \\ &= (\sigma_0^2+\sigma_y^2)^2[H_{n-1}(A) + \frac{n+1}{n-1}H_{n+3}(A) - 2H_{n+1}(A)] - \sigma_y^4 H_{n-1}(A). \end{aligned}$$

Figures 1 and 3 present the three curves of $Bias$ and $DMSE$ for $\sigma_0 = 1$, 1.5, 2 respectively with fixed $n = 10$. When σ_y increasing, then $Bias$ approaches zero; when σ_y is larger, then $DMSE$ approaches zero.

Figures 2 and 4 show the three curves of $Bias$ and $DMSE$ for $n = 5, 10, 50$ respectively with fixed $\sigma_0 = 1$. When n is increasing, $E(\hat{\sigma}_y^{2+})$ approaches σ_y^2 and $DMSE$ approaches 0. So we could increase the sample size if it is possible.

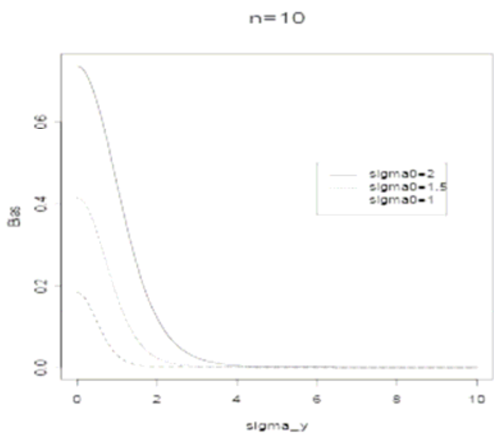

Figure 1.

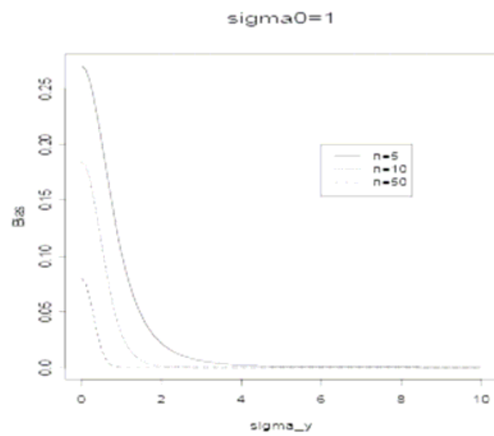

Figure 2.

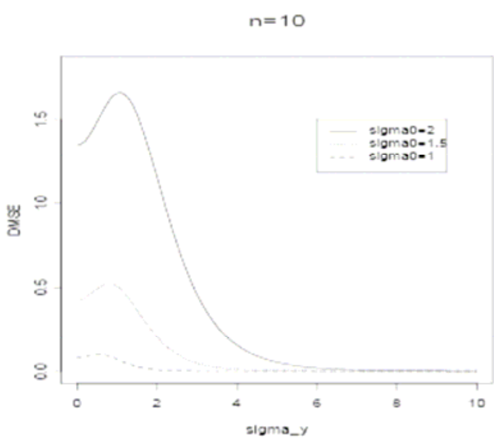

Figure 3.

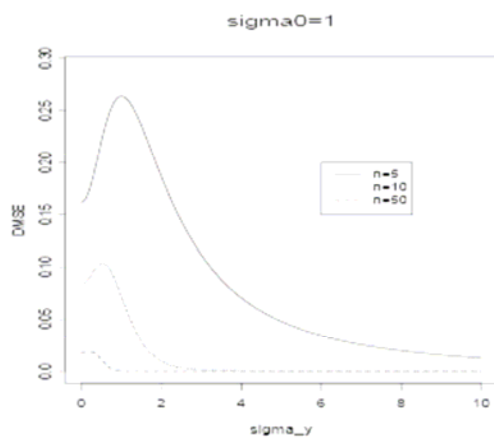

Figure 4.

5 Proof of Lemmas and Theorems

Proof of Lemma 1.

Since $E(S^2) = Var(X)$ and $S^2 \xrightarrow{p.} Var(X)$, then $E(\hat{\sigma}_y^2) = E(S^2) - \sigma_0^2 = Var(X) - \sigma_0^2 = \sigma^2 - \sigma_0^2 = \sigma_y^2$ and $\hat{\sigma}_y^2 \xrightarrow{p.} \sigma_y^2$ as $n \to +\infty$.

It can be derived by the similar way of Serfling (1980, p69-70) that

$$\begin{aligned} Var(\hat{\sigma}_y^2) &= Var(S^2) = (\frac{n}{n-1})^2[\frac{\mu_4 - \sigma^4}{n} - \frac{2(\mu_4 - 2\sigma^4)}{n^2} + \frac{\mu_4 - 3\sigma^4}{n^3}] \\ &= \frac{\mu_4}{n} - \frac{(n-3)\sigma^4}{n(n-1)}, \end{aligned}$$

and $\sqrt{n}(S^2 - Var(X)) \xrightarrow{d.} N(0, Var((X-\mu)^2))$. Thus, $\sqrt{n}(\hat{\sigma}_y^2 - \sigma_y^2) \xrightarrow{d.} N(0, Var((X-\mu)^2))$. □

Proof of Lemma 2.

Note that $\xi = \xi^+ - \xi^-$, $\xi^+\xi^- = 0$ and $Cov(\xi^+, \xi^-) = -E(\xi^+)E(\xi^-)$. Hence,

$$\begin{aligned} Var(\xi) &= Var(\xi^+ - \xi^-) = Var(\xi^+) + Var(\xi^-) - 2Cov(\xi^+, \xi^-) \\ &= Var(\xi^+) + Var(\xi^-) + 2E(\xi^+)E(\xi^-), \end{aligned}$$

it means (2.2) in Lemma 2 is true. The proof of (2.3) in Lemma 2 can be referred to Laha and Rohatgi (1979, p62). □

Proof of Theorem 1.

(i). Let $\sigma^2 = \sigma_y^2 + \sigma_0^2$, $A = (n-1)\sigma_0^2/\sigma^2$, the density of $\sum_{i=1}^n (X_i - \overline{X})^2/\sigma^2$ is $f_{n-1}(t)$, then

$$\begin{aligned} E(\hat{\sigma}_y^{2+}) &= \frac{\sigma^2}{n-1} E\Big[\frac{\sum_{i=1}^n (X_i - \overline{X})^2}{\sigma^2} - \frac{(n-1)\sigma_0^2}{\sigma^2}\Big]^+ \\ &= \frac{\sigma^2}{n-1} E\Big[\frac{\sum_{i=1}^n (X_i - \overline{X})^2}{\sigma^2} - A\Big]^+ \\ &= \frac{\sigma^2}{n-1} \int_{-\infty}^{+\infty} (t-A)^+ f_{n-1}(t)dt = \frac{\sigma^2}{n-1} \int_A^{+\infty} (t-A) f_{n-1}(t)dt \\ &= \frac{\sigma_0^2}{A} \int_A^{+\infty} (t-A) f_{n-1}(t)dt \\ &= \sigma_0^2 \int_A^{+\infty} (\frac{t}{A} - 1) f_{n-1}(t)dt. \end{aligned}$$

Note that $\int_0^{+\infty} t f_{n-1}(t)dt = E[\sum_{i=1}^n (X_i - \overline{X})^2/\sigma^2] = n-1$, then

$$\begin{aligned} E(\hat{\sigma}_y^{2+}) - \sigma_y^2 &= \sigma_0^2 \int_A^{+\infty} (\frac{t}{A} - 1) f_{n-1}(t)dt - \frac{(n-1)\sigma_0^2}{A} + \sigma_0^2 \\ &= \frac{\sigma_0^2}{A} \int_A^{+\infty} t f_{n-1}(t)dt - \frac{(n-1)\sigma_0^2}{A} + \sigma_0^2 \int_0^A f_{n-1}(t)dt \\ &= \sigma_0^2 \int_0^A f_{n-1}(t)dt - \frac{\sigma_0^2}{A} \int_0^A t f_{n-1}(t)dt \\ &= \sigma_0^2 \int_0^A (1 - \frac{t}{A}) f_{n-1}(t)dt. \end{aligned}$$

Hence,

$$0 \le E(\hat{\sigma}_y^{2+}) - \sigma_y^2 \le \sigma_0^2 \int_0^A f_{n-1}(t)dt. \tag{5.1}$$

According to the law of large numbers, as $n \to +\infty$,

$$\begin{aligned} \int_0^A f_{n-1}(t)dt &= P\{\frac{(n-1)S^2}{\sigma^2} \le A\} = P\{S^2 \le \sigma_0^2\} \\ &= P\{\sigma^2 - \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})^2 \ge \sigma_y^2\} \to 0. \end{aligned} \tag{5.2}$$

i.e. $\lim_{n\to+\infty}[E(\hat{\sigma}_y^{2+}) - \sigma_y^2] = 0$.

Using (5.1) and $F_{n-1}(t)$ is continuous at "0" uniformly for σ_y, we have $0 \le E(\hat{\sigma}_y^{2+}) - \sigma_y^2 \le \sigma_0^2 F_{n-1}(A)$ and then

$$\lim_{\sigma_y^2 \to +\infty} [E(\hat{\sigma}_y^{2+}) - \sigma_y^2] = \lim_{A\to 0} [E(\hat{\sigma}_y^{2+}) - \sigma_y^2] = 0.$$

Denote $h_1(A) = \int_A^{+\infty} (\frac{t}{A} - 1) f_{n-1}(t)dt$, $h_2(A) = \int_0^A (1 - \frac{t}{A}) f_{n-1}(t)dt$. Since $f_{n-1}(t)$ is independent of σ^2, it means $f_{n-1}(t)$ is also independent of A, we have

$$h_1'(A) = -\frac{1}{A^2}\int_A^{+\infty} t f_{n-1}(t)dt < 0, \quad h_2'(A) = \frac{1}{A^2}\int_0^A t f_{n-1}(t)dt > 0.$$

We can conclude that $E(\hat{\sigma}_y^{2+})$ is a strictly increasing function of A, so it is also a strictly increasing function of σ_y^2. Furthermore, $E(\hat{\sigma}_y^{2+}) - \sigma_y^2$ is a strictly increasing function of A, so it is a strictly decreasing function of σ_y^2.

(ii). By (5.1) and (5.2), we have

$$|E(\hat{\sigma}_y^{2+}) - \sigma_y^2| \le \sigma_0^2 P\{\sigma^2 - \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})^2 \ge \sigma_y^2\}.$$

Take $\xi = \sigma^2 - S^2$ in Lemma 2, we get

$$\begin{aligned} |E(\hat{\sigma}_y^{2+}) - \sigma_y^2| &\leq \sigma_0^2 P\{\xi - E(\xi) \geq \sigma_y^2\} \leq \frac{\sigma_0^2 Var(\xi)}{Var(\xi) + \sigma_y^4} \\ &= \frac{\sigma_0^2 V_n}{V_n + \sigma_y^4}. \quad \square \end{aligned}$$

Proof of Theorem 2.

On one hand, treat $S^2 - \sigma_0^2$ as ξ in (2.2) of Lemma 2, we get $Var[(S^2 - \sigma_0^2)^+] \leq Var(S^2 - \sigma_0^2) = V_n$, and

$$\begin{aligned} MSE(\hat{\sigma}_y^2) &= E(S^2 - \sigma_0^2 - \sigma_y^2)^2 = E[(S^2 - \sigma_0^2)^+ - (S^2 - \sigma_0^2)^- - \sigma_y^2]^2 \\ &= E[(S^2 - \sigma_0^2)^+ - \sigma_y^2]^2 - 2E[((S^2 - \sigma_0^2)^+ - \sigma_y^2)(S^2 - \sigma_0^2)^-] \\ &\quad + E[(S^2 - \sigma_0^2)^-]^2 \\ &= MSE(\hat{\sigma}_y^{2+}) + 2\sigma_y^2 E(S^2 - \sigma_0^2)^- + E[(S^2 - \sigma_0^2)^-]^2 \\ &\geq MSE(\hat{\sigma}_y^{2+}). \end{aligned}$$

On the other hand,

$$\begin{aligned} Var(\hat{\sigma}_y^{2+}) &= E[(\hat{\sigma}_y^{2+})^2] - [E(\hat{\sigma}_y^{2+})]^2 \\ &= E\Big[\frac{\sigma^2}{n-1}\Big(\frac{\sum_{i=1}^n (X_i - \overline{X})^2}{\sigma^2} - A\Big)^+\Big]^2 - [E(S^2 - \sigma_0^2)^+]^2 \\ &= \sigma_0^4 \int_A^{+\infty} \Big(\frac{t}{A} - 1\Big)^2 f_{n-1}(t)dt - \sigma_0^4 \Big[\int_A^{+\infty} \Big(\frac{t}{A} - 1\Big) f_{n-1}(t)dt\Big]^2 \\ &=: \sigma_0^4 h_3(A). \end{aligned}$$

Applying $f_{n-1}(t)$ is independent of σ^2, it is easy to get

$$\begin{aligned} h_3'(A) &= \frac{2}{A^2}\Big[-\int_A^{+\infty} t\Big(\frac{t}{A} - 1\Big) f_{n-1}(t)dt + \int_A^{+\infty} \Big(\frac{t}{A} - 1\Big) f_{n-1}(t)dt \int_A^{+\infty} t f_{n-1}(t)dt\Big] \\ &= -\frac{1}{A^2} \int_A^{+\infty}\int_A^{+\infty} \Big[\frac{t_1}{A} - 1 - \Big(\frac{t_2}{A} - 1\Big)\Big](t_1 - t_2) f_{n-1}(t_1) f_{n-1}(t_2) dt_1 dt_2 \\ &= -\frac{1}{A^3} \int_A^{+\infty}\int_A^{+\infty} (t_1 - t_2)^2 f_{n-1}(t_1) f_{n-1}(t_2) dt_1 dt_2 \\ &\leq 0, \end{aligned}$$

then $Var(\hat{\sigma}_y^{2+})$ is a decreasing function of A, so it is an increasing function of σ_y^2. □

Proof of Theorem 3.

By the law of large numbers, we have when $E(X^2) < +\infty$, $S^2 \xrightarrow{p.} \sigma^2$, and $\forall \epsilon > 0$, $P\{(S^2 - \sigma_0^2)^- > \epsilon\} \le P\{S^2 \le \sigma_0^2\} = P\{S^2 - \sigma^2 \le -\sigma_y^2\} \to 0$. Note that, $(S^2 - \sigma_0^2)^+ - \sigma_y^2 = S^2 - \sigma^2 + (S^2 - \sigma_0^2)^-$, therefore, $\hat{\sigma}_y^{2+} \xrightarrow{p.} \sigma_y^2 \quad (n \to +\infty)$. On the other hand, when $0 < Var((X-\mu)^2) < +\infty$, and $\forall \epsilon > 0$, $P\{\sqrt{n}(S^2 - \sigma_0^2)^- > \epsilon\} \le P\{S^2 \le \sigma_0^2\} \to 0$. Hence,

$$\sqrt{n}[(\hat{\sigma}_y^{2+}) - \sigma_y^2] = \sqrt{n}[S^2 - \sigma^2] + \sqrt{n}(S^2 - \sigma_0^2)^- = \sqrt{n}[S^2 - \sigma^2] + o_p(1).$$

By the central limit theorem, $\sqrt{n}(S^2 - \sigma^2) \xrightarrow{d.} N(0, Var((X-\mu)^2))$ as $n \to +\infty$, then,

$$\sqrt{n}(\hat{\sigma}_y^{2+} - \sigma_y^2) \xrightarrow{d.} N(0, Var((X-\mu)^2)). \quad \square$$

5 Acknowledgements

The research was partially supported by the Doctoral Programme of Higher Education (No: 20020027010) of China.

References

1. Fuller W. A. (1987). *Measurement Errors Models*, New York: John Wiley & Sons.

2. Serfling R. J. (1980). *Approximation Theorems of Mathematical Statistics*, New York: John Wiley & Sons.

3. Laha R. G. and Rohatgi V. K. (1979). *Probability Theory*, New York: John Wiley & Sons.