

SEQUENTIAL AND NON-SEQUENTIAL CONFIDENCE  
INTERVALS OF CONSTANT WIDTH FOR A  
SIMPLE LINEAR REGRESSION MEAN

By

JIN TZE CHEN

Bachelor of Science  
National Taiwan University  
Taipei, Taiwan  
1955

Master of Science  
Oklahoma State University  
Stillwater, Oklahoma  
1959

Submitted to the Faculty of the Graduate College  
of the Oklahoma State University  
in partial fulfillment of the requirements  
for the Degree of  
DOCTOR OF PHILOSOPHY  
May, 1974

SEQUENTIAL ANALYSIS OF SEQUENTIAL CONFIDENCE  
INTERVALS OF CONSTANT WIDTH FOR A  
SIMPLE LINEAR REGRESSION MEAN

Thesis Approved Thesis Approved Thesis Approved:

*I have read the thesis and find it acceptable*

Thesis Adviser  
Robert D. Morrison  
David J. Heub  
R. Weber  
Dean of the Graduate College

902046 902046 902046

## PREFACE

The problem considered here is to set up a confidence interval of a constant width for a simple linear regression mean at a given point which is belonging to a finite interval. The problem considered has the following property: Given the coverage probability and constant width, a rule is defined to determine the sample size so that the probability based on the determined sample size which covers a simple linear regression mean at a given point approaches a number which is no less than the preassigned coverage probability as the width goes to zero.

The author wishes to express his appreciation to his major adviser, Dr. Leroy Folks, for his suggestion of the problem studied, his guidance and his assistance in this study. Appreciation is also expressed to Dr. Lyle Broemling for his encouragement and part of the guidance before he took leave of absence to London. Thanks also to the other committee members; Dr. David Weeks, Dr. Robert Morrison, and Dr. Robert Weber for their kind suggestions for this final manuscript. Special thanks to the College Science Improvement Program (COSIP) and Central State University for their financial support for a full year of study at Oklahoma State University. Thanks is also extended to Mrs. Joyce Gazaway for her assistance in typing this manuscript.

Gratitude is expressed to my parents, Min-Da and Shu-Shieu (Chang). Without their inspiration, I could never have obtained my education.

Finally, special gratitude is expressed to my wife, Te-Hsing, our children, Jinna, Willy and Terry, for their understanding and sacrifice.

## TABLE OF CONTENTS

| Chapter  | Page |
|--|------|
| I. INTRODUCTION . . . . .  | 1    |
| Notation Used . . . . .  | 4    |
| Mathematical Model and Estimators . . . . .  | 4    |
| Description and Discussion of Taking<br>Observations on $X_1, X_2, \dots, X_m$ . . . . . | 5    |
| Summary . . . . .  | 11   |
| II. NON-SEQUENTIAL PROCEDURE . . . . .   | 12   |
| $\sigma^2$ Known, Non-Sequential Procedure . . . . .                                     | 12   |
| Comments . . . . .   | 24   |
| III. SEQUENTIAL PROCEDURE . . . . .  | 27   |
| $\sigma^2$ Unknown, Sequential Procedure,<br>Asymptotic Consistency . . . . .            | 27   |
| Comment . . . . .  | 46   |
| Asymptotic Efficiency . . . . .  | 46   |
| Conclusion . . . . .   | 56   |
| BIBLIOGRAPHY . . . . .   | 57   |
| APPENDIX . . . . .   | 59   |

## LIST OF FIGURES

| Figure  | Page |
|---|------|
| 1. Sequential Confidence Intervals of Constant Width<br>for a Simple Linear Regression Mean . . . . . | 47   |

## CHAPTER I

### INTRODUCTION

The problem considered in the present work is to set up a confidence interval of a constant width for a simple linear regression mean at a given point  $X$  which is belonging to a finite interval. The problem considered has the following property: Given the coverage probability, say  $\alpha$ , and the constant width, say  $2d$ , a rule will be defined to determine the sample size  $n$  so that the probability based on the determined sample size  $n$  which covers a simple linear regression mean at a given point  $X$  will approach a number which is no less than  $\alpha$  when  $d$  decreases. The problem will be treated in two cases: namely, when the variance of the dependent random variable  $Y$  is known and when it is unknown. A non-sequential procedure will be used to determine  $n$  when the variance is known; a sequential procedure will be used when the variance is unknown. In both cases, there is no assumption to be made about the distribution of  $Y$ , except that the variance has to be finite.

The subject of linear regression has been a popular one. A conventional method to set up a confidence interval for a regression mean at a given point can be found in almost any statistical methods textbook. The typical ones are Statistical Methods by Snedecor (1) and An Introduction to Linear Statistical Models, Volume I by Graybill (2). Further work has been done concerning the confidence bands for the

entire regression line by such prominent statisticians as Scheffé (3) Working and Hotelling (4). Then Hoel (5) considered the possibility of finding an optimal confidence band in the sense that the expected total area of an admissible class of bands is a minimum. Graybill and Bowden (6) have developed the straight line bands rather than the conventional curvilinear bands and indicated that these bands are more efficient than those curvilinear ones. Folks and Antle (7) also proved that polygon bands for general linear regression problems are more conservative than those of elliptical ones. Gafarian (8) even developed bands with constant width extended over a bounded interval. As he mentioned in his paper:

. . . ordinarily an experimenter is not interested in coverage of the whole regression curve. On the contrary, interest lies in only a bounded interval or even a finite set of points. A method for providing a band that is valid only for a finite set of interest may yield a more efficient band.

In the present work, his concept has been referred to.

There is no intention to compare the present work with those studies in the preceding paragraph. This work is merely another way to look at the problem.

Chow and Robbins (9) have used a general sequential procedure for finding a confidence interval of constant width with a given coverage probability for unknown mean  $\mu$  of a population having fixed distribution  $F$  with unknown but finite variance. Gleser (10) and Srivastava (11) have examined these results to the linear regression parameters. Here, their results will be extended to the simple linear regression mean.

A method given based on Snedecor (1) is as follows: Let  $X_1, X_2, \dots, X_n$  be a fixed set of observable points. Under the mathematical model that

$$Y_i = \beta_0 + \beta_1(X_i - \bar{X}) + e_i, i = 1, 2, 3, \dots, n$$

where  $Y_i$ 's are iid normal random variables with

$$EY_i = \beta_0 + \beta_1(X_i - \bar{X}), \text{ var } (Y_i) = \sigma^2 \text{ for all } i$$

and given the coverage probability  $\alpha$ , the fixed sample size  $n$ , the confidence interval at a given point  $X$  is given as

$$\beta_0 + \beta_1(X - \bar{X}) \pm t_{1-\alpha} \hat{\sigma} \left[ \frac{1}{n} + \frac{(X - \bar{X})^2}{nS^2} \right]^{\frac{1}{2}}$$

where

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}, \hat{\beta}_1 = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2}$$

$$\hat{\sigma}^2 = \frac{\sum (Y_i - \hat{Y})^2}{n-2} = \frac{1}{n-2} \left\{ \sum (Y_i - \bar{Y})^2 - \frac{[\sum (X_i - \bar{X})(Y_i - \bar{Y})]^2}{\sum (X_i - \bar{X})^2} \right\}$$

$$\bar{X} = n^{-1} \sum X_i, \bar{Y} = n^{-1} \sum Y_i, S^2 = n^{-1} \sum (X_i - \bar{X})^2, \hat{Y} = \hat{\beta}_0 + \hat{\beta}_1(X - \bar{X})$$

$\hat{\beta}_0$ ,  $\hat{\beta}_1$  and  $\hat{\sigma}^2$  are all unbiased maximum-likelihood estimates for  $\beta_0$ ,  $\beta_1$ , and  $\sigma^2$ .

In this method, when  $n$  and  $\alpha$  are given, the width of the confidence interval is

$$2t_{1-\alpha} \hat{\sigma} \left[ \frac{1}{n} + \frac{(X - \bar{X})^2}{nS^2} \right]^{\frac{1}{2}}$$

cannot be controlled; besides, it depends on  $X$ . The farther  $X$  is from  $\bar{X}$ , the wider the interval will be, although  $X$  can be extended from  $-\infty$  to  $\infty$  but, as Gafarian pointed out, the experimenter may not be interested. The present work considers the case that: when the coverage probability  $\alpha$  and the constant width  $2d$  are given, a sample size  $n$  will be determined by a certain rule. Based on the determined

sample size, it will be shown that the probability of covering the simple linear regression mean at a given point  $X$  by the confidence interval of constant width will be no less than  $\alpha$  when  $d$  decreases. By the restriction of the range of  $X$ , an experimenter may save his sample size as will be shown.

#### Notation Used

Let  $X_1, X_2, \dots, X_m$  be  $m$  fixed distinct observable variables where  $m \geq 2$ . Then

$$(1) \quad X = m^{-1} \sum_{i=1}^m X_i$$

$$(2) \quad [a, b] = [\bar{X} - h, \bar{X} + h]$$

where  $h$  is suitably chosen such that it has to cover  $X_1, X_2, \dots, X_m$ .

$$(3) \quad \bar{X}_n = n^{-1} \sum_{i=1}^n X_i$$

$$(4) \quad S^2 = m^{-1} \sum_{i=1}^m (X_i - \bar{X})^2$$

$$(5) \quad S_n^2 = n^{-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$$

$$(6) \quad \mu_X = \beta_0 + \beta_1 (X - \bar{X}_n)$$

$$(7) \quad X'(n) = \begin{bmatrix} 1 & 1 & 1 \\ X_1 - \bar{X}_n & X_2 - \bar{X}_n & \dots & X_n - \bar{X}_n \end{bmatrix}$$

#### Mathematical Models and Estimators

The following are the mathematical model and estimators which will be used.

(1)  $Y_i = \beta_0 + \beta_1(X_i - \bar{X}_n) + e_i$ ,  $i = 1, 2, 3, \dots$  where  $\beta_0$  and  $\beta_1$  are unknown parameters.  $Y_i$ 's are iid random variables with

$$EY_i = \beta_0 + \beta_1(X_i - \bar{X}_n), \text{ var } Y_i = \sigma^2 \text{ for } i = 1, 2, 3, \dots$$

$$(2) \hat{\mu}_X = \hat{\beta}_0 + \hat{\beta}_1(X_i - \bar{X}_n)$$

(3)  $\hat{\beta} = [X'(n)X(n)]^{-1}X(n)Y(n)$  which is the best linear unbiased estimate of  $\beta$ .

$$(4) Y'(n) = (Y_1, Y_2, \dots, Y_n)$$

$$(5) \hat{\sigma}^2(n) = N^{-1}Y'(n)\{I(n) - X(n)[X'(n)X(n)]^{-1}X'(n)\}Y(n) + n^{-1}$$

where  $I(n)$  is the  $n \times n$  identity matrix, and the additional term of  $n^{-1}$  will be explained later.

#### Description and Discussion of Taking

Observations on  $X_1, X_2, \dots, X_m$

Define the set  $\{X_1, X_2, \dots, X_m\}$  as a primary set where  $m \geq 2$ . From these points compute

$$\bar{X} = m^{-1} \sum X_i \text{ and}$$

$$S^2 = m^{-1} \sum (X_i - \bar{X})^2.$$

Then choose an appropriate  $h$  so that:

- (1)  $(\bar{X} - h, \bar{X} + h)$  covers the primary set;
- (2) the length of  $h$  serves the purpose of an experimenter's interest.

There is a reason which will be explained in the last section of Chapter II, that  $S^2$  should be as large as possible. In order to achieve this purpose, the size of  $m$  will be considered in two cases:

- (1) When the points of X's in the primary set are equally spaced;
- (2) When the points of X's in the primary set are arbitrarily spaced.

Case 1

Let  $X_1$  and  $X_m$  be the first and last points in the primary set. So,

$$X_1, X_m \in (\bar{X} - h, \bar{X} + h),$$

thus, the points of the primary set will be

$$X_1, X_1 + \frac{X_m - X_1}{m-1}, X_1 + \frac{2(X_m - X_1)}{m-1}, \dots, X_1 + \frac{(m-2)(X_m - X_1)}{m-1}, \\ X_1 + (X_m - X_1).$$

Thus, the primary set is equally spaced.

$$\bar{X} = \frac{mX_1 + \frac{m}{2}(X_m - X_1)}{m} = X_1 + \frac{X_m - X_1}{2}$$

$$S^2 = m^{-1} \sum_{k=1}^m \left\{ \left[ X_1 + \frac{k-1}{m-1}(X_m - X_1) \right] - \left[ X_1 + \frac{X_m - X_1}{2} \right] \right\}^2$$

$$= m^{-1} (X_m - X_1)^2 \sum_{k=1}^m \left[ \frac{2k - (m+1)}{2(m-1)} \right]^2$$

$$= m^{-1} \left[ \frac{(X_m - X_1)^2}{4m(m-1)^2} \right] \sum_{k=1}^m [2k - (m+1)]^2 = \frac{m+1}{12(m-1)} (X_m - X_1)^2$$

i.e.,

$$S^2 = \left[ 1 + \frac{2}{m-1} \right] \frac{(X_m - X_1)^2}{12}$$

The conclusion is

$$S_2^2 > S_3^2 > \dots > \frac{(X_m - X_1)^2}{12},$$

where  $S_k^2$  means the variance associated with a primary set of  $k$  points,  $k = 2, 3, \dots, m$ . From this result, it can be seen that when  $X_1$  and  $X_m$  are determined, more points inserted between  $X_1$  and  $X_m$ ,  $S^2$  will become smaller with the lower limit

$$\frac{(X_m - X_1)^2}{12}.$$

Therefore, fewer points should be used between  $X_1$  and  $X_m$  in the primary set whenever it is possible. Of course, the best primary set is  $\{X_1, X_m\}$ .  $S_2^2$  associated with  $\{X_1, X_m\}$  is  $\frac{1}{4}(X_m - X_1)^2$ , which is the largest.

### Case 2

Let  $X_1$  and  $X_m$  be the first and last points in the primary set,

$$X_1, X_m \in (\bar{X} - h, \bar{X} + h),$$

then

$$X_1, X_1 + t_1(X_m - X_1), X_1 + t_2(X_m - X_1), \dots, X_1 + t_{m-1}(X_m - X_1)$$

where

$$0 = t_0 < t_1 < t_2 < \dots < t_{m-1} = 1.$$

Thus, the primary set is arbitrarily spaced.

$$\bar{X} = m^{-1} \left[ mX_1 + \sum_{k=1}^m t_{k-1} (X_m - X_1) \right] = X_1 + \frac{(X_m - X_1)}{m} \sum_{k=1}^m t_{k-1}.$$

$$S^2 = m^{-1} \left\{ \sum_{k=1}^m \left[ t_{k-1}^{-m-1} \sum_{k=1}^m t_{k-1} \right]^2 \right\} (X_m - X_1)^2$$

$$= m^{-3} \left\{ \sum_{k=1}^m \left[ m t_{k-1} - \sum_{k=1}^m t_{k-1} \right]^2 \right\} (X_m - X_1)^2.$$

At this point, it was conjectured that the result would be

$$s_2^2 > s_3^2 > \dots$$

however, this is not the case in general. For an example, let

$$X_1 = 0, \text{ and } X_m = 50$$

then

$$s_2^2 = \frac{1}{2} \{ (0 - 25)^2 + (50 - 25)^2 \} = (25)^2 = 625,$$

let

$$X_1 = 0, X_2 = 1, X_m = 50$$

then

$$\begin{aligned} s_3^2 &= \frac{1}{3} \{ (0 - 17)^2 + (1 - 17)^2 + (50 - 17)^2 \} = \frac{1}{3} (17^2 + 16^2 + 33^2) \\ &= \frac{1}{3} (289 + 256 + 1089) = 544.667 \end{aligned}$$

thus

$$s_2^2 > s_3^2.$$

As another example, let

$$X_1 = 0, X_2 = 25, X_m = 50$$

$$\begin{aligned} s_3^2 &= \frac{1}{3} \{ (0 - 25)^2 + (50 - 25)^2 \} = \frac{1}{3} \{ (-25)^2 + (25)^2 \} = \frac{1}{3} (1250) \\ &= 416.667. \end{aligned}$$

Let

$$X_1 = 0, X_2 = 1, X_3 = 5, X_m = 50$$

$$\begin{aligned} s_3^2 &= \frac{1}{4} \{ (0 - 14)^2 + (1 - 14)^2 + (5 - 14)^2 + (50 - 14)^2 \} \\ &= \frac{1}{4} \{ (14)^2 + (13)^2 + (9)^2 + (36)^2 \} = \frac{196 + 169 + 81 + 1296}{4} \\ &= \frac{1742}{4} = 435.5. \end{aligned}$$

So

$$s_3^2 < s_4^2.$$

Thus, when the primary set is arbitrarily spaced, there is no trend in general such as

$$s_2^2 > s_3^2 > s_4^2 \dots$$

or

$$s_2^3 < s_3^2$$

especially when the points in the primary set clustered to either  $X_1$  or  $X_m$ . So, it is suggested that when several primary sets with the same  $X_1$  and  $X_m$  are possible for an experiment, compute each  $S^2$  associated with its primary set and use the primary set which yields the largest  $S^2$ .

In the present work, the sample size  $n$  is not predetermined. It is determined by a rule, which will be defined, based on the given coverage probability  $\alpha$  and the constant width  $2d$  for a confidence interval. The sample size  $n$  thus determined could not always be the same as  $m$ , the number of points in the primary set. In case  $n > m$ , observation must be repeated on the primary set. In the following a description is given and is explained as to how the primary set should be observed. Define the actual set  $\{X_1, X_2, \dots, X_n\}$  observed as the observed set. For example, the observed set could be

$$\{X_1, X_2, \dots, X_m, X_1, X_2\}$$

or

$$\{X_1, X_2, X_m\}$$

The object of the description concerns the order of observations on the

primary set. It is always desirable to make the variance of the estimate

$$\hat{\mu}_X = \hat{\beta}_0 + \hat{\beta}_1(X - \bar{X}_n)$$

for

$$\mu_X = \beta_0 + \beta_1(X - \bar{X}_n)$$

small.

$$\begin{aligned} \text{var} [\hat{\beta}_0 + \hat{\beta}_1(X - \bar{X}_n)] &= \text{var} \hat{\beta}_0 + (X - \bar{X}_n)^2 \text{var} \hat{\beta}_1 + 2(X - \bar{X}_n) \\ &\text{cov} (\hat{\beta}_0, \hat{\beta}_1). \end{aligned}$$

But

$$\text{cov} \hat{\beta} = E [(\hat{\beta} - \beta)(\hat{\beta} - \beta)'] = \sigma^2 [X'(n)X(n)]^{-1} = \begin{bmatrix} \frac{\sigma^2}{n} & 0 \\ 0 & \frac{2}{nS_n^2} \end{bmatrix}$$

So,

$$\text{var} [\hat{\beta}_0 + \hat{\beta}_1(X - \bar{X}_n)] = \frac{\sigma^2}{n} + \frac{\sigma^2}{nS_n^2}(X - \bar{X}_n)^2 = \sigma^2 \left[ \frac{1}{n} + \frac{(X - \bar{X}_n)^2}{\sum (X_i - \bar{X}_n)^2} \right]$$

In order to make

$$\text{var} [\hat{\beta}_0 + \hat{\beta}_1(X - \bar{X}_n)]$$

small, the quantity

$$\sum_{i=1}^n (X_i - \bar{X}_n)^2$$

has to be maximized. It is suggested that points of the primary set should be observed in the following order:

$$X_1, X_m, X_2, X_{m-1}, \dots$$

When the primary set has been exhausted, repeat the order as before. Once the order of observing  $X$ 's has been defined, the matrix  $X(n)$  associated with such a sampling will not be a random matrix for a determined  $n$ .

### Summary

Let  $X_1, X_1, \dots, X_m$  be a predetermined set, the primary set. Choose an appropriate  $h$  so that  $(\bar{X} - h, \bar{X} + h)$  will cover the primary set and, if possible, make  $h$  as short as possible.

A rule will be defined based on a coverage probability  $\alpha$  and a constant width  $2d$  for the confidence interval to determine the sample size  $n$ . Based on the determined  $n$ , set up the interval

$$I_n = [\hat{\beta}_0 + \hat{\beta}_1(X - \bar{X}_n) - d, \hat{\beta}_0 + \hat{\beta}_1(X - \bar{X}_n) + d]$$

it will be shown that

$$\lim_{d \rightarrow 0} P[\beta_0 + \beta_1(X - \bar{X}_n) \in I_n] \geq \alpha,$$

where

$$X \in (\bar{X} - h, \bar{X} + h).$$

This means that, when  $\alpha$  and  $d$  are given, the probability based on the determined size  $n$  which covers a simple linear regression mean at a given point  $X$  will approach a number which is no less than  $\alpha$  when  $d$  goes to zero.

## CHAPTER II

### NON-SEQUENTIAL PROCEDURE

As stated, a confidence interval of constant width  $2d$  for a simple linear regression mean

$$\mu_X = \beta_0 + \beta_1(X - \bar{X}_n)$$

at  $X \in (\bar{X} - h, \bar{X} + h)$  with a preassigned coverage probability  $\alpha$  will be set up which has the following property:

$$\lim_{d \rightarrow 0} P\{\mu_X \in I_n\} \geq \alpha$$

where

$$I_n = \{\hat{\beta}_0 + \hat{\beta}_1(X - \bar{X}_n) - d, \hat{\beta}_0 + \hat{\beta}_1(X - \bar{X}_n) + d\}$$

$\sigma^2$  Known, Non-Sequential Procedure

In the following, a rule is defined to determine the sample size  $n$ . Based on  $n$ , the confidence interval  $I_n$  with the constant width  $2d$  is constructed.

Choose a number  $a$ , so that

$$(1) \quad (2\pi)^{-\frac{1}{2}} \int_{\frac{-a}{\sqrt{1 + \left(\frac{h}{S}\right)^2}}}^{\frac{a}{\sqrt{1 + \left(\frac{h}{S}\right)^2}}} e^{-\frac{u^2}{2}} du = \alpha$$

where  $h$  and  $S^2$  were mentioned as before. Define the sample size  $n$  as

the smallest positive integer so that

$$(2) \quad n \geq \frac{a^2 \sigma^2}{d^2}$$

thus,  $n$  is uniquely determined when  $\alpha$  and  $d$  are given. From the rule as defined, there is a relationship between  $n$  and  $d$  which can be written down as

Lemma 1.  $d \rightarrow 0$  as  $n \rightarrow \infty$ .

Proof: As defined,  $n$  is the smallest positive integer so that

$$n \geq \frac{a^2 \sigma^2}{d^2}.$$

Thus,

$$n = \frac{a^2 \sigma^2}{d^2} + r, \quad 0 \leq r < 1$$

$$\infty = \lim_{n \rightarrow \infty} (n) = \lim_{n \rightarrow \infty} \left( \frac{a^2 \sigma^2}{d^2} \right) = a^2 \sigma^2 \lim_{n \rightarrow \infty} d^{-2}$$

$$\lim_{n \rightarrow \infty} d = 0.$$

To develop further, the following lemmas are needed.

Lemma 2.

$$\lim_{n \rightarrow \infty} \left( \frac{nd^2}{\sigma^2} \right) = a^2.$$

Proof: Since

$$n = \frac{a^2 \sigma^2}{d^2} + r, \quad 0 \leq r < 1$$

$$nd^2 = a^2 \sigma^2 + rd^2, \quad \frac{nd^2}{\sigma^2} = a^2 + \frac{rd^2}{\sigma^2}$$

$$\lim_{n \rightarrow \infty} \left( \frac{nd^2}{\sigma^2} \right) = \lim_{d \rightarrow 0} \left( \frac{nd^2}{\sigma^2} \right) = \lim_{d \rightarrow 0} \left( a^2 + \frac{rd^2}{\sigma^2} \right) = a^2.$$

or

$$\lim_{n \rightarrow \infty} \left[ \frac{\sqrt{nd}}{\sigma} \right] = a.$$

Lemma 3.

$$\lim_{n \rightarrow \infty} \bar{X}_n = \bar{X}, \quad \lim_{n \rightarrow \infty} S_n^2 = S^2.$$

Proof: Since

$$\bar{X}_n = n^{-1} \sum_{i=1}^n X_i = \frac{c \sum_{i=1}^m X_i + \text{Res}}{n} = \frac{c \sum_{i=1}^m X_i + \text{Res}}{mc + r}$$

where  $0 \leq r < 1$ ,  $c$  stands for the number of cycles repeating the primary set and Res stands for residue, which is a part of the primary set.

$$\bar{X}_n = \frac{\sum_{i=1}^m X_i + \frac{\text{Res}}{c}}{m + \frac{r}{c}}$$

$$\lim_{n \rightarrow \infty} \bar{X}_n = \lim_{c \rightarrow \infty} \bar{X}_n = m^{-1} \sum_{i=1}^m X_i = \bar{X}.$$

$$S_n^2 = n^{-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2 = n^{-1} \left[ \sum_{i=1}^n X_i^2 - n \bar{X}_n^2 \right] = n^{-1} \sum_{i=1}^n X_i^2 - \bar{X}_n^2$$

$$= \frac{c \sum_{i=1}^m X_i^2 + \text{Res}}{mc + r} - \bar{X}_n^2$$

$$\lim_{n \rightarrow \infty} S_n^2 = \lim_{c \rightarrow \infty} S_n^2 = \lim_{c \rightarrow \infty} \left[ \frac{\sum_{i=1}^m X_i^2 + \frac{\text{Res}}{c}}{m + \frac{r}{c}} - \bar{X}_n^2 \right]$$

$$= m^{-1} \sum_{i=1}^m X_i^2 - \bar{X}^2 = S^2$$

since

$$\lim_{c \rightarrow \infty} \bar{X}_n = \bar{X}.$$

Lemma 4 [Gredenko and Kolmogorov (12)].

If  $Z_1, Z_2, Z_3, \dots$  are iid random variables with  $Z_1 = 0$ ,  $\text{var } Z_1 = 1$ , and if  $\{b_{ni}\}$

$$i = 1, 2, 3, \dots, n; n = 1, 2, 3, \dots$$

is a fixed array of constants so that

$$\sum_{i=1}^n b_{ni}^2 = 1, 2, 3, \dots$$

$$\max |b_{ni}| \rightarrow 0 \text{ as } n \rightarrow \infty$$

then

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n b_{ni} Z_{ni} = N(0, 1) \text{ in distribution.}$$

With the provision of lemma 4, the following lemma is proved.

Lemma 5. Let

$$L(n, \lambda) = \lambda_0 \left[ \frac{\hat{\beta}_0 - \beta_0}{\frac{\sigma}{n}} \right] + \lambda_1 \left[ \frac{\hat{\beta}_2 - \beta_2}{\frac{\sigma}{\sqrt{ns_n^2}}} \right]$$

where

$$\lambda_0^2 + \lambda_1^2 = 1$$

then

$$L(n, \lambda) \rightarrow N(0, 1) \text{ in distribution as } n \rightarrow \infty.$$

Proof:

$$L(n, \lambda) = \lambda_0 \sigma^{-1} \sqrt{n} (\hat{\beta}_0 - \beta_0) + \lambda_1 \sigma^{-1} \sqrt{ns_n^2} (\hat{\beta}_1 - \beta_1)$$

$$= \sigma^{-1} \lambda' [X'(n)X(n)]^{\frac{1}{2}} (\hat{\beta} - \beta), \lambda' = (\lambda_0, \lambda_1)$$

$$\hat{\beta} = [X'(n)X(n)]^{-1} X'(n)Y(n),$$

$$Y(n) = X(n)\beta + e, EY(n) = X(n)\beta$$

$$X'(n)EY(n) = X'(n)X(n)\beta.$$

Since  $[X'(n)X(n)]$  is a positive definite matrix, so it is nonsingular and  $[X'(n)X(n)]^{-1}$  exists.

$$\beta = [X'(n)X(n)]^{-1} X'(n)EY(n)$$

$$\hat{\beta} - \beta = [X'(n)X(n)]^{-1} X'(n)[Y(n) - EY(n)]$$

$$L(n, \lambda) = \sigma^{-1} \lambda' [X'(n)X(n)]^{-\frac{1}{2}} X'(n)[Y(n) - EY(n)]$$

$$= \lambda' [X'(n)X(n)]^{-\frac{1}{2}} X'(n)Z(n)$$

where

$$Z(n) = \sigma^{-1}[Y(n) - EY(n)], Z'(n) = (Z_1, Z_2, \dots, Z_n).$$

Note that  $Z_i$ 's are iid random variables with

$$EZ_i = 0, EZ_i^2 = 1$$

$$i = 1, 2, 3, \dots$$

$$X'(n)X(n) = \begin{bmatrix} n & 0 \\ 0 & nS_n^2 \end{bmatrix}$$

$$[X'(n)X(n)]^{\frac{1}{2}}$$

is defined as

$$\begin{bmatrix} \sqrt{n} & 0 \\ 0 & \sqrt{nS_n^2} \end{bmatrix}$$

$$[X'(n)X(n)]^{-\frac{1}{2}}$$

is defined as the inverse of

$$[X'(n)X(n)]^{-\frac{1}{2}}.$$

Now the coefficients of the  $Z_i$ 's are the components of

$$\lambda' [X'(n)X(n)]^{-\frac{1}{2}} X'(n)$$

which can be written as

$$\begin{aligned} & n^{-\frac{1}{2}} \lambda' \{n^{-1} [X'(n)X(n)]^{-\frac{1}{2}} X'(n) \\ &= n^{-\frac{1}{2}} \lambda' \begin{bmatrix} 1 & 0 \\ 0 & S_n^2 \end{bmatrix}^{-\frac{1}{2}} \begin{bmatrix} 1 & 1 & & 1 \\ X_1 - \bar{X}_n & X_2 - \bar{X}_n & \dots & X_n - \bar{X}_n \end{bmatrix} \\ &= n^{-\frac{1}{2}} \lambda' \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{S_n^2} \end{bmatrix} \begin{bmatrix} \frac{1}{2} \begin{bmatrix} 1 & 1 & & 1 \\ X_1 - \bar{X}_n & X_2 - \bar{X}_n & \dots & X_n - \bar{X}_n \end{bmatrix} \\ &= n^{-\frac{1}{2}} \lambda' \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{S_n} \end{bmatrix} \begin{bmatrix} 1 & 1 & & 1 \\ X_1 - \bar{X}_n & X_2 - \bar{X}_n & \dots & X_n - \bar{X}_n \end{bmatrix} \\ &= \left[ n^{-\frac{1}{2}} \lambda_0 + (nS_n^2)^{-\frac{1}{2}} \lambda_1 (X_1 - \bar{X}_n), \dots, n^{-\frac{1}{2}} \lambda_0 + (nS_n^2)^{-\frac{1}{2}} \lambda_1 (X_n - \bar{X}_n) \right] \\ &= [b_{n1}, b_{n2}, b_{n3}, \dots, b_{nn}] \end{aligned}$$

where

$$\begin{aligned} b_{ni} &= n^{-\frac{1}{2}} \lambda_0 + (nS_n^2)^{-\frac{1}{2}} \lambda_1 (X_i - \bar{X}_n) \\ i &= 1, 2, \dots, n; n = 1, 2, 3, \dots \end{aligned}$$

Thus

$$L(n, \lambda) = \sum_{i=1}^n b_{ni} Z_i.$$

In order to show that

$$L(n, \lambda) \rightarrow N(0, 1) \text{ in distribution as } n \rightarrow \infty$$

the conditions of lemma 4 need to be checked.

$$\begin{aligned} \sum_{i=1}^n b_{ni}^2 &= \sum_{i=1}^n \left[ n^{-\frac{1}{2}} \lambda_0 + (nS_n)^{-\frac{1}{2}} \lambda_1 (X_i - \bar{X}_n) \right]^2 \\ &= \lambda_0^2 + (nS_n^2)^{-1} \lambda_1^2 \sum_{i=1}^n (X_i - \bar{X}_n)^2 + 2(nS_n)^{-1} \lambda_0 \lambda_1 \sum_{i=1}^n (X_i - \bar{X}_n) \\ &= \lambda_0^2 + \lambda_1^2 = \lambda' \lambda = 1. \end{aligned}$$

$$\max_i |b_{ni}| = \max_i \left| n^{-\frac{1}{2}} \lambda_0 + (nS_n^2)^{-\frac{1}{2}} \lambda_1 (X_i - \bar{X}_n) \right|$$

Since as  $n \rightarrow \infty$ ,  $\bar{X}_n \rightarrow \bar{X}$ ,  $S_n^2 \rightarrow S^2$  from Lemma 3 where  $\bar{X}$ ,  $S^2$  and all  $X_i$ 's are finite quantities,

$$\max_n |b_{ni}| \rightarrow 0 \text{ as } n \rightarrow \infty.$$

The conditions of Lemma 4 are all satisfied; hence we conclude that

$$L(n, \lambda) \rightarrow N(0, 1) \text{ in distribution as } n \rightarrow \infty.$$

Lemma 6. Let

$$L(n) = \frac{\hat{\beta}_0 - \beta_0}{\frac{\sigma}{\sqrt{n}}} + \left[ \frac{X - \bar{X}}{S} \right] \left[ \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\frac{\sigma^2}{nS_n^2}}} \right]$$

then

$$L(n) \rightarrow N \left[ 0, 1 + \left( \frac{X - \bar{X}}{S} \right)^2 \right] \text{ in distribution as } n \rightarrow \infty.$$

Proof: Let

$$\lambda_0 = \left[ 1 + \left( \frac{X - \bar{X}}{S} \right)^2 \right]^{-\frac{1}{2}}$$

$$\lambda_1 = \left[ \frac{X - \bar{X}}{S} \right] \left[ 1 + \left( \frac{X - \bar{X}}{S} \right)^2 \right]^{-\frac{1}{2}}$$

thus

$$\begin{aligned} \lambda' \lambda &= \lambda_0^2 + \lambda_1^2 = \left[ 1 + \left( \frac{X - \bar{X}}{S} \right)^2 \right]^{-1} + \left[ \frac{X - \bar{X}}{S} \right]^2 \left[ 1 + \left( \frac{X - \bar{X}}{S} \right)^2 \right]^{-1} \\ &= \left[ 1 + \left( \frac{X - \bar{X}}{S} \right)^2 \right] \left[ 1 + \left( \frac{X - \bar{X}}{S} \right)^2 \right]^{-1} = 1. \end{aligned}$$

So

$$\left[ 1 + \left( \frac{X - \bar{X}}{S} \right)^2 \right]^{-\frac{1}{2}} \left[ \frac{\hat{\beta}_0 - \hat{\beta}_0}{\frac{\sigma}{\sqrt{n}}} + \left( \frac{X - \bar{X}}{S} \right) \left[ \frac{\hat{\beta}_1 - \hat{\beta}_1}{\sqrt{nS_n^2}} \right] \right] \rightarrow N(0,1)$$

in distribution as  $n \rightarrow \infty$ .

But as it is known that if  $X_n \rightarrow X$  in distribution where  $X \sim N(0,1)$  then  $cX_n \rightarrow N(0, c^2)$  in distribution where  $c$  is a constant. Proof is given in the following: Let  $X \sim N(0,2)$ , then its characteristic function is

$$\phi_X(t) = e^{-\frac{1}{2}t^2}.$$

Since

$$X_n \rightarrow X, \phi_{X_n}(t) \rightarrow e^{-\frac{1}{2}t^2} \text{ as } n \rightarrow \infty$$

$$\phi_{cX_n}(t) = Ee^{itcX_n} = \phi_{X_n}(ct)$$

$$\lim_n \phi_{cX_n}(t) = \lim_n \phi_{X_n}(ct) = e^{-\frac{1}{2}c^2t^2}$$

$$cX_n \rightarrow N(0, c^2) \text{ in distribution as } n \rightarrow \infty.$$

Let

$$c = \left[ 1 + \left( \frac{X - \bar{X}}{S} \right)^2 \right]^{\frac{1}{2}},$$

thus

$$L_n = \left[ \frac{\hat{\beta}_0 - \beta_0}{\frac{\sigma}{\sqrt{n}}} + \left( \frac{X - \bar{X}}{S} \right) \frac{\hat{\beta}_1 - \beta_1}{\frac{\sigma}{\sqrt{ns_n^2}}} \right] \rightarrow N \left[ 0, 1 + \left( \frac{X - \bar{X}}{S} \right)^2 \right]$$

in distribution as  $n \rightarrow \infty$ .

Lemma 7 [Rao (13)].

Let  $\{X_n, Y_n\}$ ,  $n = 1, 2, \dots$  be a sequence of pairs of variables. If  $|X_n - Y_n| \rightarrow 0$  in probability and  $Y_n \rightarrow Y$  in distribution, then  $X_n \rightarrow Y$  in distribution. Lemma 7 is used to be used to show the following.

Lemma 8.

$$\frac{\hat{\beta}_0 - \beta_0}{\frac{\sigma}{\sqrt{n}}} + \left( \frac{X - \bar{X}_n}{S_n} \right) \frac{\hat{\beta}_1 - \beta_1}{\frac{\sigma}{\sqrt{ns_n^2}}} \rightarrow N \left[ 0, 1 + \left( \frac{X - \bar{X}}{S} \right)^2 \right]$$

in distribution as  $n \rightarrow \infty$ .

Proof: Let

$$L_1(n) = \frac{\hat{\beta}_0 - \beta_0}{\frac{\sigma}{\sqrt{n}}} + \left( \frac{X - X_n}{S_n} \right) \frac{\hat{\beta}_1 - \beta_1}{\frac{\sigma}{\sqrt{ns_n^2}}}$$

and

$$L_2(n) = \frac{\hat{\beta}_0 - \beta_0}{\frac{\sigma}{\sqrt{n}}} + \left( \frac{X - \bar{X}}{S} \right) \frac{\hat{\beta}_1 - \beta_1}{\frac{\sigma}{\sqrt{ns_n^2}}}$$

then

$$[L_1(n) - L_2(n)] = \left[ \left( \frac{X - \bar{X}_n}{S_n} - \frac{X - \bar{X}}{S} \right) \frac{\hat{\beta}_1 - \beta_1}{\frac{\sigma}{\sqrt{ns_n^2}}} \right]$$

Let  $\varepsilon > 0$  be an arbitrary number, and let

$$U_n = \left( \frac{X - \bar{X}_n}{S_n} - \frac{X - \bar{X}}{S} \right) \left[ \frac{\hat{\beta}_1 - \beta_1}{\sqrt{nS_n^2}} \right]$$

$$P \left[ |L_1(n) - L_2(n)| < \varepsilon \right] = P \left[ |U_n| < \varepsilon \right].$$

By lemma 5, let  $\lambda_0 = 0$ ,  $\lambda_1 = 1$ , then

$$\frac{\hat{\beta}_1 - \beta_1}{\sqrt{nS_n^2}} \rightarrow N(0,1) \text{ in distribution as } n \rightarrow \infty$$

and

$$\left( \frac{X - \bar{X}_n}{S_n} - \frac{X - \bar{X}}{S} \right) \rightarrow 0 \text{ as } n \rightarrow \infty.$$

But by a theorem from Rao (12), let  $\{X_n, Y_n\}$ ,  $n = 1, 2, 3, \dots$  be a sequence of pairs of random variables, if  $X_n \rightarrow X$  in distribution,  $Y_n \rightarrow 0$  in probability then  $X_n Y_n \rightarrow 0$  in probability. Therefore,  $U_n \rightarrow 0$  in probability which in turn implies that  $U_n \rightarrow 0$  in distribution. Let  $F_n$  be the distribution function of  $U_n$ , and  $F$  be the limiting distribution of  $F_n$ ,

$$F(u) = 1 \text{ if } u = 0$$

$$F(u) = 0 \text{ if } u \neq 0.$$

$$P \left[ |L_1(n) - L_2(n)| < \varepsilon \right] = P \left[ |U_n| < \varepsilon \right] = F_n(\varepsilon) - F_n(-\varepsilon)$$

$$\lim_n P \left[ |L_1(n) - L_2(n)| < \varepsilon \right] = \lim_n \left[ F_n(\varepsilon) - F_n(-\varepsilon) \right]$$

$$= F(\varepsilon) - F(-\varepsilon) = 0.$$

Hence,

$$\lim_n P \left[ |L_1(n) - L_2(n)| > \varepsilon \right] = 0 \text{ for all } \varepsilon > 0.$$

So by Lemma 7,

$$L_1(n) \rightarrow N\left[0, 1 + \left(\frac{X - \bar{X}}{S}\right)^2\right] \text{ in distribution as } n \rightarrow \infty$$

since as Lemma 6 indicated

$$L_2(n) \rightarrow N\left[0, 1 + \left(\frac{X - \bar{X}}{S}\right)^2\right] \text{ in distribution as } n \rightarrow \infty$$

or

$$\left[ \frac{\hat{\beta}_0 - \beta_0}{\frac{\sigma}{\sqrt{n}}} \right] + \left( \frac{X - X_n}{S_n} \right) \left[ \frac{\hat{\beta}_1 - \beta_1}{\frac{\sigma}{\sqrt{nS_n^2}}} \right] \rightarrow N\left[0, 1 + \left(\frac{X - X}{S}\right)^2\right]$$

in distribution as  $n \rightarrow \infty$ .

One more lemma is needed before stating the main theorem which has been described at the beginning of this chapter.

Lemma 9.

Let  $\{b_n\}$  be a sequence of constants so that  $b_n \rightarrow b$  as  $n \rightarrow \infty$ . Let  $\{X_n\}$  be a sequence of random variables so that  $X_n \rightarrow X$  in distribution as  $n \rightarrow \infty$ .

Let  $G_n(n)$  be the distribution function of  $X_n$  and  $G(n)$  the limiting distribution function of  $X$ . Then  $G_n(b_n)$  converges to  $G(b)$ .

Proof:

Since  $b_n \rightarrow b$ , so for all  $\epsilon > 0$  there exists a positive integer  $N(\epsilon)$ , so that when  $n > N(\epsilon)$

$$G_n(b - \epsilon) \leq G(b_n) \leq G(b + \epsilon)$$

$$\lim_{n \rightarrow \infty} G_n(b - \epsilon) \leq \liminf_{n \rightarrow \infty} G_n(b_n) \leq \limsup_{n \rightarrow \infty} G_n(b_n) \leq \lim_{n \rightarrow \infty} G_n(b + \epsilon)$$

$$G(b + \epsilon) \leq \liminf_{n \rightarrow \infty} G_n(b_n) \leq \limsup_{n \rightarrow \infty} G_n(b_n) \leq G(b + \epsilon).$$

Since  $\epsilon$  is arbitrary

$$\lim_{n \rightarrow \infty} G_n(b_n) = G(b).$$

Theorem.

$$\lim_{d \rightarrow 0} P\{\mu_x \in I_n\} \geq \alpha$$

Proof:

$$\begin{aligned} P\{\mu_x \in I_n\} &= P\{|\hat{\mu}_x - \mu_x| < d\} \\ &= P\{|\hat{\beta}_0 - \beta_0 + (\hat{\beta}_1 - \beta_1)(X - \bar{X}_n)| < d\} \\ &= P\left\{\left|\frac{\hat{\beta}_0 - \beta_0}{\frac{\sigma}{\sqrt{n}}} + \frac{\hat{\beta}_1 - \beta_1}{\frac{\sigma}{\sqrt{nS_n^2}}} \frac{X - \bar{X}_n}{S_n}\right| < \frac{\sqrt{nd}}{\sigma}\right\} \end{aligned}$$

Let

$$V_n = \frac{\hat{\beta}_0 - \beta_0}{\frac{\sigma}{\sqrt{n}}} + \frac{\hat{\beta}_1 - \beta_1}{\frac{\sigma}{\sqrt{nS_n^2}}} \frac{X - \bar{X}_n}{S_n},$$

$F_n$  be the distribution function of  $V_n$  and  $F$  be the limiting distribution function of  $F_n$ . Note,  $F$  is the normal distribution function with mean 0 and variance

$$1 + \left(\frac{X - \bar{X}}{S}\right)^2.$$

Thus

$$\begin{aligned} P\{\mu_x \in I_n\} &= P\left\{|V_n| < \frac{\sqrt{nd}}{\sigma}\right\} = F_n\left[\frac{\sqrt{nd}}{\sigma}\right] - F_n\left[-\frac{\sqrt{nd}}{\sigma}\right] \\ \lim_n P\{\mu_x \in I_n\} &= \lim_n \left\{F_n\left(\frac{\sqrt{nd}}{\sigma}\right) - F_n\left(-\frac{\sqrt{nd}}{\sigma}\right)\right\} \\ &= F(a) - F(-a) \text{ by lemma 9.} \end{aligned}$$

$$\begin{aligned}
&= (2\pi)^{-\frac{1}{2}} \int \frac{\frac{a}{\sqrt{1 + \left(\frac{X - \bar{X}}{S}\right)^2}}}{\frac{-a}{\sqrt{1 + \left(\frac{X - \bar{X}}{S}\right)^2}}} e^{-\frac{u^2}{2}} du \text{ by Lemma 8} \\
&\geq (2\pi)^{-\frac{1}{2}} \int \frac{\frac{a}{\sqrt{1 + \left(\frac{h}{S}\right)^2}}}{\frac{-a}{\sqrt{1 + \left(\frac{h}{S}\right)^2}}} e^{-\frac{u^2}{2}} du = \alpha
\end{aligned}$$

by the definition of  $a$  and since  $(X - \bar{X})^2 < h^2$ . Finally,

$$\lim_{n \rightarrow \infty} P\{\mu_{\mathbf{x}} \in I_n\} \geq \alpha.$$

From Lemma 1, it can be rewritten as

$$\lim_{d \rightarrow 0} P\{\mu_{\mathbf{x}} \in I_n\} \geq \alpha.$$

#### Comments

In the introduction it was mentioned that  $S^2$  associated with a primary set should be as large as possible and  $h$  be as small as possible.

$$(2\pi)^{-\frac{1}{2}} \int \frac{\frac{a}{\left[1 + \left(\frac{h}{S}\right)^2\right]^{\frac{1}{2}}}}{\frac{-a}{\left[1 + \left(\frac{h}{S}\right)^2\right]^{\frac{1}{2}}}} e^{-\frac{u^2}{2}} du = \alpha,$$

Let

$$Z_{\frac{1-\alpha}{2}} = \frac{a}{\sqrt{1 + \left(\frac{h}{S}\right)^2}}$$

thus  $Z_{\frac{1-\alpha}{2}}$  is a constant when  $\alpha$  is given

$$a^2 = Z_{\frac{1-\alpha}{2}}^2 \left[ 1 + \frac{h^2}{S^2} \right].$$

Recall that the sample size  $n$  is determined as the smallest positive integer  $n$  so that

$$n \geq \frac{a^2 \sigma^2}{d^2}$$

or

$$n = \frac{a^2 \sigma^2}{d^2} + r, \quad 0 \leq r < 1$$

$$n = \frac{\sigma^2}{d^2} Z_{\frac{1-\alpha}{2}}^2 \left[ 1 + \frac{h^2}{S^2} \right] + r.$$

Since

$$\sigma^2, d^2, Z_{\frac{1-\alpha}{2}}^2$$

and  $r$  are all constants,  $n$  can be considered as a function of  $h$  and  $S^2$  and can be written as

$$n = n(h, S^2) = \frac{\sigma^2}{d^2} Z_{\frac{1-\alpha}{2}}^2 \left[ 1 + \frac{h^2}{S^2} \right] + r$$

$$h > 0, S^2 > 0.$$

From the function  $n(h, S^2)$  it can be seen that under the same  $\alpha$ ,  $n$  decreases when either  $h$  decreases or  $S^2$  increases alone or  $h$  decreases

and  $S^2$  increases. Thus, when  $h$  is shorter and/or  $S^2$  is longer, the sample size  $n$  will be smaller. This implies that a smaller sample size could be used if  $h$  and  $S^2$  could be adjusted properly.

In particular, when  $h$  approaches to zero, the sample size  $n$  will approach

$$\frac{\sigma^2 Z_{\frac{1-\alpha}{2}}^2}{d^2} + r, \quad 0 \leq r < 1$$

which coincides with the elementary method for determining a sample size  $n$  when the length of confidence interval  $2d$  is preassigned.

## CHAPTER III

### SEQUENTIAL PROCEDURE

In Chapter II the case that  $\sigma^2$  is known has been assumed. A rule was defined to determine the sample size  $n$  based on a given  $\alpha$  and a constant width  $2d$ . Since  $\sigma^2$  is known and  $a$  is a fixed constant, the rule was defined that the sample size  $n$  is the smallest positive integer so that

$$n \geq \frac{a^2 \sigma^2}{d^2}$$

consequently  $n$  can be determined by a non-sequential procedure.

In the present chapter the case will be considered that  $\sigma^2$  is unknown. Since an estimate  $\hat{\sigma}^2(n)$  for  $\sigma^2$  has to be used in replacing  $\sigma^2$  for defining a rule to determine the sample size  $n$ , a sequential procedure has to be used.

#### $\sigma^2$ Unknown, Sequential Procedure

##### Asymptotic Consistency

As defined,

$$(1) \quad \hat{\sigma}^2(n) = n^{-1} Y'(n) \{I(n) - X(n) [X'(n) X(n)]^{-1} X'(n)\} Y(n) + n^{-1}.$$

The purpose of the additional term  $n^{-1}$  in (1) is to ensure that

$$\frac{\hat{\sigma}^2(n)}{\sigma^2} > 0 \text{ a.s.}$$

The positiveness of the quantity

$$\frac{\hat{\sigma}^2(n)}{\sigma^2}$$

shall be used in later proofs. In case the distribution of the random variable  $Y$  is continuous,

$$n^{-1}Y'(n)\{I(n)[X'(n)X(n)]^{-1}X'(n)\}Y(n) = V^2(n)$$

say, will not vanish, so

$$\frac{V^2(n)}{\sigma^2} > 0$$

can be insured. So, when the distribution of  $Y$  is continuous,  $V^2(n)$  serves the same purpose as  $\hat{\sigma}^2(n)$  does. In the present case, the distribution of  $Y$  is an arbitrary one; it can be either continuous or discrete.

The stopping rule is defined as follows: Start by taking  $n_0 \geq 2$  observations, so the  $n_0$  observations are

$$Y_1, Y_2, \dots, Y_{n_0}.$$

Then, sample one more observation at a time and stop when

(2)  $N =$  smallest positive integer  $k \geq n_0$  such that

$$\hat{\sigma}^2(k) \leq \frac{kd^2}{a_k^2}$$

where  $N$  is a positive integer valued random variable and

$$(3) \quad a_N^2 = a^2 \left[ 1 + \frac{|X - \bar{X}_N|}{\sqrt{NS_N^2}} \right]^2$$

From (3) it can be seen that

$$\lim_{N \rightarrow \infty} a_N^2 = a^2.$$

Based on (2) and (3) it will be shown that

$$(4) \quad \lim_{d \rightarrow 0} P\{\hat{\mu}_X \in I_N\} \geq \alpha$$

where  $\alpha$  is preassigned and

$$I_N = (\hat{\mu}_X - d, \hat{\mu}_X + d).$$

Chow and Robbins call the property (4) the "asymptotic consistency" property.

But, first of all, it is needed to show that the stopping rule defined as (2) will lead the procedure to a stop. The following lemmas are needed.

Lemma 1 [Chow and Gleser (10)]. Let

$$Z_1, Z_2, Z_3, \dots$$

be iid random variables with

$$EZ_i = 0 \text{ and } EZ_i^2 = \sigma^2 \text{ for } i = 1, 2, 3, \dots$$

Let  $b_{mn}$  be any array of real numbers,

$$m \leq n, n = 1, 2, \dots,$$

so that

$$\lim_{n \rightarrow \infty} \sum_{m=1}^n b_{mn}^2 = 1$$

then

$$-\frac{1}{2} \sum_{m=1}^n b_{mn} Z_m \rightarrow 0 \text{ a.s.}$$

Lemma 2 [Gleser (10)],

$$\hat{\sigma}^2(n) \rightarrow \sigma^2 \text{ a.s.}$$

Proof:

$$\begin{aligned} \hat{\sigma}^2(n) &= n^{-1} Y'(n) \{I(n) - X(n) [X'(n) X(n)]^{-1} X'(n)\} Y(n) + n^{-1} \\ &= n^{-1} W(n) \{I(n) - X(n) [X'(n) X(n)]^{-1} X'(n)\} W(n) + n^{-1} \end{aligned}$$

where

$$W(n) = Y(n) - EY(n)$$

$$\hat{\sigma}^2(n) = n^{-1} W'(n) [I(n) - U'(n)U(n)] W(n) + n^{-1}$$

where

$$U(n) = [X'(n)X(n)]^{-1} \frac{1}{2} X'(n)$$

$$\hat{\sigma}^2(n) = n^{-1} W'(n) W(n) - n^{-1} [U(n)W(n)]' [U(n)W(n)] + n^{-1}$$

$$n^{-1} W'(n) W(n) = n^{-1} [Y(n) - EY(n)]' [Y(n) - EY(n)]$$

$$\begin{aligned} & \sum_{i=1}^n e_i^2 \\ &= \frac{\sum_{i=1}^n e_i^2}{n} \rightarrow \sigma^2 \text{ a.s. as } n \rightarrow \infty \end{aligned}$$

by strong law of large numbers.

$$n^{-1} \rightarrow 0 \text{ as } n \rightarrow \infty.$$

So in order to show

$$\hat{\sigma}^2(n) \rightarrow \sigma^2 \text{ a.s. as } n \rightarrow \infty$$

it is necessary to show

$$n^{-1} [U(n)W(n)]' [U(n)W(n)] \rightarrow 0 \text{ a.s. as } n \rightarrow \infty.$$

This can be done, by using Lemma 1, as follows:

$$U(n)W(n) = \begin{bmatrix} U_{11} & U_{12} & \dots & U_{1n} \\ U_{21} & U_{22} & \dots & \end{bmatrix} \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_n \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^n U_{1j} W_j \\ \sum_{j=1}^n U_{2j} W_j \\ \vdots \end{bmatrix}$$

$$n^{-1} [U(n)W(n)]' [U(n)W(n)] = \sum_{i=1}^2 \left[ \frac{1}{n} \sum_{j=1}^n U_{ij} W_j \right]^2$$

$$= \sigma^2 \sum_{i=1}^2 \left[ \frac{1}{n} \sum_{j=1}^n U_{ij} Z_j \right]^2$$

where

$$Z_j = \frac{W_j}{\sigma} = \frac{Y(n) - EY(n)}{\sigma}$$

so

$$EZ_i = 0, EZ_i^2 = 1, EZ_j Z_k = 0, i \neq k.$$

$$j, k = 1, 2, \dots, n.$$

In comparing with the conditions of Lemma 2, identify  $U_{ij}$  with  $b_{mn}$ .

So, if it can be shown that

$$\lim_{n \rightarrow \infty} \sum_{j=1}^n U_{ij}^2 = 1, i = 1, 2$$

then by Lemma 1,

$$n^{-1} \sum_{j=1}^n U_{ij} Z_j \rightarrow 0 \text{ a.s. as } n \rightarrow \infty \quad i = 1, 2.$$

This is so since

$$\begin{aligned} U(n)U'(n) &= \begin{bmatrix} \sum_{j=1}^n U_{1j}^2 & \sum_{j=1}^n U_{1j}U_{2j} \\ \sum_{j=1}^n U_{1j}U_{2j} & \sum_{j=1}^n U_{2j}^2 \end{bmatrix} \\ &= [X'(n)X(n)]^{-\frac{1}{2}} X'(n)X(n) [X'(n)X(n)]^{-\frac{1}{2}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{aligned}$$

hence

$$\sum_{j=1}^n U_{ij}^2 = 1, i = 1, 2.$$

So

$$n^{-1} [U(n)W(n)]' [U(n)W(n)] = \sigma^2 \sum_{i=1}^n \left[ \frac{1}{n} \sum_{j=1}^n U_{ij} Z_j \right]^2 \rightarrow 0 \text{ a.s. as } n \rightarrow \infty.$$

Finally it can be concluded

$$\hat{\sigma}^2(n) \rightarrow \sigma^2 \quad \text{a.s. as } n \rightarrow \infty.$$

With the provision of Lemma 2, the following theorem will be shown.

Theorem 1.

$$P(N < \infty) = 1$$

where  $N$  is a positive integer valued random variable so that  $N = n$  the smallest positive integer so that

$$\hat{\sigma}^2(n) \leq \frac{nd^2}{2a_n}.$$

This means that the stopping rule defined as (2) does lead the procedure to stop.

Proof:

Instead of proving  $P[N < \infty] = 1$  it will be shown that  $P[N = \infty] = 0$ .

$$\begin{aligned} P[N = \infty] &= P\left[\hat{\sigma}^2(n) > \frac{nd^2}{2a_n} \text{ for all } n\right] \\ &= P\left[\hat{\sigma}^2(n) - \frac{nd^2}{2a_n} > 0 \text{ for all } n\right] \\ &= P\left[\frac{\hat{\sigma}^2(n)}{\hat{\sigma}^2} - \frac{nd^2}{2a_n \hat{\sigma}^2} > 0 \text{ for all } n\right]. \end{aligned}$$

There certainly exists a positive integer  $n$  so that

$$n \geq \frac{2\sigma^2}{d^2}.$$

Let  $n$  be the smallest positive integer so that

$$n \geq \frac{2\sigma^2}{d^2}$$

then

$$n = \frac{a^2 \sigma^2}{d^2} + r, \quad 0 \leq r < 1$$

$$nd^2 = a^2 \sigma^2 + r, \quad \frac{nd^2}{a^2 \sigma^2} = 1 + \frac{rd^2}{a^2 \sigma^2}$$

hence

$$\lim_{n \rightarrow \infty} \frac{nd^2}{a_n^2 \sigma^2} = \lim_{d \rightarrow 0} \left[ 1 + \frac{rd^2}{a^2 \sigma^2} \right] = 1$$

since

$$\lim_{n \rightarrow \infty} \frac{nd^2}{a_n^2 \sigma^2} = \lim_{n \rightarrow \infty} \left[ \frac{nd^2}{a^2 \sigma^2} \cdot \frac{a^2}{a_n^2} \right] = 1.$$

Since  $\hat{\sigma}^2(n) \rightarrow \sigma^2$  a.s. as  $n \rightarrow \infty$

$$(6) \quad \frac{\hat{\sigma}^2(n)}{\hat{\sigma}^2} \rightarrow 1 \text{ a.s. as } n \rightarrow \infty.$$

Assume

$$\frac{\hat{\sigma}^2(n)}{\sigma^2} - \frac{nd^2}{a_n^2 \sigma^2} = \varepsilon > 0,$$

for all  $n$ . Then

$$P\{N = \infty\} = P\left[\frac{\hat{\sigma}^2(n)}{\sigma^2} - \frac{nd^2}{a_n^2 \sigma^2} = \varepsilon \text{ for all } n\right].$$

From (5), there exists a positive integer  $M_1(\varepsilon)$ , so that

$$\left| \frac{nd^2}{a_n^2 \sigma^2} - 1 \right| < \frac{\varepsilon}{3} \text{ when } n > M_1(\varepsilon).$$

From (6), there exists a positive integer  $M_2(\varepsilon)$ , so that

$$\left| \frac{\hat{\sigma}^2(n)}{\sigma^2} - 1 \right| < \frac{\varepsilon}{3} \text{ when } n > M_2(\varepsilon).$$

Choose  $M = \max (M_1, M_2)$ , thus

$$\left| \frac{nd^2}{a_n \sigma^2} - 1 \right| + \left| \frac{\hat{\sigma}^2(n)}{\sigma^2} - 1 \right| < \epsilon \text{ when } n > M$$

or

$$\left| \frac{nd^2}{a_n \sigma^2} - \frac{\hat{\sigma}^2(n)}{\sigma^2} \right| < \epsilon \text{ when } n > M.$$

That is, there exists a number  $M$  such that when  $n > M$

$$(8) \quad \left| \frac{nd^2}{a_n \sigma^2} - \frac{\hat{\sigma}^2(n)}{\sigma^2} \right| < \epsilon$$

Compare (8) with (7). The conclusion is that

$$P[N = \infty] = 0 \text{ or } P[N < \infty] = 1.$$

This means that the sequential procedure will terminate subject to the stopping rule.

In order to demonstrate (4), the following lemmas are needed.

Lemma 3 [Chow and Robbins (9)].

Let  $y_n$ ,  $n = 1, 2, 3, \dots$  be any sequence of random variables so that

$$y_n > 0 \text{ a.s. and } \lim_{n \rightarrow \infty} y_n = 1 \text{ a.s.}$$

Let  $f(n)$  be any sequence of constants so that

$$f(n) > 0$$

$$\lim_{n \rightarrow \infty} f(n) = \infty \text{ and}$$

$$\lim_{n \rightarrow \infty} \frac{f(n)}{f(n-1)} = 1.$$

For each  $t > 0$ , define  $N = N(t) = k$ , the smallest positive integer, so that

$$y_k \leq \frac{f(k)}{t}.$$

Then (a)  $N = N(t)$  is a well-defined function of  $t$ .

(b)  $N(t)$  is a non-decreasing function of  $t$ .

$$(c) \lim_{t \rightarrow \infty} \frac{f(n)}{t} = 1 \text{ a.s.}$$

$$(d) \lim_{t \rightarrow \infty} N = \infty \text{ a.s.}$$

Lemma 4.

$$\lim_{t \rightarrow \infty} \frac{Nd^2}{a^2 \sigma^2} = 1 \text{ a.s.}$$

Proof:

Define

$$y_n = \frac{\hat{\sigma}^2(n)}{\sigma^2}, \quad f(n) = \frac{na^2}{a_n^2} \text{ and } t = \frac{a^2 \sigma^2}{d^2},$$

then

$$\lim_{n \rightarrow \infty} y_n = 1 \text{ a.s. since } \hat{\sigma}^2(n) \rightarrow \sigma^2 \text{ a.s.}$$

$$\lim_{n \rightarrow \infty} \frac{f(n)}{f(n-1)} = \lim_{n \rightarrow \infty} \left[ \frac{na^2}{a_n^2} \right] \left[ \frac{a_{n-1}^2}{(n-1)a^2} \right] = 1, \text{ since } a_n^2 \rightarrow a^2 \text{ as } n \rightarrow \infty.$$

Also, by the defined stopping rule (2)

$$k \geq \frac{a_k^2 \hat{\sigma}^2(k)}{d^2} \text{ or } \frac{\hat{\sigma}^2(k)}{\sigma^2} \leq \frac{ka^2}{a_k^2 \left[ \frac{a^2 \sigma^2}{d^2} \right]}$$

i.e.,

$$y_k \leq \frac{f(k)}{t}.$$

Thus, the conditions for Lemma 3 are all satisfied. So,

$$\lim_{n \rightarrow \infty} \frac{f(N)}{t} = 1 \text{ a.s.}$$

This implies that

$$\lim_{t \rightarrow \infty} \frac{f(N)}{t} = \lim_{t \rightarrow \infty} \left[ \frac{Na^2}{2} \cdot \frac{d^2}{a^2 \sigma^2} \right] = \lim_{t \rightarrow \infty} \left[ \frac{Nd^2}{2a^2 \sigma^2} \right] = \lim_{t \rightarrow \infty} \left[ \frac{Nd^2}{a^2 \sigma^2} \right] = 1 \text{ a.s.}$$

$$(7) \quad \lim_{t \rightarrow \infty} \frac{Nd^2}{2a^2 \sigma^2} = 1 \text{ a.s.}$$

Lemma 5.

$$\lim_{t \rightarrow \infty} \frac{N}{t} = 1 \text{ a.s.}$$

Proof:

Recalling that

$$t = \frac{a^2 \sigma^2}{d^2},$$

substitute  $t$  in (7) then

$$\lim_{t \rightarrow \infty} \frac{N}{t} = 1, \text{ a.s.}$$

Lemma 6 [Wijsman, see Srivastava (11)].

Let  $Z_1, Z_2, Z_3, \dots$  be iid random variables with

$$EZ_i = 0, EZ_i^2 = 1, i = 1, 2, 3, \dots$$

Let  $b_1, b_1, \dots, b_n$  be a sequence of constants so that

$$\lim_{n \rightarrow \infty} \left[ n^{-1} \sum_{i=1}^n b_i^2 \right] = 1.$$

Let  $N$  be a positive integer valued random variable so that  $\frac{N}{t} \rightarrow 1$  in probability as  $t \rightarrow \infty$ . Then

$$\lim_{t \rightarrow \infty} N^{-\frac{1}{2}} \sum_{i=1}^N b_i Z_i = N(0,1) \text{ in distribution.}$$

Lemma 7.

Let  $N$  be a positive valued random variable so that  $\frac{N}{t} \rightarrow 1$  in probability as  $t \rightarrow \infty$ . Then

$$N^{-\frac{1}{2}} \sum_{i=1}^N \left[ \lambda_0 + \lambda_1 \left( \frac{X_i - \bar{X}}{S} \right) \right] Z_i \rightarrow N(0,1) \text{ in distribution as } t \rightarrow \infty$$

where

$$\lambda_0^2 + \lambda_1^2 = 1, Z_i \text{'s are the same as in Lemma 6.}$$

Proof:

From Lemma 5,

$$\lim_{t \rightarrow \infty} \frac{N}{t} = 1$$

in probability since convergence a.s. implies convergence in probability [see Roussas (14)]. Define

$$b_i = \left[ \lambda_0 + \lambda_1 \left( \frac{X_i - \bar{X}}{S} \right) \right]^2,$$

then

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n b_i^2 &= \frac{1}{n} \sum_{i=1}^n \left[ \lambda_0 + \lambda_1 \left( \frac{X_i - \bar{X}}{S} \right) \right]^2 \\ &= \frac{1}{n} \left[ n \lambda_0^2 + \frac{\lambda_1^2}{S^2} \sum_{i=1}^n (X_i - \bar{X})^2 + 2 \lambda_0 \lambda_1 \sum_{i=1}^n (X_i - \bar{X}) \right] \\ &= \frac{1}{n} \left[ n \lambda_0^2 + \frac{\lambda_1^2}{S^2} \sum_{i=1}^n (X_i - \bar{X}_n)^2 + \frac{\lambda_1^2}{S^2} (\bar{X}_n - \bar{X})^2 + 2 \lambda_0 \lambda_1 \sum_{i=1}^n (X_i - \bar{X}) \right] \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{n} \left[ n\lambda_0^2 + \frac{\lambda_1^2}{s^2} \sum_1^n (x_1 - \bar{x}_n)^2 + \frac{n\lambda_1^2}{s^2} (x_n - \bar{x})^2 + 2\lambda_0\lambda_1 \sum_1^n (x_1 - \bar{x}) \right] \\
\frac{1}{n} &= \left[ n\lambda_0^2 + \frac{n\lambda_1^2}{s^2} s_n^2 + \frac{n\lambda_1^2}{s^2} (\bar{x}_n - \bar{x})^2 + 2\lambda_0\lambda_1 \sum_1^n (x_1 - \bar{x}_n) + 2\lambda_0\lambda_1 \sum_1^n (\bar{x}_n - \bar{x}) \right] \\
&= \lambda_0^2 + \lambda_1^2 \frac{s_n^2}{s^2} + \frac{\lambda_1^2}{s^2} (\bar{x}_n - \bar{x})^2 + 2\lambda_0\lambda_1 (\bar{x}_n - \bar{x})
\end{aligned}$$

i.e.,

$$\frac{1}{n} \sum_1^n b_1^2 = \lambda_0^2 + \lambda_1^2 \left[ \frac{s_n^2}{s^2} \right] + \frac{\lambda_1^2}{s^2} (\bar{x}_n - \bar{x}) + 2\lambda_0\lambda_1 (\bar{x}_n - \bar{x}).$$

Since

$$s_n^2 \rightarrow s^2, \bar{x}_n \rightarrow \bar{x} \text{ and } \lambda_0^2 + \lambda_1^2 = 1,$$

therefore,

$$\lim_{n \rightarrow \infty} \left[ \frac{1}{n} \sum_1^n b_1^2 \right] = \lambda_0^2 + \lambda_1^2 = 1.$$

Thus, the conditions of Lemma 6 are all satisfied and it can be concluded that

$$\frac{1}{\sqrt{N}} \sum_{i=1}^N \left[ \lambda_0 + \lambda_1 \left( \frac{x_i - \bar{x}}{s} \right) \right] Z_i \rightarrow N(0,1) \text{ in distribution as } t \rightarrow \infty.$$

Lemma 8.

Let  $N(t)$ ,  $Z_i$ 's and  $\lambda_0, \lambda_1$  are defined as Lemma 7, then

$$\frac{1}{\sqrt{N}} \sum_{i=1}^N \left[ \lambda_0 + \lambda_1 \left( \frac{x_i - \bar{x}_n}{s_n} \right) \right] Z_i \rightarrow N(0,1) \text{ in distribution as } t \rightarrow \infty.$$

Proof:

From Lemma 7,

$$\frac{1}{\sqrt{N}} \sum_{i=1}^N \left( \frac{x_i - \bar{x}}{s} \right) Z_i \rightarrow N(0,1) \text{ in distribution as } t \rightarrow \infty,$$

by setting

$$\lambda_0 = 0, \lambda_1 = 1, \text{ so } \lambda_0^2 + \lambda_1^2 = 1.$$

Also

$$\frac{1}{\sqrt{N}} \sum_{i=1}^N Z_i \rightarrow N(0,1)$$

in distribution by Anscombe, [see Chow and Robbins (9)]. Thus,

$$\begin{aligned} & \frac{1}{\sqrt{N}} \sum_{i=1}^N \left[ \lambda_0 + \lambda_1 \left( \frac{X_i - \bar{X}_n}{S_n} \right) \right] Z_i \\ &= \frac{1}{\sqrt{N}} \sum_{i=1}^N \left\{ \lambda_0 + \lambda_1 \left( \frac{S}{S_N} \right) \left( \frac{X_i - \bar{X}}{S} + \frac{\bar{X} - \bar{X}_n}{S} \right) \right\} Z_i \\ &= \lambda_0 \left[ \frac{1}{\sqrt{N}} \sum_{i=1}^N Z_i \right] \\ &+ \lambda_1 \left( \frac{S}{S_N} \right) \left[ \frac{1}{N} \sum_{i=1}^N \left( \frac{X_i - \bar{X}}{S} \right) Z_i \right] \\ &+ \lambda_1 \left( \frac{S}{S_N} \right) \left( \frac{\bar{X} - \bar{X}_n}{S} \right) \frac{1}{\sqrt{N}} \sum_{i=1}^N Z_i \\ &\lim_{t \rightarrow \infty} \left\{ \frac{1}{\sqrt{N}} \sum_{i=1}^N \left[ \lambda_0 + \lambda_1 \left( \frac{X_i - \bar{X}_N}{S_N} \right) \right] Z_i \right\} = \lambda_0 N(0,1) + \lambda_1 N(0,1) \\ &= N(0, \lambda_0^2) + N(0, \lambda_1^2) = N(0, \lambda_0^2 + \lambda_1^2) = N(0,1) \text{ in distribution} \end{aligned}$$

i.e.,

$$\frac{1}{\sqrt{N}} \sum_{i=1}^N \left[ \lambda_0 + \lambda_1 \left( \frac{X_i - \bar{X}_N}{S_N} \right) \right] Z_i \rightarrow N(0,1) \text{ in distribution as } t \rightarrow \infty.$$

Lemma 9.

$$\lambda_0 \left[ \frac{\hat{\beta}_0 - \beta_0}{\frac{\sigma}{\sqrt{N}}} \right] + \lambda_1 \left[ \frac{\hat{\beta}_1 - \beta_1}{\frac{\sigma}{\sqrt{NS_N^2}}} \right] \rightarrow N(0,1) \text{ in distribution as } t \rightarrow \infty.$$

Proof:

$$\begin{aligned}
& \lambda_0 \left[ \frac{\hat{\beta}_0 - \beta_0}{\frac{\sigma}{\sqrt{N}}} \right] + \lambda_1 \left[ \frac{\hat{\beta}_1 - \beta_1}{\frac{\sigma}{\sqrt{NS_N^2}}} \right] = (\lambda_0, \lambda_1) \begin{bmatrix} \frac{\hat{\beta}_0 - \beta_0}{\frac{\sigma}{\sqrt{N}}} \\ \frac{\hat{\beta}_1 - \beta_1}{\frac{\sigma}{\sqrt{NS_N^2}}} \end{bmatrix} \\
& = \sigma^{-1} (\lambda_0, \lambda_1) \begin{bmatrix} N & 0 \\ 0 & NS_N^2 \end{bmatrix} \begin{bmatrix} \hat{\beta}_0 - \beta_0 \\ \hat{\beta}_1 - \beta_1 \end{bmatrix} \\
& = \sigma^{-1} \begin{bmatrix} N & 0 \\ 0 & NS_N^2 \end{bmatrix} [\hat{\beta} - \beta], \quad \lambda = \begin{bmatrix} \lambda_0 \\ \lambda_1 \end{bmatrix}, \quad [\hat{\beta} - \beta] = \begin{bmatrix} \hat{\beta}_0 - \beta_0 \\ \hat{\beta}_1 - \beta_1 \end{bmatrix} \\
& = \sigma^{-1} \lambda' \left[ X'(N)X(N) \right]^{\frac{1}{2}} (\hat{\beta} - \beta) \\
& = \sigma^{-1} \lambda' \left[ X'(N)X(N) \right]^{\frac{1}{2}} \left[ X'(N)X(N) \right]^{-1} X'(N) [Y(N) - EY(N)] \\
& = \sigma^{-1} \lambda' \left[ X'(N)X(N) \right]^{-\frac{1}{2}} X'(N) [Y(N) - EY(N)] \\
& = \sigma^{-1} \lambda' N^{-1} \left[ X'(N)X(N) \right]^{-\frac{1}{2}} X'(N) [Y(N) - EY(N)] \\
& = \frac{\sigma^{-1}}{\sqrt{N}} \lambda' \begin{bmatrix} 1 & 0 \\ 0 & S_N^2 \end{bmatrix}^{-\frac{1}{2}} X'(N) [Y(N) - EY(N)] \\
& = \frac{1}{\sqrt{N}} \lambda' \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{S_N^2} \end{bmatrix} X'(N) \left[ \frac{Y(N) - EY(N)}{\sigma} \right]
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{\sqrt{N}} \left[ \lambda_0 + \lambda_1 \left( \frac{X_1 - \bar{X}_N}{S_N} \right), \dots, \lambda_0 + \lambda_1 \left( \frac{X_N - \bar{X}_N}{S_N} \right) \right] \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_N \end{bmatrix} \\
&= \frac{1}{\sqrt{N}} \sum_{i=1}^N \left[ \lambda_0 + \lambda_1 \left( \frac{X_i - \bar{X}_N}{S_N} \right) \right] Z_i
\end{aligned}$$

i.e.,

$$\lambda_0 \left[ \frac{\hat{\beta}_0 - \beta_0}{\frac{\sigma}{\sqrt{N}}} \right] + \lambda_1 \left[ \frac{\hat{\beta}_1 - \beta_1}{\frac{\sigma}{\sqrt{NS_N^2}}} \right] = \frac{1}{\sqrt{N}} \sum_{i=1}^N \left[ \lambda_0 + \lambda_1 \left( \frac{X_i - \bar{X}_N}{S_N} \right) Z_i \right]$$

From Lemma 8,

$$\lambda_0 \left[ \frac{\hat{\beta}_0 - \beta_0}{\frac{\sigma}{\sqrt{N}}} \right] + \lambda_1 \left[ \frac{\hat{\beta}_1 - \beta_1}{\frac{\sigma}{\sqrt{NS_N^2}}} \right] \rightarrow N(0,1) \text{ in distribution as } t \rightarrow \infty.$$

Lemma 10.

$$\frac{\hat{\beta}_0 - \beta_0}{\frac{\sigma}{\sqrt{N}}} + \frac{\hat{\beta}_1 - \beta_1}{\frac{\sigma}{\sqrt{NS_N^2}}} \left( \frac{X - \bar{X}}{S} \right) \rightarrow N \left[ 0, 1 + \left( \frac{X - \bar{X}}{S} \right)^2 \right]$$

in distribution as  $t \rightarrow \infty$ .

Proof:

From Lemma 9, assign

$$\begin{aligned}
\lambda_0 &= \left[ 1 + \left( \frac{X - \bar{X}}{S} \right)^2 \right]^{-\frac{1}{2}} \\
\lambda_1 &= \left( \frac{X - \bar{X}}{S} \right) \left[ 1 + \left( \frac{X - \bar{X}}{S} \right)^2 \right]^{-\frac{1}{2}}
\end{aligned}$$

then

$$\lambda' \lambda = \lambda_0^2 + \lambda_1^2 = 1$$

$$\lambda_0 \left[ \frac{\hat{\beta}_0 - \beta_0}{\frac{\sigma}{\sqrt{N}}} \right] + \lambda_1 \left[ \frac{\hat{\beta}_1 - \beta_1}{\frac{\sigma}{\sqrt{NS_N^2}}} \right] = \left[ 1 + \left( \frac{X - \bar{X}}{S} \right)^2 \right]^{-\frac{1}{2}}$$

$$\left| \frac{\hat{\beta}_0 - \beta_0}{\frac{\sigma}{\sqrt{N}}} + \frac{\hat{\beta}_1 - \beta_1}{\frac{\sigma}{\sqrt{NS_N^2}}} \frac{X - \bar{X}}{S} \right| \rightarrow N(0,1) \text{ in distribution as } t \rightarrow \infty.$$

$$\frac{\hat{\beta}_0 - \beta_0}{\frac{\sigma}{\sqrt{N}}} + \frac{\hat{\beta}_1 - \beta_1}{\frac{\sigma}{\sqrt{NS_N^2}}} \frac{X - \bar{X}}{S} \rightarrow \left[ 1 + \left( \frac{X - \bar{X}}{S} \right)^2 \right]^{-\frac{1}{2}} N(0,1)$$

$$= N \left[ 0, 1 + \left( \frac{X - \bar{X}}{S} \right)^2 \right] \text{ in distribution as } t \rightarrow \infty.$$

Lemma 11.

$$\frac{\hat{\beta}_0 - \beta_0}{\frac{\sigma}{\sqrt{N}}} + \frac{\hat{\beta}_1 - \beta_1}{\frac{\sigma}{\sqrt{NS_N^2}}} \frac{X - \bar{X}_N}{S_N} = N \left[ 0, 1 + \left( \frac{X - \bar{X}}{S} \right)^2 \right]$$

in distribution as  $t \rightarrow \infty$ .

Proof:

$$\begin{aligned} & \frac{\hat{\beta}_0 - \beta_0}{\frac{\sigma}{\sqrt{N}}} + \frac{\hat{\beta}_1 - \beta_1}{\frac{\sigma}{\sqrt{NS_N^2}}} \frac{X - \bar{X}_N}{S_N} \\ &= \frac{\hat{\beta}_0 - \beta_0}{\frac{\sigma}{\sqrt{N}}} + \frac{\hat{\beta}_1 - \beta_1}{\frac{\sigma}{\sqrt{NS_N^2}}} \left( \frac{S}{S_N} \right) \left[ \frac{X - \bar{X}}{S} + \frac{X - \bar{X}_N}{S} \right] \\ &= \frac{\hat{\beta}_0 - \beta_0}{\frac{\sigma}{\sqrt{N}}} + \left( \frac{S}{S_N} \right) \frac{\hat{\beta}_1 - \beta_1}{\frac{\sigma}{\sqrt{NS_N^2}}} \left( \frac{X - \bar{X}}{S} \right) + \left( \frac{S}{S_N} \right) \left[ \frac{\hat{\beta}_1 - \beta_1}{\frac{\sigma}{\sqrt{NS_N^2}}} \right] \left( \frac{X - \bar{X}_N}{S} \right) \end{aligned}$$

From Lemma 9, assign

$$\lambda_0 = 1, \lambda_1 = 0$$

then

$$\frac{\hat{\beta}_0 - \beta_0}{\frac{\sigma}{\sqrt{N}}} \rightarrow N(0,1) \text{ in distribution as } t \rightarrow \infty.$$

If assign

$$\lambda_0 = 0, \lambda_1 = 1,$$

then

$$\frac{\hat{\beta}_1 - \beta_1}{\frac{\sigma}{\sqrt{NS_N^2}}} \rightarrow N(0,1) \text{ in distribution as } t \rightarrow \infty.$$

Also,

$$\frac{S}{S_N} \rightarrow 1, \bar{X}_N \rightarrow \bar{X} \text{ as } t \rightarrow \infty \text{ since } \frac{N}{t} \rightarrow 1 \text{ a.s.}$$

therefore

$$\begin{aligned} & \frac{\hat{\beta}_0 - \beta_0}{\frac{\sigma}{\sqrt{N}}} + \frac{\hat{\beta}_1 - \beta_1}{\frac{\sigma}{\sqrt{NS_N^2}}} \frac{X - X_N}{S_N} \rightarrow N(0,1) + \left( \frac{X - \bar{X}}{S} \right) N(0,1) \\ & = N \left[ 0, 1 + \left( \frac{X - \bar{X}}{S} \right)^2 \right] \text{ in distribution as } t \rightarrow \infty. \end{aligned}$$

Now it is ready to prove

Theorem 2.

$$\lim_{d \rightarrow 0} P\{\mu_X \in I_N\} \geq \alpha$$

Proof:

Since, by the stopping rule that  $N = N(t) = n$ , the smallest positive integer so that

$$n \geq \frac{a_n^2 \hat{\sigma}^2(n)}{d^2} \quad \text{or} \quad \frac{a_n \hat{\sigma}(n)}{\sqrt{n}} \leq d$$

then

$$(\hat{\mu}_X - d, \hat{\mu}_X + d) \supset \left[ \hat{\mu}_X - \frac{a_n \hat{\sigma}(n)}{\sqrt{n}}, \hat{\mu}_X + \frac{a_n \hat{\sigma}(n)}{\sqrt{n}} \right].$$

Define

$$J_n = \left[ \hat{\mu}_X - \frac{a_n \hat{\sigma}(n)}{\sqrt{n}}, \hat{\mu}_X + \frac{a_n \hat{\sigma}(n)}{\sqrt{n}} \right]$$

and

$$I_n = (\mu_X - d, \mu_X + d)$$

as defined before, then

$$P\{\mu_X \in I_n\} \geq P\{\mu_X \in J_n\}$$

So, if

$$\lim_{d \rightarrow 0} P\{\mu_X \in J_n\} \geq \alpha$$

can be shown then Theorem 2 is proved.

$$\begin{aligned} P\{\mu_X \in J_n\} &= P\left\{[(\hat{\beta}_0 - \beta_0) + (\hat{\beta}_1 - \beta_1)(X - \bar{X}_N)] < \frac{\hat{\sigma}_N a_N}{\sqrt{N}}\right\} \\ &= P\left\{\left[\frac{\hat{\beta}_0 - \beta_0}{\frac{\sigma}{\sqrt{N}}} + \frac{\hat{\beta}_1 - \beta_1}{\frac{\sigma}{\sqrt{NS_N^2}}} \left(\frac{X - \bar{X}_N}{S_N}\right)\right] \leq \frac{a_N}{2}\right\}. \end{aligned}$$

Let

$$L_N = \frac{\hat{\beta}_0 - \beta_0}{\frac{\sigma}{\sqrt{N}}} + \frac{\hat{\beta}_1 - \beta_1}{\frac{\sigma}{\sqrt{NS_N^2}}} \left(\frac{X - \bar{X}_N}{S_N}\right),$$

$F_N$  be the distribution function of  $L_N$ , and  $F$  be the limiting distribution of  $F$ . From Lemma 11,

$$L_N \rightarrow N \left[ 0, 1 + \left( \frac{X - \bar{X}}{S} \right)^2 \right] \text{ in distribution as } t \rightarrow \infty.$$

So

$$F = N \left[ 0, 1 + \left( \frac{X - \bar{X}}{S} \right)^2 \right]$$

$$P \left\{ |L_N| < \frac{a_N \hat{\sigma}(N)}{\sigma} \right\} = F_N \left[ \frac{a_N \hat{\sigma}(N)}{\sigma} \right] - F_N \left[ - \frac{a_N \hat{\sigma}(N)}{\sigma} \right]$$

$$\lim_{t \rightarrow \infty} P \left\{ |L_N| < \frac{a_N \hat{\sigma}(N)}{\sigma} \right\} = \lim_{t \rightarrow \infty} \left[ F_N \left[ \frac{a_N \hat{\sigma}(N)}{\sigma} \right] - F_N \left[ - \frac{a_N \hat{\sigma}(N)}{\sigma} \right] \right]$$

$$= [F(a) - F(-a)] \text{ since } a_N \rightarrow a, \hat{\sigma}(N) \rightarrow \hat{\sigma}^2 \text{ a.s. as } t \rightarrow \infty.$$

$$= (2\pi)^{-\frac{1}{2}} \int \frac{\frac{a}{\sqrt{1 + \left( \frac{X - \bar{X}}{S} \right)^2}}}{\frac{-a}{1 + \left( \frac{X - \bar{X}}{S} \right)^2}} e^{-\frac{u^2}{2}} \geq du$$

$$(2\pi)^{-\frac{1}{2}} \int \frac{\frac{a}{\sqrt{1 + \left( \frac{h}{S} \right)^2}}}{\frac{-a}{\sqrt{1 + \left( \frac{h}{S} \right)^2}}} e^{-\frac{u^2}{2}} du = \alpha$$

by the definition of  $a$  and since

$$h^2 > (X - \bar{X})^2$$

That is

$$\lim_{t \rightarrow \infty} P\{\mu_X \in I_N\} \geq \lim_{t \rightarrow \infty} P\{\mu_X \in J_N\} \geq \alpha.$$

Since  $t$  and  $d$  are related by

$$t = \frac{a^2 \sigma^2}{d^2}$$

the last result can be written as

$$\lim_{d \rightarrow 0} P\{\mu_X \in I_N\} \geq \lim_{d \rightarrow 0} P\{\mu_X \in J_N\} \geq \alpha.$$

#### Comment

The quantities of  $h$  and  $S^2$  used in this chapter have the same significance as discussed in Chapter II. Namely, if it is possible, the shorter  $h$  and larger  $S^2$  are recommended. The confidence interval

$$I_n = (\hat{\mu}_X - d, \hat{\mu}_X + d)$$

is used and

$$J_n = \hat{\mu}_X - \frac{a_n \hat{\sigma}(n)}{\sqrt{n}}, \hat{\mu}_X + \frac{a_n \hat{\sigma}(n)}{\sqrt{n}}$$

defined where

$$a_n = a \left( 1 + \frac{|X - X_n|}{\sqrt{nS_n^2}} \right)$$

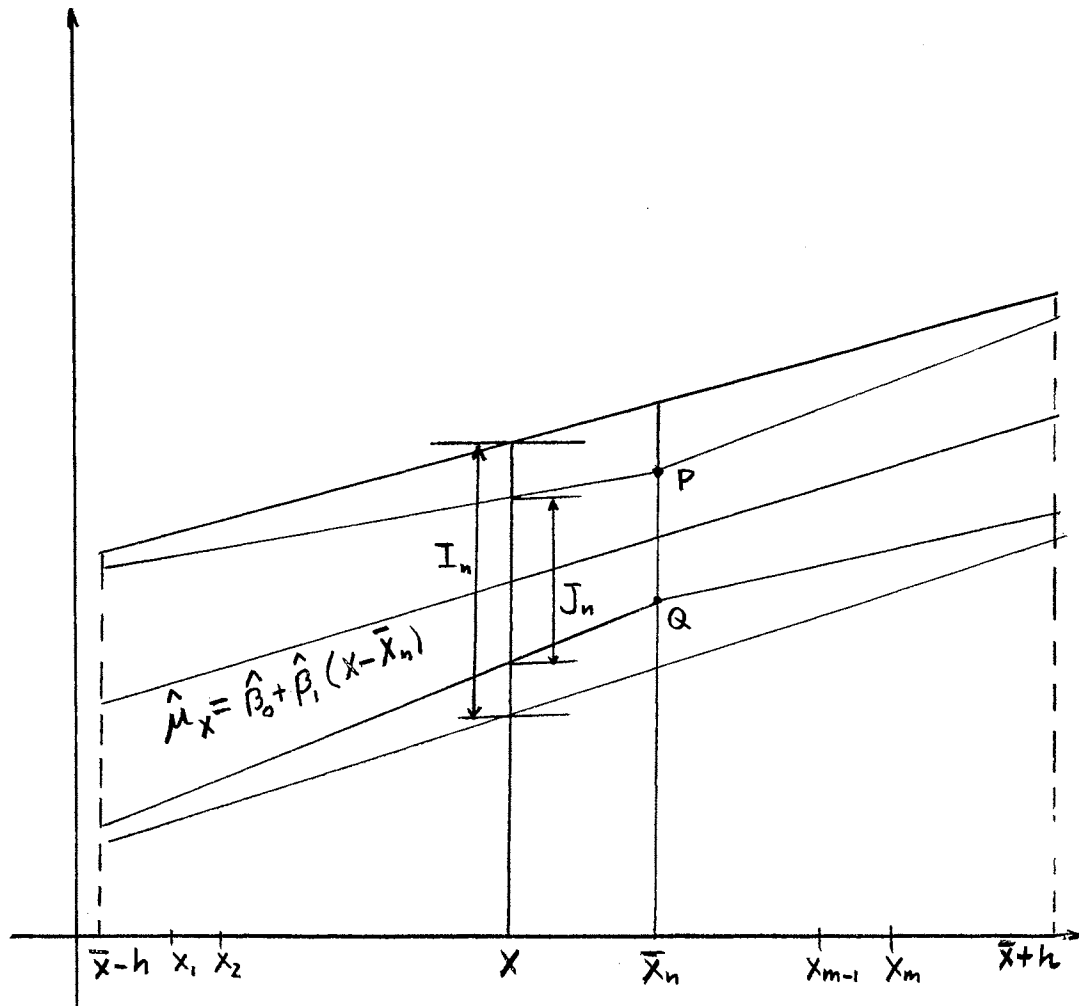
are related as  $I_n \supset J_N$  since  $n$  is defined as the smallest positive integer so that

$$\frac{\hat{\sigma}^2(n)}{a_n^2} \geq \frac{nd^2}{2} \quad \text{or} \quad \frac{a_n \hat{\sigma}(n)}{\sqrt{n}} \leq d.$$

The picture of this relationship is shown on page 47.

#### Asymptotic Efficiency

The sequential procedure used has another property, Chow and Robbins (9) called it the asymptotic efficiency property, namely



$$I_n = \left\{ \hat{\mu}_X - d, \hat{\mu}_X + d \right\}$$

$$J_n = \left\{ \hat{\mu}_X - \frac{a_n \hat{\sigma}(n)}{\sqrt{n}}, \hat{\mu}_X + \frac{a_n \hat{\sigma}(n)}{\sqrt{n}} \right\}$$

$$a_n = a \left( 1 + \frac{|X - \bar{x}_n|}{\sqrt{n} S_n^2} \right)$$

$$P = P \left( \bar{x}_n, \hat{\mu}_X + \frac{a \hat{\sigma}(n)}{\sqrt{n}} \right)$$

$$Q = Q \left( \bar{x}_n, \hat{\mu}_X - \frac{a \hat{\sigma}(n)}{\sqrt{n}} \right)$$

Figure 1. Sequential Confidence Intervals of Constant Width for a Simple Linear Regression Mean

$$\lim_{d \rightarrow 0} \frac{d^2 EN}{a^2 \sigma^2} = 1.$$

Recall that when  $\sigma^2$  is known, the sample size  $n$  is determined as the smallest positive integer so that

$$(1) \quad n \geq \frac{a^2 \sigma^2}{d^2}$$

when  $\sigma^2$  is unknown; the sample size  $k$  is determined as the smallest positive integer so that

$$(2) \quad N = N(t) = k,$$

so that

$$k \geq \frac{a_k^2 \sigma^2}{d^2}$$

where  $N$  is a positive integer valued random variable. This property means that when  $\sigma^2$  is unknown, the expected value of the sample size  $N$  determined by (2) is the same as the sample size  $n$  determined by (1) when  $\sigma^2$  is known as  $d$  goes to zero.

The following demonstration is essentially based on Chow and Robbins' (9) work.

Lemma 12 [Chow and Robbins (9)].

If (a)  $y_n$  is a sequence of random variables so that  $y_n > 0$  a.s.

(b)  $f(n)$  is a sequence of constants so that

$$f(n) > 0, \lim_{n \rightarrow \infty} f(n) = \infty \text{ and } \lim_{n \rightarrow \infty} \left[ \frac{f(n)}{f(n-1)} \right] = 1.$$

(c) For all  $t > 0$ , define  $N(t) = k$ , the smallest positive integer,

$$y_k \leq \frac{f(k)}{t}$$

$$(d) \quad \lim_{n \rightarrow \infty} \left[ \frac{f(n)}{n} \right] = 1$$

(e)  $E(N) < \infty$  for all  $t > 0$

(f)  $\lim_{t \rightarrow \infty} \sup [E(Ny_N)/E(N)] \leq 1$

(g) there exists a sequence of constants  $g(n)$  so that  $g(n) > 0$ ,

$$\lim_{n \rightarrow \infty} g(n) = 1, g(n) > 0, \lim_{n \rightarrow \infty} g(n) = 1, y_n \geq g(n)G(n-1)$$

when the conditions of this list are all satisfied, then

$$\lim_{t \rightarrow \infty} \frac{E(N)}{t} = 1.$$

Theorem 3:

The sequential procedure is asymptotically efficient, i.e.,

$$\lim_{d \rightarrow 0} \frac{d^2 E(N)}{a^2 \sigma^2} = 1.$$

Proof:

Define

$$y_n = \frac{\hat{\sigma}^2(n)}{\sigma^2}, f(n) = \frac{na^2}{a_n^2} \text{ and } t = \frac{a^2 \sigma^2}{d^2}.$$

All that is needed is to show

$$(3) \lim_{t \rightarrow \infty} \frac{E(N)}{t} = 1.$$

Once (3) is true and since

$$t = \frac{a^2 \sigma^2}{d^2},$$

then

$$\lim_{d \rightarrow 0} \frac{d^2 E(N)}{a^2 \sigma^2} = 1.$$

In the following, (3) will be shown by Lemma 12. The above list of conditions of Lemma 12 will be checked and it will be seen that each of

them is satisfied. As in the proof of Lemma 4, conditions (a), (b), and (c) have been checked. Check on (d), i.e.,

$$\lim_{n \rightarrow \infty} \frac{f(n)}{n} = 1.$$

$$\lim_{n \rightarrow \infty} \left[ \frac{f(n)}{n} \right] = \lim_{n \rightarrow \infty} \left[ \frac{1}{n} \left( \frac{na^2}{a_n^2} \right) \right] = \lim_{n \rightarrow \infty} \left[ \frac{a^2}{a_n^2} \right] = 1$$

since  $a_n^2 \rightarrow a^2$ .

Check on (g), i.e., there exists a sequence of constants  $g(n)$  so that

$$g(n) > 0, \lim_{n \rightarrow \infty} g(n) = 1, y_n \geq g(n)y_{n-1}.$$

Define

$$g(n) = \left[ \frac{n-1}{n} \right], \text{ then } y_n = \frac{\sigma^2(n)}{\sigma^2} = \frac{1}{n\sigma^2} \left[ \sum_{i=1}^n (Y_i - \hat{Y}_n)^2 + 1 \right]$$

where  $\hat{Y}_n = \hat{\mu}_X$

$$\begin{aligned} y_n &\geq \frac{1}{n\sigma^2} \left[ \sum_{i=1}^{n-1} (Y_i - \hat{Y}_{n-1})^2 + 1 \right] = \left[ \frac{n-1}{n} \right] \left[ \frac{1}{(n-1)\sigma^2} \sum_{i=1}^{n-1} (Y_i - \hat{Y}_{n-1})^2 + 1 \right] \\ &= \left[ \frac{n-1}{n} \right] \left[ \frac{\sigma^2(n-1)}{\sigma^2} \right] = \left[ \frac{n-1}{n} \right] y_{n-1} = g(n)y_{n-1}. \end{aligned}$$

Thus,

$$g(n) > 0, \lim_{n \rightarrow \infty} g(n) = 1, \text{ and } y_n \geq g(n)y_{n-1}.$$

Check on (e), i.e.,  $E(N) < \infty$  for all  $t > 0$ . For fixed  $t > 0$ , choose  $m$  so that

$$\frac{f(n)}{t} \geq 1 \text{ when } n > m.$$

This choosing is possible since

$$f(n) = \frac{na^2}{a_n^2} \text{ and } t = \frac{a^2\sigma^2}{d^2}.$$

So

$$\frac{f(n)}{t} = \left[ \frac{na^2}{a_n^2} \right] \left[ \frac{d^2}{a^2 \sigma^2} \right] = \frac{nd^2}{a_n^2 \sigma^2}$$

when  $n \rightarrow \infty$ ,  $a_n \rightarrow a$

$$\lim_{n \rightarrow \infty} \frac{f(n)}{t} = \lim_{n \rightarrow \infty} \left[ \frac{d^2 n}{a^2 \sigma^2} \right] = \infty.$$

Therefore, a number  $m$  can be chosen so that when  $n > m$ ,

$$\frac{f(n)}{t} \geq 1.$$

Choose  $0 < \delta < 1$  such that  $(n-1)f(n-1) \geq \delta n^2$  when  $n \geq 2$ . This choosing is possible since

$$\begin{aligned} \frac{(n-1)f(n-1)}{n^2} &= \left( \frac{n-1}{n^2} \right) \left[ \frac{(n-1)a^2}{a_{n-1}^2} \right] = \left( \frac{n-1}{n} \right)^2 \left[ \frac{a^2}{a_{n-1}^2} \right] \\ &= \left( 1 - \frac{1}{n} \right)^2 \left( \frac{a^2}{a_{n-1}^2} \right) \end{aligned}$$

$$\lim_{n \rightarrow \infty} \left[ \frac{(n-1)f(n-1)}{n^2} \right] = 1.$$

So  $\delta$  can be chosen when  $n \geq 2$ , so that

$$(n-1)f(n-1) \geq \delta n^2.$$

Now define for any  $r \geq m$ ,  $M = \text{Min}(N, r)$ , then by Wald's theorem for cumulative sums [see Ferguson (15) on p. 374],

$$E \left[ \sum_{i=1}^M (Y_i - \mu_X)^2 \right] = E(M) E(Y_1 - \mu_X)^2 = \sigma^2 E(M)$$

$$E(MY_M) = E \left\{ M \frac{1}{M\sigma^2} \left[ \sum_{i=1}^M (Y_i - \hat{Y}_M)^2 + 1 \right] \right\}$$

$$\begin{aligned}
&= \frac{1}{\sigma^2} E \left[ \sum_{i=1}^M (Y_i - \hat{Y}_M)^2 + 1 \right] \leq \frac{1}{\sigma^2} E \left[ \sum_{i=1}^M (Y_i - \mu_X)^2 + 1 \right] \\
&= \frac{1}{\sigma^2} \left[ \sigma^2 E(M) + 1 \right] = EM + \frac{1}{\sigma^2}
\end{aligned}$$

i.e.,

$$(4) \quad E(My_M) \leq EM + \frac{1}{\sigma^2}.$$

Recalling  $M = \min(N, r)$ , so

$$\begin{aligned}
E(My_M) &= \int_{\{r < N\}} ry_r + \int_{\{N \leq r\}} Ny_N \\
E(My_M) &\geq r \frac{f(r)}{t} P\{r < N\} + \int_{\{2 \leq N \leq r\}} Ny_N
\end{aligned}$$

since

$$\frac{f(r)}{t} < y_r \text{ and } \int_{\{2 \leq N \leq r\}} Ny_N < \int_{\{N \leq r\}} Ny_N$$

Also, recalling that  $M$  is a number so that a fixed  $t > 0$

$$\frac{f(n)}{t} \geq 1, \text{ when } n > M.$$

Now since  $r \geq M$ , so

$$\frac{f(r)}{t} \geq 1.$$

Thus,

$$E(My_M) \geq rP\{r < N\} + \int_{\{2 \leq N \leq r\}} Ny_N \geq rP\{r < N\} + \int_{\{2 \leq N \leq r\}} Ng(N) \frac{f(n-1)}{t}$$

since

$$\begin{aligned} y_n &\geq g(n)y_{n-1} \geq g(n) \frac{f(n-1)}{t} = rP\{r < N\} + \frac{1}{t} \int_{\{2 \leq N \leq r\}} \frac{N^2 g(N) f(N-1)}{N} \\ &= rP\{r < N\} + \frac{1}{t} \int_{\{2 \leq N \leq r\}} \frac{N^2 \frac{N-1}{N} f(N-1)}{N} \\ &= rP\{r < N\} + \frac{\delta}{t} \int_{\{2 \leq N \leq r\}} N^2 \end{aligned}$$

since  $\delta$  is chosen so that  $(n-1)f(n-1) \geq \delta n^2$  or

$$\frac{(N-1)f(N-1)}{N^2} \geq \delta, N \geq 2.$$

Therefore,

$$(5) \quad E(My_M) \geq rP\{N > r\} + \frac{\delta}{t} \int_{\{2 \leq N \leq r\}} N^2$$

From (4) and (5)

$$(6) \quad rP\{N > r\} + \frac{\delta}{t} \int_{\{2 \leq N \leq r\}} N^2 \leq E(My_M) \leq EM + \frac{1}{\sigma^2}$$

Recalling that  $M = \min(N, r)$ , thus

$$\begin{aligned}
 EM &= \int_{\{r < N\}} r + \int_{\{N \leq r\}} N < \int_{\{r \leq N\}} r + \int_{\{2 \leq N \leq r\}} N \\
 &= rP\{N > r\} + \int_{\{2 \leq N \leq r\}} N
 \end{aligned}$$

Thus

$$(7) \quad E(M) \leq \int_{\{2 \leq N \leq r\}} N + rP\{N > r\}$$

Substitute (7) into (6)

$$rP\{N > r\} + \frac{\delta}{t} \int_{\{2 \leq N \leq r\}} N^2 \leq \int_{\{2 \leq N \leq r\}} N + rP\{N > r\} + \frac{1}{\sigma^2}$$

$$\frac{\delta}{t} \int_{\{2 \leq N \leq r\}} N^2 \leq \int_{\{2 \leq N \leq r\}} N + \frac{1}{\sigma^2}$$

$$\{2 \leq N \leq r\} \quad \{2 \leq N \leq r\}$$

$$\frac{\delta}{t} \int_{\{N \leq r\}} N^2 \leq \int_{\{N \leq r\}} N + \frac{1}{\sigma^2}$$

$$\frac{\delta}{t} \left[ \int_{\{N \leq r\}} N \right]^2 \leq \left[ \int_{\{N \leq r\}} N \right] + \frac{1}{\sigma^2}$$

$$\left[ \int_{\{N \leq r\}} N \right] \left[ \frac{\delta}{t} \int_{\{N \leq r\}} (N) - 1 \right] \leq \frac{1}{\sigma^2}$$

$$\frac{1}{\sigma^2} \geq \lim_{r \rightarrow \infty} \left[ \int_{\{N \leq r\}} N \right] \left[ \frac{\delta}{t} \int_{\{N \leq r\}} (N) - 1 \right]$$

$$\frac{1}{\sigma^2} \geq (EN) \left[ \frac{\delta}{t} (EN) - 1 \right]$$

this implies that  $E(N) < \infty$ .

Check on (f): From (4)

$$E(Ny_N) \leq E(N) + \frac{1}{\sigma^2}$$

$$\frac{E(Ny_N)}{E(N)} \leq 1 + \frac{1}{\sigma^2 E(N)}$$

$$\lim_{t \rightarrow \infty} \sup \frac{E(Ny_N)}{E(N)} \leq 1 + \lim_{t \rightarrow \infty} \sup \left( \frac{1}{\sigma^2 EN} \right) = 1$$

since  $E(N) = \infty$  as  $t \rightarrow \infty$

i.e.,

$$\lim_{t \rightarrow \infty} \sup \frac{E(Ny_N)}{E(N)} \leq 1.$$

Thus, the conditions for Lemma 12 are all satisfied; consequently, according to the same Lemma

$$\lim_{t \rightarrow \infty} \frac{E(N)}{t} = 1$$

or

$$\lim_{d \rightarrow 0} \frac{d^2 E(N)}{a^2 \sigma^2} = 1 \text{ since } t = \frac{a^2 \sigma^2}{d^2}.$$

### Conclusion

Given a primary set  $\{X_1, X_2, \dots, X_m\}$  which is bounded by  $\bar{X} - h$  and  $\bar{X} + h$ , also given a preassigned coverage probability  $\alpha$  and a constant width  $2d$ , the confidence interval  $(\hat{\mu}_X - d, \hat{\mu}_X + d)$  for a simple linear regression mean  $\mu_X = \beta_0 + \beta_1(X - \bar{X}_n)$  subject to the non-sequential procedure or sequential procedure has been shown with a property that

$$\lim_{d \rightarrow 0} P\{\hat{\mu}_X - d < \mu_X < \hat{\mu}_X + d\} \geq \alpha.$$

When  $\sigma^2$  is known, where a non-sequential procedure as mentioned in Chapter II is used; when  $\sigma^2$  is unknown, where a sequential procedure as mentioned in Chapter III is used.

However, before this property could be actually put into practice, the lower bound for  $d$  when  $\alpha$  is given must be computed. That is, under a fixed distribution of the random variable  $Y$  or of the random error  $e$  and a probability coverage  $\alpha$ , it is necessary to find the lower bound of  $d$ , say  $d_0$ , so that

$$P\{\hat{\mu}_X - d_0 < \mu_X < \hat{\mu}_X + d_0\}$$

will come to  $\alpha$  to a satisfactory closeness. In this paper, this computation work has not been done.

# BIBLIOGRAPHY

- (1) Snedecor, George. Statistical Methods. 5th ed. Ames, Iowa: Iowa State College Press, 1956.
- (2) Graybill, Franklin A. An Introduction to Linear Statistical Models. New York: McGraw-Hill Book Company, Inc., 1961.
- (3) Scheffé, Henry. The Analysis of Variance. New York: John Wiley and Sons, Inc., 1959.
- (4) Working, Holbrook and Harold Hotelling. "Application of the Theory of Error to the Interpretation of Trends." Journal of the American Statistical Association, 24 (1929), pp. 73-85.
- (5) Hoel, Paul G. "Confidence Regions for Linear Regression." Proceedings of the Second Berhelen Symposium on Mathematical Statistics and Probability. Berkeley, California: University of California Press, 1951, pp. 75-81.
- (6) Graybill, Franklin A., and David C. Bowden. "Linear Segment Confidence Bands for Simple Linear Models." Journal of the American Statistical Association, 62 (1967), pp. 403-408.
- (7) Folks, John Leroy and Charles E. Antle. "Straight Line Confidence Regions for Linear Models." Journal of the American Statistical Association, 62 (December, 1967), pp. 1365-1374.
- (8) Gafarian, A. V. "Confidence Bands in Straight Line Regression." Journal of the American Statistical Association, 59 (1964), pp. 182-213.
- (9) Chow, Y. S., and Herbert Robbins. "On the Asymptotic Theory of Fixed-Width Sequential Confidence Intervals for the Mean." Ann. Math. Statist., 36 (1965), pp. 457-462.
- (10) Gleser, Leon Jay. "On the Asymptotic Theory of Fixed-Size Sequential Confidence Bounds for Linear Regression Parameters." Ann. Math. Statist., 37, (1966), pp. 1053-1055.
- (11) Srivastava, M. S. "On Fixed-Width Bounds for Regression Parameters and Mean Vector." Journal of Royal Stat. Society, B (1967), pp. 132-138.

- (12) Gnedenko, B. V., and A. N. Kalmogorov. Limit Distribution for Sums of Independent Random Variables. (Translated from the Russian by K. L. Chung.) Cambridge: Addison-Wesley, 1954.
- (13) Rao, C. Radhakrishna. Linear Statistical Inference and Its Applications. New York: John Wiley and Sons, Inc., 1968.
- (14) Roussas, George G. A First Course in Mathematical Statistics. Reading, Massachusetts: Addison-Wesley Publishing Company, 1973.
- (15) Ferguson, Thomas S. Mathematical Statistics, A Decision Theoretic Approach. New York: Academic Press, 1967.

## APPENDIX

### EXAMPLES FOR THE DETERMINATION OF SAMPLE SIZE

n WHEN  $\sigma^2$  IS KNOWN AS WELL AS WHEN  
IT IS UNKNOWN

In the following, two examples will be given about the determination of the sample size  $n$  when  $\sigma^2$  is known and when it is unknown.

#### Example 1

Let the primary set be  $\{0, 50\}$ , then the mean and variance associated with this primary set are  $\bar{X} = 25$ ,  $S^2 = 625$ . Let  $h = 25$ ,  $\alpha = 0.95$ ,  $d = 1$  and  $\sigma^2 = 1$ , then  $a$  is determined by

$$a = z_{\frac{1-\alpha}{2}} \sqrt{1 + \frac{h^2}{S^2}} = (1.96) \sqrt{1 + \frac{625}{625}} = (1.96)(1.41) = 2.76.$$

Thus,

$$\frac{a^2 d^2}{\sigma^2} = (2.76)^2 = 7.62.$$

So  $n = 8$ . This means, under the conditions given, eight observations are needed. The determination that  $n = 8$  is independent from  $X$  which is a point in  $\{0, 50\}$ .

#### Example 2

Let the primary set be  $\{0, 50\}$ , then the mean and variance associated with the primary set are  $\bar{X} = 25$ ,  $S^2 = 625$ . Let  $h = 25$ ,  $\alpha = 0.95$ ,  $d = 1$ ,  $X = 30$  and  $\sigma^2$  is unknown. The sample size  $n$  will be determined (so a confidence interval for a simple linear regression mean at  $X = 30$  could be constructed). Here  $a$  is defined the same way as in Example 1. So  $a = 2.76$ . Since  $\sigma^2$  is unknown, the sample size  $n$  will be determined sequentially subject to the stopping rule that  $N = N(t) = k$ , the smallest positive integer such that

$$\hat{\sigma}^2(k) \leq \frac{kd^2}{a_k^2}$$

where

$$\hat{\sigma}^2_{(n)} = \frac{1}{n} \left\{ \Sigma (Y_i - \bar{Y}_n)^2 - \frac{|\Sigma (X_i - \bar{X}_n)(Y_i - \bar{Y}_n)|^2}{\Sigma (X_i - \bar{X}_n)^2} + 1 \right\}$$

and

$$a_k^2 = a^2 \left| 1 + \frac{|X - \bar{X}_n|}{\sqrt{\Sigma (X_i - \bar{X}_n)^2}} \right|^2$$

Here the determination for the sample size  $n$  depends on  $X$  which is a point in  $\{0, 50\}$ .

Start with  $n = 2$ .

The following is the data as well as the computations.

| $X_1$                                 | $Y_1$                                | $X_1 - \bar{X}_2$ | $Y_1 - \bar{Y}_2$ | $(X_1 - \bar{X}_2)^2$                  | $(Y_1 - \bar{Y}_2)^2$                | $(X_1 - \bar{X}_2)(Y_1 - \bar{Y}_2)$                 |
|---------------------------------------|--------------------------------------|-------------------|-------------------|--|--------------------------------------|--|
| 0                                     | 4.00                                 | -25               | -4                | 625                                    | 16                                   | 100  |
| 50                                    | 12.00                                | 25                | 4                 | 625                                    | 16                                   | 100  |
| $\Sigma X_1 = 50$<br>$\bar{X}_2 = 25$ | $\Sigma Y_1 = 16$<br>$\bar{Y}_2 = 8$ |                   |                   | $\Sigma (X_1 - \bar{X}_2)^2$<br>= 1250 | $\Sigma (Y_1 - \bar{Y}_2)^2$<br>= 32 | $\Sigma (X_1 - \bar{X}_2)(Y_1 - \bar{Y}_2)$<br>= 200 |

$$\hat{\sigma}^2(2) = \frac{1}{2} \left\{ 32 - \frac{40,000}{1,250} + 1 \right\} = \frac{1}{2} (32 - 32 + 1) = 0.50$$

$$a_2^2 = (2.76)^2 \left| 1 + \frac{-30 - 25}{1250} \right|^2 = (7.62)(1.34) = 10.21$$

$$\frac{2d^2}{a_2^2} = \frac{2}{10.21} = 0.19$$

Since

$$\hat{\sigma}^2(2) = \frac{2d^2}{a_2^2},$$

so it is necessary to consider  $n = 3$ .

The following is the data as well as the computations.

$$n = 3$$

| $X_i$                                    | $Y_i$                                     | $X_i - \bar{X}_3$ | $Y_i - \bar{Y}_3$ | $(X_i - \bar{X}_3)$                         | $(Y_i - \bar{Y}_3)^2$                     | $(X_i - \bar{X}_3)(Y_i - \bar{Y}_3)$                      |
|--|---|-------------------|-------------------|---|---|---|
| 0  | 4.0                                       | -16.67            | -3.1              | 277.89                                      | 9.61                                      | 51.68   |
| 50                                       | 12.0                                      | 33.33             | 4.9               | 1110.89                                     | 24.01                                     | 163.32  |
| 0  | 5.3                                       | 16.67             | -1.8              | 277.89                                      | 3.24                                      | 30.01   |
| $\Sigma X_i = 50$<br>$\bar{X}_3 = 16.67$ | $\Sigma Y_i = 21.3$<br>$\bar{Y}_3 = 7.10$ |                   |                   | $\Sigma (X_i - \bar{X}_3)^2$<br>$= 1666.67$ | $\Sigma (Y_i - \bar{Y}_3)^2$<br>$= 36.86$ | $\Sigma (X_i - \bar{X}_3)(Y_i - \bar{Y}_3)$<br>$= 245.01$ |

$$\hat{\sigma}^2(3) = \frac{1}{3} \left\{ 36.86 - \frac{(245.01)^2}{1666.67} + 1 \right\} = \frac{1}{3} (36.86 - 36.02 + 1) = \frac{1}{3}(1.84)$$

$$= 0.61$$

$$a_3^2 = (2.76)^2 \left[ 1 + \frac{|30 - 16.67|}{166.67} \right]^2 = (7.62)(1.33) = 10.13$$

$$\frac{3d^2}{a_3^2} = \frac{3}{10.13} = 0.30$$

Since

$$\hat{\sigma}^2(3) > \frac{3d^2}{a_3^2},$$

so it is necessary to consider  $n = 4$ .

The following is the data as well as the computations.

$$n = 4$$

| $X_i$                                  | $Y_i$                                  | $X_i - \bar{X}_4$ | $Y_i - \bar{Y}_4$ | $(X_i - \bar{X}_4)$                      | $(Y_i - \bar{Y}_4)$                       | $(X_i - \bar{X}_4)(Y_i - \bar{Y}_4)$                      |
|--|--|-------------------|-------------------|--|---|---|
| 0                                      | 4.0                                    | -25               | -4.2              | 625                                      | 17.64                                     | 105.00  |
| 50                                     | 12.0                                   | 25                | 3.8               | 625                                      | 14.44                                     | 95.00   |
| 0                                      | 5.3                                    | -25               | -2.9              | 625                                      | 8.41                                      | 72.50   |
| 50                                     | 11.5                                   | 25                | 3.3               | 625                                      | 10.89                                     | 82.50   |
| $\Sigma X_i = 100$<br>$\bar{X}_4 = 25$ | $\Sigma Y_i = 32.8$<br>$\bar{Y} = 8.2$ |                   |                   | $\Sigma (X_i - \bar{X}_4)^2$<br>$= 2500$ | $\Sigma (Y_i - \bar{Y}_4)^2$<br>$= 51.38$ | $\Sigma (X_i - \bar{X}_4)(Y_i - \bar{Y}_4)$<br>$= 355.00$ |

$$\begin{aligned} \hat{\sigma}^2(4) &= \frac{1}{4} \left\{ 51.38 - \frac{(355)^2}{2500} + 1 \right\} = \frac{1}{4} \left\{ 51.38 - \frac{126025}{2500} + 1 \right\} \\ &= \frac{1}{4} (51.38 - 50.41 + 1) = \frac{1}{4} (1.97) = .49 \end{aligned}$$

$$a_4^2 = (2.76)^2 \left[ 1 + \frac{|30 - 25|}{2500} \right]^2 = (7.62)(1.21) = 9.22$$

$$\frac{4d^2}{a_4^2} = \frac{4}{9.22} = 0.43.$$

Since

$$\hat{\sigma}^2(4) > \frac{4d^2}{a_4^2},$$

so it is necessary to consider  $n = 5$ .

The following is the data as well as the computations.

$n = 5$

| $x_i$                                  | $y_i$                                    | $x_i - \bar{x}_5$ | $y_i - \bar{y}_5$ | $(x_i - \bar{x}_5)^2$                       | $(y_i - \bar{y}_5)^2$                     | $(x_i - \bar{x}_5)(y_i - \bar{y}_5)$                      |
|--|--|-------------------|-------------------|---|---|---|
| 0                                      | 4.0                                      | -20.00            | -3.3              | 400.00                                      | 10.89                                     | 66.00   |
| 50                                     | 12.0                                     | 30.00             | 4.7               | 900.00                                      | 22.09                                     | 141.00  |
| 0                                      | 5.3                                      | -20.00            | -2.0              | 400.00                                      | 4.00                                      | 40.00   |
| 50                                     | 11.5                                     | 30.00             | 4.2               | 900.00                                      | 17.64                                     | 126.00  |
| 0                                      | 3.7                                      | -20.00            | -3.6              | 400.00                                      | 12.96                                     | 72.00   |
| $\Sigma x_i = 100$<br>$\bar{x}_5 = 20$ | $\Sigma y_i = 36.5$<br>$\bar{y}_5 = 7.3$ |                   |                   | $\Sigma (x_i - \bar{x}_5)^2$<br>$= 3000.00$ | $\Sigma (y_i - \bar{y}_5)^2$<br>$= 67.58$ | $\Sigma (x_i - \bar{x}_5)(y_i - \bar{y}_5)$<br>$= 445.00$ |

$$\hat{\sigma}^2(5) = \frac{1}{5} \left\{ 67.58 - \frac{(445)^2}{3000} + 1 \right\} = \frac{1}{5} (67.58 - 66.01 + 1) = \frac{1}{5} (2.57) = 0.51$$

$$a_5^2 = (2.76)^2 \left[ 1 + \frac{|30 - 20|}{3000} \right]^2 = (7.62)(1.19) = 9.08$$

$$\frac{5d^2}{2a_5} = \frac{5}{9.08} = 0.55.$$

Thus,

$$\hat{\sigma}^2(5) \leq \frac{5d^2}{2a_5}.$$

By the stopping rule, the sample size  $n$  determined is  $k = 5$ .

Summary for example 2.

| $n$ | $\hat{\sigma}^2(n)$ | $a_n^2$ | $\frac{nd^2}{a_n^2}$ |
|-----|---------------------|---------|----------------------|
| 2   | 0.50                | 10.21   | 0.19                 |
| 3   | 0.61                | 10.13   | 0.30                 |
| 4   | 0.49                | 9.22    | 0.43                 |
| 5   | 0.51                | 9.08    | 0.55                 |

VITA

Jin Tze Chen

Candidate for the Degree of

Doctor of Philosophy

Thesis: SEQUENTIAL AND NON-SEQUENTIAL CONFIDENCE INTERVALS OF CONSTANT WIDTH FOR A SIMPLE LINEAR REGRESSION MEAN

Major Field: Statistics

Biographical:

Personal Data: Born in Sa-La-Chi Hsien, Suiyan, China, August 11, 1929, the son of Min-Da and Shu-Shieu Chen.

Education: Attended Nan-Kai High School, Chungking, China; graduated in 1948; received the Bachelor of Science degree with a major in Mathematics from National Taiwan University, Taipei, Taiwan in July, 1955; received the Master of Science degree with a major in Statistics from Oklahoma State University, Stillwater, Oklahoma in August, 1959; completed the requirements for the Doctor of Philosophy degree in May, 1974.

Professional Experience: Employed as an assistant professor by the Department of Mathematics, Central State University, Edmond, Oklahoma from 1959 to the present.

Professional Organization: Member of the Institute of Mathematical Statistics.