The Annals of Statistics
1974, Vol. 2, No. 3, 535-539

A SEQUENTIAL SOLUTION TO THE INVERSE LINEAR REGRESSION PROBLEM[1]

By S. K. Perng and Yung Liang Tong

Kansas State University and University of Nebraska

In this note we apply the sequential theory developed by Chow and Robbins [1] and Gleser [3], [4] to the inverse linear regression problem. A two-stage sequential procedure has been proposed for the construction of a fixed-width confidence interval for x (an unknown parameter). It is shown that the limiting probabilities of "correct decision" are equal to P^* (pre-assigned).

1. Introduction. Consider the following model of the inverse linear regression problem:

$$Y_{1i} = \alpha + \beta x_i + \varepsilon_{1i} \qquad i = 1, 2, \cdots, n, \cdots \tag{1.1}$$

$$Y_{2j} = \alpha + \beta x + \varepsilon_{2j} \qquad j = 1, 2, \cdots, m, \cdots \tag{1.2}$$

where $\{\varepsilon_{1i}\}$, $\{\varepsilon_{2j}\}$ are two sequences of independent, identically distributed random variables with means zero and finite unknown variances σ_1^2, σ_2^2 respectively, $\{x_i\}$ is a sequence of known constants, α, β and x are unknown. The problem is to estimate x based on the observed $Y_{11}, \cdots, Y_{1n}$ and $Y_{21}, \cdots, Y_{2m}$. The solution of this problem has various statistical applications.

In previous literature ([5], [6]) the point and interval estimation of x has been considered when the random variables are normally distributed, the sample size n and m are fixed, and the variances σ_1^2 and σ_2^2 are equal. Due to the undesirable facts that the mean square error of maximum likelihood estimator of x is infinite, and the length of the confidence interval of x may be infinite, other approaches to the estimation of x should be considered. In this note we apply the sequential sampling rules developed by Chow and Robbins [1] and Gleser [3], [4] to construct a fixed-width confidence interval for x. In the first stage we observe the sequence $\{Y_{1i}\}$ sequentially for the estimation of α and β. If the estimator of β is not significantly different for 0 we do not proceed to the second stage and conclude that this model is not suitable for the estimation of x. Otherwise, we proceed to the second stage to observe $\{Y_{2j}\}$ sequentially. When the experiment terminates, a fixed-width confidence interval for x is constructed so that the probability of coverage is approximately P^* (pre-assigned) when the length of the interval is small.

Received January 1972; revised July 1973.

[1] Contribution number 192. Department of Statistics, Statistical Laboratory, Kansas Agricultural Experiment Station, Manhattan, Kansas 66506.

AMS 1970 *subject classifications*. Primary 62L10.

Key words and phrases. Inverse linear regression problem, two-stage sequential procedure, fixed-width confidence interval.

In Section 2 notations and assumptions are introduced. The sequential procedure is specified in Section 3, and certain asymptotic properties of this procedure are proved in Section 4.

2. Notations and assumptions. $d_1 > 0$, $d_2 > 0$ and $P^* \in (0, 1)$ are specified constants, and a satisfies

$$P^* = \int_{-a}^{a} \frac{1}{(2\pi)^{\frac{1}{2}}} \exp\left(-\frac{x^2}{2}\right) dx .$$

For $n = 1, 2, 3, \cdots$

$$X_n = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{bmatrix}, \tag{2.1}$$

$$\bar{x}(n) = n^{-1} \sum_{i=1}^{n} x_i , \qquad [S(n)]^2 = \sum_{i=1}^{n} (x_i - \bar{x}(n))^2 .$$

For observed $Y_{11}, \cdots, Y_{1n}$ and $Y_{21}, \cdots, Y_{2m}$, $\hat{\alpha}(n)$, $\hat{\beta}(n)$ are the least-squares estimators of α and β respectively based on $Y_{11}, \cdots, Y_{1n}$; $\bar{Y}_1(n) = n^{-1} \sum_{i=1}^{n} Y_{1i}$, $\bar{Y}_2(m) = m^{-1} \sum_{j=1}^{m} Y_{2j}$ are the sample means and

$$\hat{\sigma}_1^2(n) = n^{-1} \sum_{i=1}^{n} [Y_{1i} - \hat{\alpha}(n) - \hat{\beta}(n)x_i]^2 , \tag{2.2}$$

$$\hat{\sigma}_2^2(m) = m^{-1} \sum_{j=1}^{m} (Y_{2j} - \bar{Y}_2(m))^2 \tag{2.3}$$

are estimators of σ_1^2, σ_2^2 respectively.

Throughout this note we shall make the following assumptions on X_n:

ASSUMPTION A. X_n is of rank 2 for every n.

ASSUMPTION B. There exists a (2×2) positive definite matrix Σ such that

$$\lim_{n\to\infty} n^{-1}(X_n X_n') = \Sigma .$$

We observe that Assumption B is Assumption 3.1 of Gleser [3], which implies

$$\lim_{n\to\infty} [S(n)]^2/n = \theta$$

for some positive real number θ.

3. The procedure and its probability of correct decision. For given constants d_1, d_2 and P^* we state the sequential procedure:

(1) First Stage: (a) Start by observing $Y_{11}, \cdots, Y_{1n_0}$ where $n_0 \geqq 2$ is predetermined. Then sample one at a time and stop according to the stopping variable N where

$$N = \text{the first integer} \quad n \geqq n_0 \quad \text{such that} \tag{3.1}$$
$$(\hat{\sigma}_1^2(n) + n^{-1}) \leqq d_1^2[S(n)]^2/a^2 .$$

(b) If $|\hat{\beta}(N)| < d_1$, then conclude that β is not significantly different from zero and the model is not suitable for estimating x. Otherwise, proceed to the second stage.

(2) Second Stage: (a) Start by observing $Y_{21}, \cdots, Y_{2m_0}$ where $m_0 \geqq 2$ is predetermined. Then with the observed $\hat{\beta}(N)$ from the first stage, sample one at a

time according to the following stopping variable M:

$$M = \text{the integer} \quad m \geqq m_0 \quad \text{such that} \quad (\hat{\sigma}_2^2(m) + m^{-1}) \leqq [d_2\hat{\beta}(N)]^2 m/a^2 . \tag{3.2}$$

(b) When sampling is stopped at $M = m$, construct

$$I = (\hat{x} - d_2, \hat{x} + d_2) \tag{3.3}$$

and conclude that I covers x, where $\hat{x} = [\bar{Y}_2(M) - \hat{\alpha}(N)]/\hat{\beta}(N)$.

Note that the stopping variable M also depends on d_1 through $\hat{\beta}(N)$; and that if $[S(n)]^2$ is replaced by n in (3.1) then the first-stage stopping rule is similar to the stopping rule considered in [3] and the second-stage stopping rule is almost the same as the stopping rule in [1].

It seems reasonable to define a correct decision as not to proceed to the second stage if $\beta = 0$, and to proceed to the second stage and to have $x \in I$ if $\beta \neq 0$. Hence, under the sequential procedure for every α, β, σ_1^2, σ_2^2 and x the probability of correct decision (CD) is

$$P[\text{CD}] = P[|\hat{\beta}(N)| \leqq d_1] \qquad \text{if} \quad \beta = 0 , \tag{3.4}$$

$$P[\text{CD}] = P[|\hat{\beta}(N)| > d_1, x \in I] \qquad \text{if} \quad \beta \neq 0 . \tag{3.5}$$

4. Asymptotic results. In this section we investigate the asymptotic properties of the procedure.

LEMMA 1. *Under the proposed sequential procedure*:

$$\lim_{d_1 \to 0} N = \infty \quad \text{a.s.} , \qquad \lim_{d_1 \to 0} \frac{d_1^2 N}{a^2 \sigma_1^2} = \frac{1}{\theta} \quad \text{a.s.} , \qquad \lim_{d_1 \to 0} \frac{d_1 S(N)}{\sigma_1} = a \quad \text{a.s.} ; \tag{4.1}$$

$$\lim_{d_2 \to 0} M = \infty \quad \text{a.s.} \quad \textit{and} \quad \lim_{d_2 \to 0} \frac{(d_2\hat{\beta}(N))^2 M}{a^2 \sigma_2^2} = 1 \quad \text{a.s.} \tag{4.2}$$

PROOF. The proof follows immediately from Lemma 1 of [1].

LEMMA 2. *Under the stopping rule specified in* (3.1)

$$\lim_{d_1 \to 0} \hat{\alpha}(N) = \alpha \quad \text{a.s.} , \qquad \lim_{d_1 \to 0} \hat{\beta}(N) = \beta \quad \text{a.s.} \tag{4.3}$$

PROOF. Obviously for every $n \geqq 2$, we have

$$\hat{\beta}(n) = \beta + c_n(n)^{-\frac{1}{2}} \sum_{i=1}^{n} b_{ni} Z_i = \beta + U_n \quad \text{(say)} ,$$

where $c_n = \sigma_1 n^{\frac{1}{2}}/S(n)$, $b_{ni} = (x_i - \bar{x}(n))/S(n)$ and $Z_1, \cdots, Z_n, \cdots$ is a sequence of i.i.d. random variables with zero mean and unit variance. Since $c_n \to \sigma_1/(\theta)^{\frac{1}{2}}$ for some $\theta > 0$, applying Lemma 2 of [4] it follows that $U_n \to 0$ a.s. and $\hat{\beta}(n) \to \beta$ a.s. Therefore, by (4.1) we have $\hat{\beta}(N) \to \beta$ a.s. The proof of the a.s. convergence of $\hat{\alpha}(N)$ is similar.

We now prove two theorems regarding the expected sample sizes and the limiting probabilities of correct decision.

THEOREM 1. *For every finite* σ_1^2 *and* σ_2^2,

$$P[N < \infty] = 1\,, \qquad P[M < \infty] = 1\,; \tag{4.4}$$

$$\lim_{d_1 \to 0} \frac{d_1^2(EN)}{a^2\sigma_1^2} = \frac{1}{\theta}\,, \qquad \textit{and} \tag{4.5}$$

$$\lim_{d_2 \to 0} \frac{(d_2\hat{\beta}(N))^2(EM)}{a^2\sigma_2^2} = 1 \qquad \textit{for every observed} \quad \hat{\beta}(N)\,. \tag{4.6}$$

PROOF. (4.4) follows from the a.s. convergences of $\hat{\sigma}_1^2(N)$ to σ_1^2 ([3]) and $\hat{\sigma}_2^2(M)$ to σ_2^2. (4.6) follows from Lemma 3 of [1]. (4.5) also follows from Lemma 3 of [1]; clearly the discussion following the lemma (page 460 in [1]) applies to the proof of (4.5) with the aid of the inequality

$$\begin{aligned}\textstyle\sum_{i=1}^{n+1} [y_i - \hat{\alpha}(n+1) - \hat{\beta}(n+1)x_i]^2 &\geqq \textstyle\sum_{i=1}^{n} [y_i - \hat{\alpha}(n+1) - \hat{\beta}(n+1)x_i]^2 \\ &\geqq \textstyle\sum_{i=1}^{n} [y_i - \hat{\alpha}(n) - \hat{\beta}(n)x_i]^2\,,\end{aligned}$$

which follows from a property of the least squares estimators.

THEOREM 2. *Under the sequential procedure*: (a) *If* $\beta = 0$, *then* $\lim_{d_1 \to 0} P[\mathrm{CD}] = P^*$. (b) *If* $\beta \neq 0$ *and if* $\{Y_{2j}\}$ *defined in* (1.2) *is a sequence of continuous random variables, then* $\lim_{d_2 \to 0} \lim_{d_1 \to 0} P[\mathrm{CD}] = P^*$.

PROOF. If $\beta = 0$, then by Corollary B_2 of [7], $S(N)\hat{\beta}(N)/\sigma_1$ has a limiting standard normal distribution as $d_1 \to 0$. Therefore

$$\lim_{d_1 \to 0} P[\mathrm{CD}] = \lim_{d_1 \to 0} P\left[\left|\frac{S(N)\hat{\beta}(N)}{\sigma_1}\right| \leqq \frac{d_1 S(N)}{\sigma_1}\right] = P^*\,,$$

where the second equality follows from (4.1) and a convergence theorem of Cramér ([2] page 254). This proves (a).

To prove (b) we first define a new stopping variable M^* with $\hat{\beta}(N)$ replaced by β in (3.2) (clearly we would not be able to apply this stopping rule in practice because β is unknown), and show that for fixed $d_2 > 0$, M converges almost surely to M^* as $d_1 \to 0$. Let $w = (y_{11}, y_{12}, \cdots, y_{21}, y_{22}, \cdots)$ be a point in the sample space. Then $M^*(w) = m$ for some m iff

$$T_m(w) \leqq \frac{d_2^2}{a^2}\beta^2 < \min_{m_0 \leqq k < m} T_k(w) \tag{4.7}$$

holds, where $T_k = (k)^{-1}[\hat{\sigma}_2^2(k) + (k)^{-1}]$.
Using the facts that $\hat{\beta}(N)$ converges to β almost surely as $d_1 \to 0$, and T_k is a continuous random variable for every k, we have that for sufficiently small $d_1 > 0$,

$$T_m(w) < \frac{d_2^2}{a^2}\hat{\beta}(N)(w) < \min_{m_0 \leqq k < m} T_k(w)$$

holds with probability one. This implies that for fixed $d_2 > 0$, $\lim_{d_1 \to 0} M = M^*$ a.s. It follows that for fixed $d_2 > 0$, $\bar{Y}_2(M)$ converges a.s. to $\bar{Y}_2(M^*)$ as $d_1 \to 0$. Applying a convergence theorem of [2], page 254, Lemma 1 of [4], (4.3) and the fact that

$$\lim_{d_2\to 0}\left[\frac{d_2|\beta|(M^*)^{\frac{1}{2}}}{\sigma_2}\right] = a \quad \text{a.s.}\,,$$

we have

$$\lim_{d_2\to 0}\lim_{d_1\to 0} P[\mathrm{CD}] = \lim_{d_2\to 0}\lim_{d_1\to 0} P\left[|\hat\beta(N)| > d_1, \left|\frac{\bar Y_2(M)-\hat\alpha(N)}{\hat\beta(N)} - x\right| \leqq d_2\right]$$

$$= \lim_{d_2\to 0} P\left[\left|\frac{\bar Y_2(M^*)-\alpha-\beta x}{\sigma_2/(M^*)^{\frac{1}{2}}}\right| \leqq \frac{d_2|\beta|(M^*)^{\frac{1}{2}}}{\sigma_2}\right] = P^*\,.$$

This completes the proof of the theorem.

5. Acknowledgment. We wish to thank the referees for their helpful comments.

REFERENCES

[1] Chow, Y. S. and Robbins, H. (1965). On the asymptotic theory of fixed-width sequential confidence intervals for the mean. *Ann. Math. Statist.* **36** 457–462.

[2] Cramér, H. (1947). *Mathematical Methods of Statistics*. Princeton Univ. Press.

[3] Gleser, L. J. (1965). On the asymptotic theory of fixed-size sequential confidence bounds for linear regression parameters. *Ann. Math. Statist.* **36** 463–467.

[4] Gleser, L. J. (1966). Correction to "On the asymptotic theory of fixed-size sequential confidence bounds for linear regression parameters." *Ann. Math. Statist.* **37** 1053–1055.

[5] Hoadley, B. (1970). A Bayesian look at inverse linear regression. *J. Amer. Statist. Assoc.* **65** 356–369.

[6] Krutchkoff, R. G. (1967). Classical and inverse regression methods of calibration. *Technometrics* **9** 425–439.

[7] Srivastava, M. S. (1967). On fixed-width confidence bounds for regression parameters and mean vector. *J. Roy. Statist. Soc. Ser. B* **29** 132–140.

[8] Wijsman, R. A. (1965). Review #2643 in *Math. Reviews* **30** 508–509.

Department of Statistics
Kansas State University
Calvin Hall 19
Manhattan, Kansas 66506

Department of Mathematics and Statistics
University of Nebraska
Lincoln, Nebraska 68508