Some Methods of Estimating Regression Parameters

Kiyoshi Takeuchi

1

Recently we developed the two-point method of estimating linear regression parameters by Z. Hellwig in [5]. In [5], assuming two-dimentional normal population, we independently proved theorems concerning the properties of estimators obtained by two-point method proposed by Hellwig, and as our own new results we derived the test statistics and the method of constructing confidence limits about population parameters.

In this paper we will intend to extend the above results to new situations.

2

In this section, assuming that the distribution of population is normal, we will define the density function of two-dimentional normal distribution as follows:

$$f(x,y)=\frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}exp\left[-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{x-\mu_x}{\sigma_x}\right)^2-\frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}+\left(\frac{y-\mu_y}{\sigma_y}\right)^2\right\}\right], \quad (1)$$

where

$$E(x)=\mu_x, \quad E(y)=\mu_y \ ,$$

$$E(x-\mu_x)^2=\sigma_x^2, \quad E(y-\mu_y)^2=\sigma_y^2 \ ,$$
$$E(x-\mu_x)(y-\mu_y)=\rho \ ,$$

In this case, conditional expectation $E(y \mid x)$ will be given as follows:

$$E(y \mid x)=\int_{-\infty}^{\infty} y f(y \mid x)\, dy$$
$$=\alpha+\beta x \ , \tag{2}$$

where $\alpha=\mu_y$ and $\beta=\rho\sigma_y/\sigma_x$.
In this case we regard the x's as fixed variates.

We shall assume that a random sample comprising of $n=2m$ items is drawn from the above population. Then we get pairs of (x_1, y_1), $(x_2, y_2)\cdots$, (x_m, y_m), $(x_{m+1}, y_{m+1})\cdots$, (x_n, y_n). We shall divide the set into two subgroups in such a way that in the first subgroup G_1 any random k points are included, and into the second subgroup G_2 all the remaining $(n-k)$ points are included. Then

$$\left.\begin{aligned} \bar{x}_1&=\sum_{i=1}^{k} x_i/k\ , \quad \bar{x}_2=\sum_{j=k+1}^{n} x_j/(n-k)\ ,\\ \bar{y}_1&=\sum_{i=1}^{k} y_i/k\ , \quad \bar{y}_2=\sum_{j=k+1}^{n} y_j/(n-k) \end{aligned}\right\} \tag{3}$$

are computed. As an estimator for population parameter β, b will be expressed as follows:

$$b=\frac{\bar{y}_2-\bar{y}_1}{\bar{x}_2-\bar{x}_1} \tag{4}$$

or

$$b=\frac{\bar{y}_1-\bar{y}_2}{\bar{x}_1-\bar{x}_2} \ . \tag{5}$$

In this paper we shall use expression (4).

Then we shall derive mathematical expectation and variance of estimator b. In this case we regard the x's as fixed variates.

$$E(b \mid x)=E\left(\frac{\bar{y}_2-\bar{y}_1}{\bar{x}_2-\bar{x}_1} \mid x\right)$$

$$=\frac{1}{\bar{x}_2-\bar{x}_1}E(\bar{y}_2-\bar{y}_1 \mid x)$$

$$=\frac{1}{\bar{x}_2-\bar{x}_1}\left\{E(\bar{y}_2 \mid x_2)-E(\bar{y}_1 \mid x_1)\right\} . \tag{6}$$

While

$$E(\bar{y}_1 \mid x_1)=\frac{1}{k}\sum_1^k E(y_1 \mid x_1) , \tag{7}$$

and

$$E(\bar{y}_2 \mid x_2)=\frac{1}{n-k}\sum_1^{n-k} E(y_2 \mid x_2) . \tag{8}$$

Therefore we need to derive $E(y_1 \mid x_1)$ and $E(y_2 \mid x_2)$.

$$E(y_1 \mid x_1)=\int_{-\infty}^{\infty} y_1 f(y \mid x)\,dy$$

$$=\mu_y+\rho\frac{\sigma_y}{\sigma_x}x_1 , \tag{9}$$

and in the same way

$$E(y_2 \mid x_2)=\mu_y+\rho\frac{\sigma_y}{\sigma_x}x_2 . \tag{10}$$

Therefore

$$E(\bar{y}_2 \mid x_2)=\frac{1}{k}\sum_1^k \left(\mu_y+\rho\frac{\sigma_y}{\sigma_x}x_1\right)$$

$$=\mu_y+\rho\frac{\sigma_y}{\sigma_x}\bar{x}_1 , \tag{11}$$

and in the same way

$$E(\bar{y}_2 \mid x_2)=\mu_y+\rho\frac{\sigma_y}{\sigma_x}\bar{x}_2 . \tag{12}$$

From (6), (11) and (12), we get the following result:

$$E(b \mid x)=\frac{1}{\bar{x}_2-\bar{x}_1}\left(\mu_y+\rho\frac{\sigma_y}{\sigma_x}\bar{x}_2-\mu_y-\rho\frac{\sigma_y}{\sigma_x}\bar{x}_1\right)$$

$$=\rho\frac{\sigma_y}{\sigma_x}=\beta \ . \tag{13}$$

Therefore it is derived that b is an unbiassed estimator for regression coefficient β in population.

Next we shall derive the variance of b.

$$\sigma^2(b \mid x)=E(b-\beta \mid x)^2$$

$$=\frac{1}{(\bar{x}-\bar{x})^2}\sigma^2(\bar{y}_2-\bar{y}_1 \mid x) \tag{14}$$

Since the sample is drawn at random from the population, n pairs of (x_i, y_i) are independently distributed. Thus

$$\sigma^2(b \mid x)=\frac{1}{(\bar{x}_2-\bar{x}_1)^2}\left\{\sigma^2(\bar{y}_2 \mid x_2)+\sigma^2(\bar{y}_1 \mid x_1)\right\} \ . \tag{15}$$

Then we need to derive $\sigma^2(\bar{y}_2 \mid x_2)$ and $\sigma^2(\bar{y}_1 \mid x_1)$.

$$\sigma_2(\bar{y}_1 \mid x_1)=\int^{\infty}(\bar{y}_1-\mu_y)^2 f(y \mid x)\mathrm{d}y$$

$$=\sigma_y(1-\rho^2)/k \ . \tag{16}$$

In the same way

$$\sigma^2(\bar{y}_2 \mid x_2)=\sigma_y{}^2(1-\rho^2)\ /\ (n-k) \ . \tag{17}$$

Therefore we get the following result:

$$\sigma^2(b \mid x)=\frac{\sigma_y{}^2(1-\rho^2)}{(\bar{x}_2-\bar{x}_1)^2}\left(\frac{1}{k}+\frac{1}{n-k}\right) \ . \tag{18}$$

From (18) it will be derived that apart from $(\bar{x}_2-\bar{x}_1)^2$, $\sigma^2(b \mid x)$ will be minimized when $k=n-k$, i.e., $k=m=n/2$. Therefore it will be desirable to take $k=n/2=m$.

When $k=m=n/2$, the above results will be expressed as follows:

$$E(b \mid x)=\rho\frac{\sigma_y}{\sigma_x}=\beta \quad ,$$

$$\sigma^2(b \mid x)=\frac{2\sigma_y{}^2(1-\rho^2)}{m(\bar{x}_2-\bar{x}_1)^2}=\frac{4\sigma_y{}^2(1-\rho^2)}{n(\bar{x}_2-\bar{x}_1)^2} \quad . \tag{19}$$

Then we shall derive that b is a consistent estimator for β. We shall assume the following condition:

$$\left|\frac{(x_1+x_2+\cdots+x_m)-(x_{m+1}+x_{m+2}+\cdots+x_n)}{n}\right|>c_0>0 \ , \tag{20}$$

where c_0 is positive constant. Then it is necessary that the following relation is proved:

$$P_r\{\mid b-\beta \mid > \varepsilon\}\to 1 \ , \tag{21}$$

when $\varepsilon>0$ is given and $n\to\infty$.

According to Чебышев's inequality, we have

$$P_r\{\mid b-\beta \mid \leq\varepsilon\}\geq 1-\frac{4\sigma_y{}^2(1-\rho^2)}{\varepsilon^2 n(\bar{x}_2-\bar{x}_1)^2} \quad . \tag{22}$$

It is easily derived that when ε is given and the condition (20) is valid,

$$\lim_{n\to\infty}\frac{4\sigma^2{}_y(1-\rho^2)}{\varepsilon^2 n(\bar{x}_2-\bar{x}_1)^2}\to 0 \quad . \tag{23}$$

Then we have (21) .

3

From the assumption that distribution of population is normal, regression coefficient b derived in (14) is distributed according to $N(\beta,/4\sigma_y{}^2(1-\rho^2)/\{n(\bar{x}_2-\bar{x}_1)^2\})$. Hence

$$\frac{b-\beta}{\sigma(b \mid x)}=\frac{2(b-\beta)(\bar{x}_2-\bar{x}_1)}{\sqrt{n}\ \sigma_y\sqrt{1-\rho^2}} \tag{24}$$

is distributed according to normalized normal distribution, $N(0,1)$. When σ_y and ρ are known, we can easily test the hypothesis on regression coefficient β.

In general σ_y^2 and ρ are unknown. However, if sample size n is large enough, we can use sample standard deviation s_y for σ_y, and r for ρ, constructing the following statistics:

$$\frac{b-\beta}{\hat{\sigma}(b \mid x)}=\frac{2(b-\beta)(\bar{x}_2-\bar{x}_1)}{\sqrt{n}\ s_y\sqrt{1-r^2}} \quad . \tag{25}$$

Above statistic will be approximately normally distributed with $N(0, 1)$, when sample size is large enough. Using it, we can make inferences about regression coefficient β.

However, when sample size n is not so large, it is not possible to use (25), in order to make exact inferences about the regression coefficient β in population. It is our purpose to construct a valid statistic in such a case. To simplify our discussion, we shall limit ourselves to the case in which the fixed variable x obeys the relation $\Sigma x=0$. This may always be possible by a shift of origir of x.

When x is fixed, $y_1, y_2, \cdots, y_n$ are independently distributed according to $N(\mu_y+\beta x_1,\ \sigma_y^2(1-\rho^2))$, $N(\mu_y+\beta x_2,\ \sigma_y^2(1-\rho^2))$,……, $N(\mu_y+\beta x_n,$ $\sigma_y^2(1-\rho^2))$. We shall consider the following transformation:

$$z=y-\alpha-\beta x\,. \tag{26}$$

It is easily understood that z is independently distributed according to $N(0, \sigma_y^2(1-\rho^2))$.

We shall use the following orthogonal transformation:

$$u_1=\frac{1}{\sqrt{n}}(z_{21}+z_{22}+\cdots+z_{2m}-z_{11}-z_{12}-\cdots-z_{1m})$$

$$=\frac{1}{\sqrt{n}}\left\{\sum_{i=1}^{m}(y_{2i}-\alpha-\beta x_{2i})-\sum_{i=1}^{m}(y_{1i}-\alpha-\beta x_{1i})\right\}$$

$$=\frac{m}{\sqrt{n}}\{(\bar{y}_2-\bar{y}_1)-\beta(\bar{x}_2-\bar{x}_1)\}$$

$$=\frac{m}{\sqrt{n}}(\bar{x}_2-\bar{x}_1)\,(b-\beta)\ , \tag{27}$$

$$u_2=\frac{1}{\sqrt{n}}(z_{21}+z_{22}+\cdots+z_{2m}+z_{11}+z_{12}+\cdots+z_{1m})$$

$$=\sqrt{n}\,(a-\alpha)\ , \tag{28}$$

where $a=\bar{y}$.

$$u_3=\frac{1}{\sqrt{\Sigma(x_2-\bar{x}_2)^2+\Sigma(x_1-\bar{x}_1)^2}}\left\{\Sigma(x_2-\bar{x}_2)\,z_2+\Sigma(x_1-\bar{x}_1)z_1\right\}$$

$$=\frac{1}{\sqrt{\Sigma(x_2-\bar{x}_2)^2+\Sigma(x_1-\bar{x}_1)^2}}\Big[\Sigma(x_2-\bar{x}_2)\,(y_2-\bar{y}_2)+$$

$$\Sigma(x_1-\bar{x}_1)\,(y_1-\bar{y}_1)-\beta\,\{\Sigma(x_2-\bar{x}_2)^2+\Sigma(x_1-\bar{x}_1)^2\}\Big]$$

$$=\frac{\sqrt{n}}{s'_x}\{s'_{xy}-\beta(s'_x)^2\}\ , \tag{29}$$

where

$$(s'_x)^2=\frac{\Sigma(x_1-\bar{x}_1)^2+\Sigma(x_2-\bar{x}_2)^2}{n}\ . \tag{30}$$

$$s'_{xy}=\frac{\Sigma(x_1-\bar{x}_1)\,(y_1-\bar{y}_1)+\Sigma(x_2-\bar{x}_2)\,(y_2-\bar{y}_2)}{n}\ . \tag{31}$$

$$\begin{aligned} u_4&=\sum_{i=1}^{m}a_{42i}\,z_{2i}+\sum_{i=1}^{m}a_{41i}z_{1i}\ ,\\ &\ \vdots\qquad\qquad\ \vdots\\ u_n&=\sum_{i=1}^{m}a_{n2i}z_{2i}+\sum_{i=1}^{m}a_{n1i}z_{1i}\ . \end{aligned} \tag{32}$$

Coefficients in the equations (32) are assumed to be so selected as to give an orthogonal transformation.

As z_i is independently distributed according to $N(0,\ \sigma_y^2(1-\rho^2))$, it is easily derived that u_i is also independently distributed according to $N(0,\ \sigma_y^2(1-\rho^2))$. From the above results, it is derived that

$$a=\frac{u_2}{\sqrt{n}}+\alpha \tag{33}$$

$$b=\frac{u_1}{\sqrt{n}\,(\bar{x}_2-\bar{x}_1)}+\beta \tag{34}$$

are independently distributed according to $N(\alpha,\ \sigma_y^2(1-\rho^2)/n)$ and $N(\beta,\ 4\sigma_y^2(1-\rho^2)/n(\bar{x}_2-\bar{x}_1)^2)$, respectively. And they are independently distributed with $(u_4^2+u_5^2+\cdots+u_n^2)/\{\sigma_y^2(1-\rho^2)\}$. In this case

$$\frac{u_4^2+u_5^2+\cdots+u_n^2}{\sigma_y^2(1-\rho^2)} \tag{35}$$

is distributed according to χ^2-distribution with degree of freedom $(n-3)$. Then we obtain the following relation:

$$\begin{aligned} w=\sum_{i=1}^{4}u_i^2 &= \sum_{j=1}^{n}z_j^2-u_1^2-u_2^2-u_3^2 \\ &=\Sigma(y-\alpha-\beta x)^2-\frac{n}{4}(\bar{x}_2-\bar{x}_1)^2(b-\beta)^2- \\ &\quad n(a-\alpha)^2-\frac{n}{(s'_x)^2}\left\{s'_{xy}-\beta(s'_x)^2\right\}^2 \\ &=n(s'_y)^2\left\{1-\frac{(s'_{xy})^2}{(s'_x)^2(s'_y)^2}\right\}, \end{aligned} \tag{36}$$

where

$$(s'_y)^2=\frac{\Sigma(y_2-\bar{y}_2)^2+\Sigma(y_1-\bar{y}_1)^2}{n}. \tag{37}$$

This result is the same form as we derived in [5]. The above result is similar to the variation of error in usual regression line. When we use least squares method in estimating regression parameters, we have the following relation:

$$\begin{aligned} \sum_{i=1}^{n}(y_i-a-bx_i)^2 &= n\,s_y^2\left(1-\frac{s_{xy}^2}{s_x^2\,s_y^2}\right) \\ &=ns_y^2(1-r^2)\ . \end{aligned} \tag{38}$$

In this case degree of freedom is $(n-2)$, but it is noted that our result (36) has degree of freedom $(n-3)$.

As we have already seen, w is independently distributed with a and b, and $w/\{\sigma_y^2(1-\rho^2)\}$ is distributed according to χ^2-distribution with degree of freedom $(n-3)$. $\sqrt{n}(a-\alpha)/(\sigma_y\sqrt{1-\rho^2})$ and $\sqrt{n}\cdot(\bar{x}_2-\bar{x}_1)(b-\beta)/\{2\sigma_y\sqrt{1-\rho^2}\}$ are distributed according to $N(0, 1)$, respectively. Therefore we have the following results:

$$t_1=\frac{\sqrt{n}(a-\alpha)}{\sigma_y\sqrt{1-\rho^2}}\Bigg/\sqrt{\frac{w}{\sigma_y^2(1-\rho^2)}\Big/(n-3)}$$

$$=\sqrt{n}(a-\alpha)\Bigg/\sqrt{\frac{w}{(n-3)}}\ , \tag{39}$$

and

$$t_2=\frac{\sqrt{n}(\bar{x}_2-\bar{x}_1)(b-\beta)}{2\sigma_y\sqrt{1-\rho^2}}\Bigg/\sqrt{\frac{w}{\sigma_y^2(1-\rho^2)}\Big/(n-3)}$$

$$=\frac{\sqrt{n}}{2}(\bar{x}_2-\bar{x}_1)(b-\beta)\Bigg/\sqrt{\frac{w}{(n-3)}} \tag{40}$$

are distributed according to t-distribution with degree of freedom $(n-3)$, respectively.

Using t_1, we can test the hypothesis on regression coefficient α, and construct confidence interval with respect to α. Also we can test the hypothesis on regression coefficient β, and construct confidence interval with respect to β. However, computation of these statistics is somewhat complex compared with computation of estimators themselves.

4

In sections 2 and 3 we estimated α and β on the basis of dividing the set of observations into two subgroups G_1 and G_2. G_1 consists of pairs of observations $(x_1, y_1), \cdots, (x_m, y_m)$, and G_2 of pairs of observations $(x_{m+1}, y_{m+1}), \cdots, (x_n, y_n)$.

The same result as above will be derived on the basis of any division of the set of observations into two subgroups G_1 and G_2 of the same size, under the condition that division does not depend on selection of observations (x_i, y_i) $(i=1, 2, \cdots, n)$. When we consider fluctuation of estimator b, it is desirable to take estimator b having minimum standard error. From (19) it is easily understood that maximum $|\bar{x}_2 - \bar{x}_1|$ will lead to minimum standard error of b. Maximum $|\bar{x}_2 - \bar{x}_1|$ will be obtained, when we enumerate observations in an ascending order : $x_1 \leq x_2 \leq \cdots \leq x_n$.(1) If half number of x_n takes sign $+$, and the other half number takes sign $-$, maximum value of $|\bar{x}_2 - \bar{x}_1|$ will be obtained, when m smallest members have the same signs, but the other m largest members have other signs. Although such arrangements bring dependence on observation, practically we could use the preceding formula when n is large and variance of estimator is not large.

5

With respect to the method of dividing the set of observations into two subgroups G_1 and G_2, there is a difference between two-point method [2], [5] and the preceding method. However, it is

(1) For example, see Ю. В. Линник [3] стр. 301.

noted that on the remaining points there are similarities between them.

According to Z. Hellwig [2], estimation of regression parameters by two-point method is made as follows. A random sample comprising of n items is drawn from two-dimentional general population of which regession lines are assumed to be straight. After $\bar{x}$ and $\bar{y}$ are computed, he divides the set into two subgroups in such a way that in the first subgroup the points with abscissa x not greater than $\bar{x}$ are included, and into the second subgroup all the remaining points are included. If in the second subgroup there are n_2 points, then in the first subgroup there will be $n_1=n-n_2$ points. In this case n_2 is a random variable which may assume the values 1, 2, $\cdots$, $n-1$. The followings are defined:

$$\begin{aligned} x_1&=x \mid x\leq\bar{x}, \quad y_1=y \mid x\leq\bar{x}, \\ x_2&=x \mid x>\bar{x}, \quad y_2=y \mid x>\bar{x}. \end{aligned} \tag{41}$$

Then the following statistics are computed:

$$\begin{aligned} \bar{x}_1&=\frac{1}{n-n_2}\Sigma x_1, \quad \bar{x}_2=\frac{1}{n_2}\Sigma x_2, \\ \bar{y}_1&=\frac{1}{n-n_2}\Sigma y_1, \quad \bar{y}_2=\frac{1}{n_2}\Sigma y_2. \end{aligned} \tag{42}$$

As the estimator b for population parameter β, any one of the following three formulae is utilized:

$$b=\frac{\bar{y}_2-\bar{y}}{\bar{x}_2-\bar{x}}, \tag{43}$$

$$b=\frac{\bar{y}-\bar{y}_1}{\bar{x}-\bar{x}_1}, \tag{44}$$

$$b=\frac{\bar{y}_2-\bar{y}_1}{\bar{x}_2-\bar{x}_1}. \tag{45}$$

An estimator a for population parameter α is expressed in one of the following three ways:

$$a=\bar{y}-b\bar{x}_2\,, \tag{46}$$

$$a=\bar{y}-b\bar{x}_1\,, \tag{47}$$

$$a=\bar{y}-b\bar{x}\,. \tag{48}$$

Then he gives, without proofs, theorems on consistency, unbiasedness, and efficiency of estimator b. However, he does not derive exact distributions of estimators a and b. Therefore he does not make such inferences as statistical testing or confidence limits about population parameters. In [5] We have derived test statistics and the method of confidence limits about population parameters, assuming two-dimentional normal population. The main results, corresponding to (39) and (40), are as follows:

$$t_1=\sqrt{n}\,(a-\alpha)\Bigg/\sqrt{\frac{w}{n-3}} \tag{49}$$

and

$$t_2=\sqrt{\frac{n_1n_2}{n}}(\bar{x}_2-\bar{x}_1)\;(b-\beta)\Bigg/\sqrt{\frac{w}{n-3}} \tag{50}$$

are distributed according to t-distribution with degree of freedom, $(n-3)$, respectively. In this case

$$a=\bar{y}\;,\qquad b=\frac{\bar{y}_2-\bar{y}_1}{\bar{x}_2-\bar{x}_1}\;, \tag{51}$$

$$\bar{x}_1=\frac{\Sigma x_1}{n_1}\;,\qquad \bar{x}_2=\frac{\Sigma x_2}{n_2}\;, \tag{52}$$

$$w=n(s'_y)^2\left\{1-\frac{(s'_{xy})^2}{(s'_x)^2(s'_y)^2}\right\}\;, \tag{53}$$

$$(s'_x)^2=\frac{\Sigma(x_1-\bar{x}_1)^2+\Sigma(x_2-\bar{x}_2)^2}{n}\;, \tag{54}$$

$$(s'_y)^2 = \frac{\Sigma(x_1-\bar{x}_1)^2+\Sigma(x_2-\bar{x}_2)^2}{n}, \tag{55}$$

$$s'_{xy} = \frac{\Sigma(x-\bar{x}_1)(y_1-\bar{y}_1)+\Sigma(x_2-\bar{x}_2)(y_2-\bar{y}_2)}{n}, \tag{56}$$

$$\Sigma x=0, \quad \bar{x}=0,$$

$$\begin{aligned} &x_1=x \mid x\leq 0, \quad y_1=y \mid x\leq 0, \\ &x_2=x \mid x>0, \quad y_2=y \mid x>0. \end{aligned} \tag{57}$$

6

There are several methods to divide the set of observations into two subgroups G_1 and G_2. In [4] J. Ogawa shows that in an unpublished manuscript G. Brown has proposed an estimator of the correlation coefficient. His method is as follows. x and y are distributed according to (1) with equal variances σ^2 and means equal to zero. Equality of variances is not necessary, but the ratio of the variances must be known. Only those observations for which $x_1>k\sigma$ are retained, and then the following statistics is introduced:

$$\rho_B = \frac{\bar{y}_+ - \bar{y}_-}{\bar{x}_+ - \bar{x}_-}, \tag{58}$$

where $\bar{y}_+$ and $\bar{x}_+$ are the averages of the n_1 x's and y's for which $x_i>k\sigma$, and $\bar{y}_-$ and $\bar{x}_-$ are similarly defined for the n_2 observations for which $x_i<-k\sigma$. ρ_B is described to be an unbiased estimator for ρ. Regarding the x's as fixed variates it is derived that

$$\sigma^2(\rho_B \mid x) = \frac{(1-\rho^2)\sigma^2}{(\bar{x}_+-\bar{x}_-)^2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right). \tag{59}$$

As easily seen, (58) is similar to (4), (43), (44), and (45). The expression (59) is similar to (18) and (19).

When we divide the set of observations into two subgroups in such a way as described above, we can use the followlng statistic as an estimator for population parameter β:

$$b=\frac{\bar{y}_{+}-\bar{y}_{-}}{\bar{x}_{+}-\bar{x}_{-}} \quad . \tag{60}$$

We shall prove that b defined in (60) is an unbiased estimator for β. Regarding the $x's$ are fixed variates,

$$E(b\mid x)=\frac{1}{\bar{x}_{+}-\bar{x}_{1}}\left\{E(\bar{y}_{+}\mid x_{+})-E(\bar{y}_{-}\mid x_{-})\right\} \quad . \tag{61}$$

In this case

$$E(\bar{y}_{+}\mid x_{+})=\frac{1}{n_1}\Sigma E(y_{+}\mid x_{+})=\int_{-\infty}^{\infty}yf(y\mid x_{+})dy$$

$$=\mu_y+\rho\frac{\sigma_y}{\sigma_x}(\bar{x}_{+}-\mu_x) \quad , \tag{62}$$

and in the same way

$$E(\bar{y}_{-}\mid x_{-})=\mu_y+\rho\frac{\sigma_y}{\sigma_x}(\bar{x}_{-}-\mu_x) \quad . \tag{63}$$

Therefore

$$E(b\mid x)=\frac{1}{\bar{x}_{+}-\bar{x}_{-}}\rho\frac{\sigma_y}{\sigma_x}(\bar{x}_{+}-\bar{x}_{-})$$

$$=\rho\frac{\sigma_y}{\sigma_x}=\beta \quad . \tag{64}$$

Variance of b is derived as follows:

$$\sigma^2(b\mid x)=\frac{\sigma_y^2(1-\rho^2)}{(\bar{x}_{+}-\bar{x}_{-})^2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right) \quad . \tag{65}$$

In the similar way as described above, we can systematically derive test statistics and confidence limits about population parameters.

References

[1] H. W. Alexander, *Elements of Mathematical Statistics*, 1961.

[2] Z. Hellwig, *Linear Regression and its Application to Economics*, 1963.

[3] Ю. В. Линник, Метод Наименьших Квадратов и Основы Математико-Статистцческой Теории Обработки Наблюдений, издание второе, дополненное и исправленное, 1962, стр. 291—303.

[4] *Contributions to Order Statistics*, ed. by A. E. Sarhan, 1962, pp. 290—291.

[5] K. Takeuchi, "*On estimation of regression parameters* (II)," *The Economic Review*, Vol. 14, No. 4, pp. 17—33. (in Japanese).