Ann. Inst. Statist. Math.
34 (1982), Part A, 309-317

MULTI-AUXILIARY REGRESSION ESTIMATION BASED ON CONDITIONAL SPECIFICATION

GRACE ESIMAI AND CHIEN-PAI HAN

(Received Sept. 25, 1979; revised Dec. 18, 1981)

1. Introduction

Consider a $(p+1)\times 1$ random vector $\begin{pmatrix} Y \\ \boldsymbol{X} \end{pmatrix}$ which follows a multivariate normal distribution where Y is a scalar and $\boldsymbol{X}$ is a $p\times 1$ vector ($p\geqq 1$). In estimating the population mean μ_y of Y, it is well known that the precision of the estimator can be increased if $\boldsymbol{X}$ is used as an auxiliary variable. In this paper, we shall consider the linear regression estimator of μ_y. To use the regression estimator, we need to know the population mean, $\boldsymbol{\mu}_x$, of $\boldsymbol{X}$. In certain situations, an investigator may have partial information about $\boldsymbol{\mu}_x$. In order to utilize this partial information, the investigator can perform a preliminary test about the hypothesis H_0: $\boldsymbol{\mu}_x=\boldsymbol{\mu}_0$ versus H_1: $\boldsymbol{\mu}_x\neq\boldsymbol{\mu}_0$ where $\boldsymbol{\mu}_0$ is some constant vector that he believes $\boldsymbol{\mu}_x$ should be.

As an example consider the estimation of the average yield per acre of a certain crop. It is known that the yield is highly correlated with the moisture and nitrogen content of the soil. Hence these can be used as the auxiliary variable $\boldsymbol{X}$. The experimenter usually does not know $\boldsymbol{\mu}_x$ but from the amount of rainfall reported by the weather bureau or other sources and from analysis by some soil scientist, he believes that $\boldsymbol{\mu}_x$ should be $\boldsymbol{\mu}_0$. Once a preliminary sample is available, the investigator may test H_0. He then will use $\boldsymbol{\mu}_0$ in the regression estimator if H_0 is accepted, otherwise he uses the simple mean $\bar{y}$ to estimate μ_y. The estimator resulting from this procedure is usually referred to as a preliminary test estimator. Studies on the efficiency of the preliminary test estimator show that in practice, it is desirable to use the preliminary test estimator when the investigator's prior information is reliable. Preliminary test estimator was first studied by Bancroft [1] and later by Bennett [3], [4], Han [7], [8], Han and Bancroft [10], Kale and Bancroft [12], Kitagawa [13], Mosteller [15] and others. It belongs to the area of inference based on conditional specification. A note and a bibliography on inference based on conditional specification was compiled by Bancroft and Han [2].

Suppose $\begin{pmatrix} Y \\ \boldsymbol{X} \end{pmatrix}$ is distributed as $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu}=\begin{pmatrix} \mu_y \\ \boldsymbol{\mu}_x \end{pmatrix}$, $\boldsymbol{\Sigma}=\begin{pmatrix} \sigma^2 & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}$. Let $(Y_i, X_{1i}, X_{2i}, \cdots, X_{pi})'$, $i=1, \cdots, n$, be a random sample from $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\bar{y}=\frac{1}{n}\sum_{i=1}^{n} Y_i$, $\bar{\boldsymbol{X}}=\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{X}_i$. If $\boldsymbol{\mu}_x$ and $\boldsymbol{\Sigma}$ are known, then an unbiased estimator of μ_y is $\bar{y}+\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{\mu}_x-\bar{\boldsymbol{X}})$ with variance $\frac{1}{n}\{\sigma^2-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\}$. If $\frac{1}{n}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}$ is considerably large, we have an appreciable gain in precision. If $\boldsymbol{\mu}_x$ is unknown but from certain sources, the experimenter expects but is not certain that $\boldsymbol{\mu}_x=\boldsymbol{\mu}_0$, then he may perform a preliminary test of H_0 and construct a regression estimator depending on the result of this test. Without loss of generality we let $\boldsymbol{\mu}_0=\boldsymbol{0}$. The preliminary test estimator is defined as

$$(1.1) \qquad \bar{y}^*=\begin{cases} \bar{y}-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\bar{\boldsymbol{X}}, & \text{if } n\bar{\boldsymbol{X}}'\boldsymbol{\Sigma}_{22}^{-1}\bar{\boldsymbol{X}}\leqq\chi^2_{p,\alpha}, \\ \bar{y}, & \text{if } n\bar{\boldsymbol{X}}'\boldsymbol{\Sigma}_{22}^{-1}\bar{\boldsymbol{X}}>\chi^2_{p,\alpha}, \end{cases}$$

where $\chi^2_{p,\alpha}$ is the $100(1-\alpha)$ percentage point of the Chi-squared distribution with p degrees of freedom and α is the level of significance of the preliminary test. Han [7] studied the estimator $\bar{y}^*$ when $p=1$. This paper considers the general case with $p\geqq1$. The bias, mean squared error (MSE) and relative efficiency of $\bar{y}^*$ are derived in Section 2.

We also consider a regression estimator of μ_y by using a shrunken estimator of the form $\gamma\bar{\boldsymbol{X}}$, $0<\gamma\leqq1$, for $\boldsymbol{\mu}_x$ when prior information that $\boldsymbol{\mu}_x$ is close to $\boldsymbol{\mu}_0$ is available. For the case $p=1$, assuming σ_x^2, σ^2, ρ known, the shrunken regression estimator of μ_y is defined as, letting $\mu_0=0$,

$$(1.2) \qquad \hat{\mu}=\bar{y}-\beta\gamma\bar{X}$$

where $\beta=\frac{\sigma_{xy}}{\sigma_x^2}$, and

$$(1.3) \qquad \text{MSE}(\hat{\mu})=\text{E}(\bar{y}-\beta\gamma\bar{X}-\mu_y)^2.$$

Following Thompson [16], we find the optimal value of γ which minimizes (1.3). This yields the shrunken regression estimator for $p=1$ as

$$(1.4) \qquad \hat{\mu}=\bar{y}-\frac{\beta\bar{X}\sigma_x^2}{n\bar{X}^2+\sigma_x^2}.$$

The case $p=2$ can be treated similarly but the derivations are more difficult. This case will not be treated here. For the case $p\geqq3$, we assume that $\boldsymbol{\Sigma}$ is known and $\boldsymbol{\Sigma}_{22}=I$ and $\sigma^2=1$ without loss of generality. Following James and Stein [11], we use $\bar{\boldsymbol{X}}\left(1-\frac{p-2}{n\bar{\boldsymbol{X}}'\bar{\boldsymbol{X}}}\right)$ as an estimator of $\boldsymbol{\mu}_x$. We then define the shrunken regression estimator for $p\geqq 3$ as

$$\tilde{\mu}=\bar{y}-\boldsymbol{\Sigma}_{12}\bar{\boldsymbol{X}}\frac{p-2}{n\bar{\boldsymbol{X}}'\bar{\boldsymbol{X}}}\,. \tag{1.5}$$

The MSE of $\hat{\mu}$ and $\tilde{\mu}$ and the efficiency of the preliminary test estimator, $\bar{y}^*$, relative to $\hat{\mu}$ and $\tilde{\mu}$ are derived and discussed respectively in Section 3.

2. Bias, MSE and relative efficiency of $\bar{y}^*$

Let $c=\chi^2_{p,\alpha}$ and $A=[n\bar{\boldsymbol{X}}'\boldsymbol{\Sigma}_{22}^{-1}\bar{\boldsymbol{X}}:\ n\bar{\boldsymbol{X}}'\boldsymbol{\Sigma}_{22}^{-1}\bar{\boldsymbol{X}}\leqq c]$ so that the rejection region of the preliminary test is the complement $\bar{A}$. The expected value of $\bar{y}^*$ can be written as

$$\begin{aligned}\mathrm{E}\,(\bar{y}^*)&=\mathrm{E}\,\{(\bar{y}-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\bar{\boldsymbol{X}})|A\}\,\mathrm{P}\,(A)+\mathrm{E}\,(\bar{y}|\bar{A})\,\mathrm{P}\,(\bar{A})\\&=\mathrm{E}\,(\bar{y})-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\,\mathrm{E}\,(\bar{\boldsymbol{X}}|A)\,\mathrm{P}\,(A)\\&=\mu_y+B_1\,.\end{aligned} \tag{2.1}$$

To evaluate $\mathrm{E}\,(\bar{\boldsymbol{X}}|A)\,\mathrm{P}\,(A)$, we express $\mathrm{P}\,(A)$ in terms of the noncentral Chi-squared distribution and then in terms of the normal distribution. Therefore we have

$$\mathrm{P}\,(A)=\int_0^c e^{-\lambda/2}\sum_{j=0}^{\infty}\frac{1}{j!}\left(\frac{\lambda}{2}\right)^j h_{p+2j}(t)dt\,,$$

where $h_{p+2j}(\cdot)$ is the density function of χ^2_{p+2j} and $\lambda=n\boldsymbol{\mu}_x'\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\mu}_x$. Also

$$\mathrm{P}\,(A)=\int\cdots\int_A\left(\frac{2\pi}{n}\right)^{-p/2}|\boldsymbol{\Sigma}_{22}|^{-1/2}\exp\left\{-\frac{n}{2}(\boldsymbol{X}-\boldsymbol{\mu}_x)'\boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{X}-\boldsymbol{\mu}_x)\right\}d\boldsymbol{x}\,.$$

Differentiating the two expressions of $\mathrm{P}\,(A)$ with respect to $\boldsymbol{\mu}_x$ and equating the results, we find

$$B_1=-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\mu}_x H_{p+2}(c\,;\lambda)\,, \tag{2.2}$$

where $H_{p+2}(c\,;\lambda)$ is the comulative distribution function of the noncentral Chi-squared distribution with $p+2$ degrees of freedom and noncentrality parameter λ.

As a partial check, when $c=0$, we always reject the null hypothesis and use $\bar{y}$ and $B_1=0$. When $c=\infty$, $B_1=-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\mu}_x$ which is the bias of always using $\bar{y}-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\bar{\boldsymbol{X}}$. Without loss of generality, we let $\boldsymbol{\Sigma}_{22}=I$ and $\sigma^2=1$. When $p=1$, since B_1 changes sign with $\boldsymbol{\Sigma}_{12}$ and μ_x, we need only study the bias for $\mu_x \geq 0$ and $\rho > 0$. The bias was also studied by Han [7] who expressed it in terms of the cumulative distribution function of the standard normal distribution. However the two expressions are equivalent (see Han [9]). The general behavior of $-B_1$ is as follows. The bias is zero when $\mu_x = 0$. It is an increasing function of ρ but a decreasing function of α. For fixed n_1, α and ρ, the bias increases from zero to a maximum and then decreases to zero as μ_x increases. The values of $-\sqrt{n}B_1$ for $p=2$ and certain values of $\boldsymbol{\Sigma}_{12}$, $\boldsymbol{\mu}_x\sqrt{n}$ and α are given in Table 1. The properties of the bias are found to be identical with those recorded for $p=1$.

Table 1. Values of $-\sqrt{n}B_1$ for $p=2$

α	.05			.20			.50		
$(\boldsymbol{\mu}_x\sqrt{n})'$ \ $\boldsymbol{\Sigma}'_{12}$	$\begin{pmatrix}.5\\.5\end{pmatrix}$	$\begin{pmatrix}-.5\\.7\end{pmatrix}$	$\begin{pmatrix}.7\\.7\end{pmatrix}$	$\begin{pmatrix}.5\\.5\end{pmatrix}$	$\begin{pmatrix}-.5\\.7\end{pmatrix}$	$\begin{pmatrix}.7\\.7\end{pmatrix}$	$\begin{pmatrix}.5\\.5\end{pmatrix}$	$\begin{pmatrix}-.5\\.7\end{pmatrix}$	$\begin{pmatrix}.7\\.7\end{pmatrix}$
(0, 0)	0	0	0	0	0	0	0	0	0
(.5, .5)	.37	.08	.52	.21	.04	.29	.06	.01	.09
(1.0, 1.0)	.58	.12	.82	.27	.06	.38	.07	.01	.10
(1.5, 1.5)	.55	.11	.76	.19	.04	.27	.04	.01	.05
(2.0, 2.0)	.34	.07	.47	.09	.02	.12	.01	0	.02
(2.5, 2.5)	.14	.03	.19	.02	.01	.03	0	0	0
(3.0, 3.0)	.04	.01	.05	0	0	.01	0	0	0

To obtain the MSE of $\bar{y}^*$, we use the equation

$$M_1 = \text{MSE}(\bar{y}^*) = \text{V}(\bar{y}^*) + B_1^2 . \tag{2.3}$$

By using a similar method for the bias, i.e. differentiating the two expressions of P(A) twice, we can evaluate V$(\bar{y}^*)$. We found that

$$M_1 = \frac{1}{n}\sigma^2[1+h_1] , \tag{2.4}$$

where

$$h_1 = \frac{n}{\sigma^2}\Big\{-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\mu}_x\boldsymbol{\mu}_x'\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}H_{p+4}(c\,;\,\lambda) - \frac{1}{n}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}H_{p+2}(c\,;\,\lambda)$$
$$+2\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\mu}_x\boldsymbol{\mu}_x'\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}H_{p+2}(c\,;\,\lambda)\Big\} .$$

We now compare the preliminary test estimator, $\bar{y}^*$, with the usual estimator $\bar{y}$. The relative efficiency of $\bar{y}^*$ to $\bar{y}$ is

$$e_1 = \frac{\text{MSE}(\bar{y})}{\text{MSE}(\bar{y}^*)} = \frac{1}{1+h_1} . \tag{2.5}$$

The selection of the level of the preliminary test such that the relative efficiency is the largest when μ_x equals 0 and is at least as large as some $e_{\min}$ when $\mu_x \neq 0$ was first recommended by Han and Bancroft [10]. The values of $e_{\max}$ and $e_{\min}$ for $p=1$ are given in Han [7]. The values of e_1 for $p=2$ are given in Table 2 for some choices of Σ_{12}, $\mu_x\sqrt{n}$ and α. The values of $e_{\max}$ and $e_{\min}$ for some α, $\Sigma_{12}=(.5, .5)$ and $p=2$ are given in Table 3 which also gives μ_x^*, the value of μ_x about which $e_{\min}$ occurs (to accuracy within 0.05). We note that $-\mu_x^*$ also gives the same values of $e_{\min}$. In general we observe from the Tables that e_1 is maximum when $\mu_x=0$ for fixed n, α and Σ_{12}. $e_{\max}$ is always an increasing function of the absolute value of any components of Σ_{12} and a decreasing function of α while $e_{\min}$ is an increasing function of α. μ_x^* decreases as α increases.

Table 2. Values of e_1 for $p=2$

α	.05			.20			.50		
$(\mu_x\sqrt{n})'$ \ Σ'_{12}	$\binom{.5}{.5}$	$\binom{-.5}{.7}$	$\binom{.7}{.7}$	$\binom{.5}{.5}$	$\binom{-.5}{.7}$	$\binom{.7}{.7}$	$\binom{.5}{.5}$	$\binom{-.5}{.7}$	$\binom{.7}{.7}$
(0, 0)	1.67	2.45	4.64	1.31	1.55	1.88	1.08	1.13	1.18
(.5, .5)	1.15	2.18	1.34	1.05	1.43	1.10	1.01	1.10	1.01
(1.0, 1.0)	.67	1.67	.51	.77	1.23	.63	.92	1.05	.85
(1.5, 1.5)	.51	1.29	.35	.71	1.08	.56	.92	1.02	.85
(2.0, 2.0)	.53	1.09	.36	.79	1.02	.66	.96	1.00	.92
(2.5, 2.5)	.67	1.02	.51	.91	1.00	.84	.99	1.00	.98
(3.0, 3.0)	.86	1.00	.76	.98	1.00	.96	1.00	1.00	1.00

Table 3. Values of $e_{\min}$ and $e_{\max}$ for $p=2$, $\Sigma_{12}=(.5, .5)$

α	.05	.10	.20	.30	.40	.50
$e_{\max}$	1.67	1.50	1.31	1.20	1.13	1.08
$e_{\min}$	.50	.59	.71	.79	.86	.91
μ_x^*	$\binom{1.65}{1.65}$	$\binom{1.60}{1.60}$	$\binom{1.40}{1.40}$	$\binom{1.35}{1.35}$	$\binom{1.35}{1.35}$	$\binom{1.35}{1.35}$

3. MSE and relative efficiencies of $\hat{\mu}$ and $\tilde{\mu}$

We define the MSE of $\hat{\mu}$ as

$$(3.1)\qquad M_2=\mathrm{E}\left[\bar{y}-\frac{\beta\bar{X}\sigma_x^2}{n\bar{X}^2+\sigma_x^2}-\mu_y\right]^2$$
$$=\frac{1}{n}\sigma^2+2\beta^2\mu_x\sigma_x^2\,\mathrm{E}\left\{\frac{\bar{X}}{n\bar{X}^2+\sigma_x^2}\right\}-2\beta^2\sigma_x^2\,\mathrm{E}\left\{\frac{\bar{X}^2}{n\bar{X}^2+\sigma_x^2}\right\}$$

$$+\beta^2\sigma_x^4 \mathrm{E}\left\{\frac{\bar{X}^2}{(n\bar{X}^2+\sigma_x^2)^2}\right\}.$$

The efficiency of the preliminary test estimator $\bar{y}^*$ relative to the shrunken regression estimator $\hat{\mu}$ is

$$e_2=\frac{\mathrm{MSE}\,(\hat{\mu})}{\mathrm{MSE}\,(\bar{y}^*)}=\frac{M_2}{M_1}. \tag{3.2}$$

Without loss of generality, we let $\sigma_x^2=\sigma^2=1$ in the computation of e_2. The Gauss-Hermite quadrature is used to evaluate the above expected values. For the relevant approximation used, one is referred to Davis and Polonsky [6]. The values of e_2 are given in Table 4 for $n=9$ and certain choices of μ_x, ρ, and α. From the table, we observe that e_2 has a maximum greater than unity at $\mu_x=0$. For fixed n, μ_x and α, e_2 is generally a decreasing function of ρ and for fixed n, μ_x and ρ, e_2 is also a decreasing function of α. For fixed n, ρ and α, e_2 first decreases to a minimum, then increases to above unity and then finally drops back to unity as μ_x increases.

Table 4. Values of e_2 for $n=9$

α	.05		.10		.25	
μ_x \ ρ	.7	.9	.7	.9	.7	.9
0	1.143	1.367	1.019	1.042	.855	.733
.3	.830	.727	.822	.716	.834	.733
.6	.644	.533	.732	.632	.897	.845
.9	.703	.606	.844	.778	1.025	1.039
1.6	1.048	1.078	1.058	1.096	1.062	1.103
2.5	1.026	1.043	1.026	1.043	1.026	1.043

We now consider the MSE of $\tilde{\mu}$ when $p\geqq 3$, $\boldsymbol{\Sigma}_{22}=\boldsymbol{I}$ and $\sigma^2=1$ which is

$$M_3=\mathrm{MSE}\,(\tilde{\mu})=\mathrm{E}\left[\bar{y}-\boldsymbol{\Sigma}_{12}\bar{\boldsymbol{X}}\frac{p-2}{n\bar{\boldsymbol{X}}'\bar{\boldsymbol{X}}}-\mu_y\right]^2 \tag{3.3}$$
$$=\frac{1}{n}[1+h_2],$$

where

$$h_2=n\{2(p-2)\boldsymbol{\Sigma}_{12}[\boldsymbol{D}\boldsymbol{\mu}_x'-\boldsymbol{T}]\boldsymbol{\Sigma}_{21}+(p-2)^2\boldsymbol{\Sigma}_{12}\boldsymbol{G}\boldsymbol{\Sigma}_{21}\},$$

with

$$\boldsymbol{D}=\boldsymbol{\mu}_x\,\mathrm{E}\left(\frac{1}{p+2K}\right),$$

$$\boldsymbol{T}=\frac{1}{n}\left\{\mathrm{E}\left(\frac{1}{p+2K}\right)\boldsymbol{I}+\mathrm{E}\left(\frac{1}{p+2K+2}\right)n\boldsymbol{\mu}_x\boldsymbol{\mu}_x'\right\},$$

$$\boldsymbol{G}=\frac{1}{n}\left\{\mathrm{E}\left[\frac{1}{(p+2K)(p+2K-2)}\right]\boldsymbol{I}+\mathrm{E}\left[\frac{1}{(p+2K)(p+2K+2)}\right]n\boldsymbol{\mu}_x\boldsymbol{\mu}_x'\right\},$$

and K has a Poisson distribution with mean $\frac{n}{2}\boldsymbol{\mu}_x'\boldsymbol{\mu}_x$.

The efficiency of $\bar{y}^*$ relative to $\tilde{\mu}$ is

$$e_3=\frac{\mathrm{MSE}\,(\tilde{\mu})}{\mathrm{MSE}\,(\bar{y}^*)}=\frac{1+h_2}{1+h_1}\,.$$

where h_1 is given in (2.4). Table 5 gives the values of e_3 for $p=4$. It is easily seen that e_3 depends on the parameter values through λ, $d_1=\Sigma_{12}\Sigma_{21}$ and $d_2=\sqrt{n}\,\Sigma_{12}\mu_x$ only. Therefore the table is given for several values of d_1, d_2, λ and α. The expectations in $\boldsymbol{D}$, $\boldsymbol{T}$ and $\boldsymbol{G}$ are obtained by the method given in Chao and Straderman [5] and Lepage [14]. For example, if K has a Poisson distribution with mean m, then

$$\mathrm{E}\left(\frac{1}{K+A}\right)=\begin{cases}\dfrac{1}{m}(1-e^{-m}) & \text{if } A=1\\[2ex] \dfrac{1}{m}+\Gamma(A)\left[\left(\dfrac{-1}{m}\right)^{A}e^{-m}-\sum\limits_{i=1}^{A-1}\left(\dfrac{-1}{m}\right)^{i+1}\Big/\Gamma(A-i)\right] & \\ \qquad\qquad \text{if } A=2,3,\cdots\end{cases}$$

and

Table 5. Values of e_3 for $p=4$

d_1	d_2	λ	α		
			0.05	0.10	0.25
0.5	0.5	0.1	1.05	1.01	0.94
		0.2	1.05	1.00	0.93
		0.5	1.03	0.98	0.92
		1.0	1.00	0.96	0.90
		2.0	0·97	0.93	0.88
		5.0	0·92	0.89	0.85
		10.0	0.90	0.89	0.88
		20.0	0.92	0.92	0.92
		30.0	0.94	0.94	0.94
1.0	0.5	0.1	1.53	1.25	0.91
		0.2	1.50	1.22	0.90
		0.5	1.42	1.16	0.86
		1.0	1.31	1.08	0.81
		2.0	1.15	0.96	0.76
		5.0	0.93	0.83	0.72
		10.0	0.84	0.79	0.75
		20.0	0.85	0.84	0.84
		30.0	0.88	0.88	0.88

$$\mathrm{E}\left[\frac{1}{(K+A)(K+A+1)}\right]=\Gamma(A)\left[\left(\frac{-1}{m}\right)^{A}e^{-m}-\sum_{j=1}^{A}\left(\frac{-1}{m}\right)^{j}G_j\Big/\Gamma(A-j+1)\right]$$

where $G_j=\mathrm{E}\left(\frac{1}{K+A-j+1}\right)$.

We observe from Table 5 (and tables not presented here) that the values of e_3 are large when λ is small, i.e. when μ_x is close to the null value. The efficiency decreases when λ increases and falls below unity before it increases. When λ tends to infinity the efficiency goes to one since both estimators reduce to $\bar{y}$. After studying the behavior of e_3, the conclusion is that if the investigator is certain that μ_x is close to the null value, he should use the preliminary test estimator; otherwise the shrunken regression estimator should be used.

We wish to thank the referee for his valuable comments.

UNIVERSITY OF NIGERIA
IOWA STATE UNIVERSITY AND UNIVERSITY OF TEXAS AT ARLINGTON

REFERENCES

[1] Bancroft, T. A. (1944). On biases in estimation due to the use of preliminary tests of significance, *Ann. Math. Statist.*, **15**, 190-204.

[2] Bancroft, T. A. and Han, C. P. (1977). Inference based on conditional specification: A note and a bibliography, *Int. Statist. Rev.*, **45**, 117-127.

[3] Bennett, B. M. (1952). Estimation of means on the basis of preliminary tests of significance, *Ann. Inst. Statist. Math.*, **4**, 31-43.

[4] Bennett, B. M. (1956). On the use of preliminary tests in certain statistical procedures, *Ann. Inst. Statist. Math.*, **8**, 45-52.

[5] Chao, M. T. and Strawderman, W. E. (1972). Negative moments of Poisson random variables, *J. Amer. Statist. Ass.*, **67**, 429-431.

[6] Davis, P. J. and Polonsky, I. (1964). Numerical interpolation, differentiation and integration, *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables* (eds. M. Abramowitz and J. A. Stegun), *Nat. Bur. Stand. Appl. Math. Ser.* 55, 875-924.

[7] Han, C. P. (1973a). Regression estimation for bivariate normal distributions, *Ann. Inst. Statist. Math.*, **25**, 335-344.

[8] Han, C. P. (1973b). Double sampling with partial information on auxiliary variables, *J. Amer. Statist. Ass.*, **68**, 914-918.

[9] Han, C. P. (1975). Some relationships between noncentral chi-squared and normal distributions, *Biometrika*, **62**, 213.

[10] Han, C. P. and Bancroft, T. A. (1968). On pooling means when variance is unknown, *J. Amer. Statist. Ass.*, **63**, 1333-1342.

[11] James, W. and Stein, C. (1960). Estimation with quadratic loss, *Proc. 4th Berkeley Symp. Math. Statist. Prob.*, **I**, 361-379.

[12] Kale, B. K. and Bancroft, T. A. (1967). Inference for some incompletely specified models involving normal approximations to discrete data, *Biometrics*, **23**, 335-348.

[13] Kitagawa, T. (1963). Estimation after preliminary tests of significance, *Univ. Calif. Pub. Statist.*, **3**, 147-186.

[14] Lepage, Y. (1978). Negative factorial moments of positive random variables, *Industr. Math.*, **28**, 95-100.

[15] Mosteller, F. (1948). On pooling data, *J. Amer. Statist. Ass.*, **43**, 231-242.
[16] Thompson, J. R. (1968). Some shrinkage techniques for estimating the mean, *J. Amer. Statist. Ass.*, **63**, 113-122.