Journal of Statistical Theory & Mothods
1995, Vol. 6, pp 1~11

승법모형의 모수추정[1]

장 석 환[2]

요약 승법모형식 $Y_j = \alpha_0 \prod_{k=1}^{p} X_{kj}^{\beta_k} v_j$ 의 모수는 일반적으로 대수변환한 후에최소제곱법에 의하여 추정되나 $E(\exp(\beta_0)) \neq \alpha_0$ 이므로 $\exp(\beta_0)$은 α_0의 편의추정량이다. 본 연구에서는 α_0의 불편추정량을 (1) 최소제곱추정량을 수정하는 방법과 (2) Finney의 결과를 이용하는 방법으로 추정하였고, 이들 추정량의 분산을 비교하여 효율성을 검토하였다. 아울러 벼의 수량과 수량구성요소와의 관계를 설명할 때 승법모형의 이용 가능성을 검토하였다.

주제어 : 승법모형, 모수, Cobb-Douglas

1. 서 언

승법모형(multiplicative model)은 계량경제 연구에서 생산과 수요를 예측할 때 많이 이용되는 모형으로서 그 대표적인 예로서 Cobb-Douglas 생산함수를 들 수 있고 (Teekens & Koerts. 1971), 일반적으로 다음과 같이 나타낸다.

$$Y_j = \alpha_0 \prod_{k=1}^{p} X_{kj}^{\beta_k} v_j \tag{1}$$

식 (1)에서 Y_j는 생산량, $X_k (k = 1,2,\cdots,p)$는 투자재를 나타내며 α_0는 상수항이고, β_k는 X_k 투자재의 탄력성(elasticity)을 나타내는 모수, 그리고 v_j는 대수정규분포를 하는 확률변수이다. 식(1)의 양변을 자연대수로 변형하면

$$y_j = \beta_0 + \sum_{k=1}^{p} \beta_k x_{kj} + u_j \tag{2}$$

로 쓸 수 있으며 여기서 $y_j = \ln Y_j$, $\beta_0 = \ln \alpha_0$, $x_{kj} = \ln X_{kj}$, $u_j = \ln v_j$ 이다. 식(1)의 모수는 일반적으로 식(2)에서 최소제곱 추정량에 의하여 추정되나 α_0의 추정량으로서의 $\hat{\beta}_0$은 편의 추정량이므로 많은 연구자들 (Goldberger 1968, Heien 1968, Bradu & Mundlak 1970, and Teekens & Koerts 1972)의 관심을 끌어 왔으며 이들은 주로

---

[1] 본 연구는 1993년도 계명대학교 비사연구기금으로 이루어졌음.
[2] 대구시 달서구 신당동 1000 계명대학교 통계학과 704-701

Finney(1941)의 $g_m(\cdot)$함수를 이용하여 α_0의 불편추정량을 추정하였다. 특히 Teekens & Koerts (1972)는 여러가지 추정량과 그들의 MSE를 비교하여 효율성을 검토 한바 있고 Bradu & Mundlak (1970)은 $-2 \le t \le 2$ 에서 $g_m(t)$와 $G_m(t)$함수의 값을 표로 구하여 α_0의 불편추정량과 분산을 구하였으나 특히 $t>2.0$이면 $g_m(t)$와 $G_m(t)$는 발산하고 $t<-2.0$이면 0에 수렴하여 모수 추정에 부적합한 것으로 생각된다.

본 연구에서는 승법모형의 모수의 불편추정량과 분산을 추정하고 농업분야에서 쌀의 수량과 수량구성요소와의 관계를 나타내는 모형식으로서의 사용가능성을 알아보고져 한다.

쌀의 품종개량의 목표는 양질의 다수성 품종에 있으며 이를 위해서는 일반적으로 수량구성요소 (X_1=이삭수, X_2=이삭당 벼알 수, X_3=등숙율, X_4=천립중, X_5=정백비율), 내병충성(耐病蟲性)과 미질(米質)을 개선하는데 중점을 두고 있다. 특히 쌀의 수량과 수량 구성 요소의 관계를 연구함에 있어서 주로 직선다중회귀식(multiple linear regression)을 이용하고 있으나 쌀의 수량(Y)은 이론적으로 수량구성요소 ($X_k, k=1,2$ $3,4,5$)를 곱하여 얻을 수 있으므로 식(1)의 모형식으로 생각하는 것이 타당할 것으로 생각된다. 따라서 본 연구에서는 식(1)의 모수에 대한 불편추정량을 구하고 편의 추정량들과의 효율을비교하며 쌀 연구에서 식(1)의 사용가능성을 규명하므로서 시험결과 분석에 필요한 정보를 제공하고져 한다.

2. 통계적 모형

앞에서 언급했드시 쌀의 단위 면적당 수량(收量) 또는 생산량은 단위면적당 이삭수(穗數;X_1) × 이삭당 입수 (穗當 粒數 ;X_2) × 등숙율(登熟率 ;X_3) × 천립중(千粒重;X_4)에 의하여 산출될 수 있다. 단위면적 즉 $1m^2$(또는 1평)당 이삭 수는 벼를 심는 밀도에 따라 다르기 때문에 일반적으로 권장되는 재식밀도는 m^2당 22주(평당 73주)이므로 일반적으로 주당 수수(株當 穗數)를 많이 쓰며 정현비는 벼를 현미로 찌을 때 비율이다. 실험연구에서는 백미는 현미에 정백비율(精白比率)을 92%로 하여 산출한다. 따라서 단위 면적당 쌀 생산량(수량)은 벼 한포기의 수량에 면적의 환산 계수 Q를 곱하여 추정된다. 즉

수량(Y) = Q × [한 포기(株)의 쌀 수량]
= Q × (주당 수수;X_1) × (수당 입수;X_2) × (등숙율;X_3) × (천립중;X_4)

로 산출되므로 각 수량구성 요소 ($X_k; k=1,2,\cdots,4$)가 수량에 관여하는 가중치를 회귀계수 ($\beta_k, k=1,2,\cdots,4$)를 생각하면 결국 쌀의 수량(Y)는 식(1)과 같이 나타낼 수 있겠다.

식(1)과 (2)에서 $v=e^u$이고 u는 정규분포를 한다고 가정하였으므로 Finney(1941)에 의하여 $u \sim N(m, \sigma^2)$일 때 r차 적률(r-th moment)은 다음과 같다.

$$E(v^r) = \exp(mr + r^2\sigma^2/2) \tag{3}$$

또 식(1)에서 $E(v) = \alpha_0 \prod_{k=1}^{4} X_k^{\beta_k}$ 이기 위해서는 $E(v)=1$의 조건이 충족되어야 한다. 따

라서 식(3)에 의하여 u는 식(4)의 정규분포를 하게 된다.

$$u \sim N\,(-\sigma^2/2, \sigma^2) \tag{4}$$

$\xi_j = u_j + \sigma^2/2$ 로 놓으면 식 (2)는

$$y_j = \beta_0^* + \sum \beta_k x_{kj} + \xi_j \tag{5}$$

로 쓸 수 있고 $\beta_0^* = \beta_0 - \sigma^2/2$ 이며 ξ_j는 $N(0, \sigma^2)$인 확률변수임을 알 수 있다.

3. 모수와 분산추정

3.1 모수 추정

(1) 최소제곱추정량의 변형

식(5)를 행렬로 나타내면

$$\underline{y} = \underline{Z}\underline{\beta} + \underline{\xi} \tag{6}$$

과 같고, 여기서 $\underline{y}' = [\,y_1\, y_2 \cdots y_n\,]$, $\underline{Z}$는 첫번째 열이 $\underline{1}_n$, $x_{kj} = \ln X_{kj}$, $k = 1, 2, \cdots, p$, $j = 1, 2, \cdots, n$로 구성된 n×(p+1) 행렬, $\underline{\beta}' = \lfloor \beta_0^* \beta_1 \cdots \beta_k \rfloor$인 모수벡터이며 $\underline{\xi}$는 $N(\underline{0}, \underline{I}\sigma^2)$의 분포를 하는 n×1 열 벡터이다. 식(6)의 $\underline{\beta}$는 잘 알려진바와 같이 최소제곱법에 의하여

$$\hat{\underline{\beta}} = (\underline{Z'Z})^{-1}\underline{Z'y} \tag{7}$$

로 추정되며 그 기대치는 $E(\beta) = \underline{\beta}$로 불편추정량임을 알 수 있다.
여기서 $\underline{\beta}' = \lfloor \beta_0^*\, \beta_1 \cdots \beta_p \rfloor$의 $\hat{\beta}_0^*$는 식(5)에서 $(\beta_0 - \sigma^2/2)$의 불편추정량이므로 β_0의 편의추정량은 아니다(Goldberger 1968, Heien 1968, Thoni 1969, Bradu & Mundlak 1970, Teekens & Koerts 1972). 그러나 $\hat{\beta}_i\,(i = 1, 2, \cdots, p)$는 β_i의 불편추정량이다. 따라서 α_0의 불편추정량을 구하기 위해서 다음과 같이 생각할 수 있을 것이다. 즉 식(1) 과 식(2)에서 $\alpha_0 = \exp(\beta_0)$이고 실제로 α_0는 $\exp(\hat{\beta}_0^*)$로 추정된다. 그러나 $E(\exp(\hat{\beta}_0^*)) = \exp(\beta_0 - \sigma^2/2 + V(\hat{\beta}_0^*)/2)$ 이므로 $\exp(\hat{\beta}_0^*)$은 α_0의 편의추정량이다. 따라서 $E(\hat{\alpha}_0) = \exp(\beta_0)$를 얻기 위해서는 $\hat{\alpha}_0 = \exp(\hat{\beta}_0^*)$의 성질을 살펴볼 필요가 있다. 즉 그 기대치는

$$\begin{aligned} E(\exp(\hat{\beta}_0^*)) &= \exp(\beta_0^*) + V(\hat{\beta}_0^*)/2 \\ &= \exp(\beta_0 - \tfrac{1}{2}\sigma^2 + \tfrac{1}{2} m_{00}\sigma^2) \\ &= \exp(\beta_0 + \tfrac{1}{2}(m_{00} - 1)\sigma^2) \end{aligned} \tag{8}$$

이며, 여기서 m_{00}는 $(\underline{Z'Z})^{-1}$의 첫번째 대각선 원소이다. 그러므로 α_0의 불편추정량은 다음과 같이 구할 수 있을 것이다.

$$\hat{\alpha}_{01} = \exp(\hat{\beta}_0^*)\exp\{-(m_{00} - 1)s^2/2\} \tag{9}$$

여기서 $s^2 = (1/m) \cdot \underline{y}'[\underline{I} - \underline{Z}(\underline{Z'Z})^{-1}\underline{Z'}]\underline{y}$ 이다.

(2) Finney의 결과에 의한 추정

Finney(1941)는 정수 m에 의한 무한급수를

$$g_m(t)=\sum_{r=0}^{\infty}\frac{m^r(m+2r)}{m(m+2)\cdots(m+2r)}\left(\frac{mt}{m+1}\right)^r\frac{1}{r!} \tag{10}$$

로 정의하고

$$E\left[g_m(As^2)\right]=\exp\left[\frac{mA}{m+1}\sigma^2\right] \tag{11}$$

이 됨을 보였다. 이 결과를 이용하면

$$\begin{aligned}E\left[\exp(\hat{\beta}_0^*)g_m(As^2)\right]&=\exp(\beta_0)\cdot\exp\left[(m_{00}-1)/2+mA/(m+1)\right]\sigma^2\\&=\exp(\beta_0+\lambda\sigma^2)\end{aligned} \tag{12}$$

이고, $\lambda=(m_{00}-1)/2+mA/(m+1)$ 이다. 그러므로 α_0의 불편추정량을 $\hat{\alpha}_0$이라고 하면 $\lambda=0$일 때 $\alpha_0=\exp(\beta_0)$이므로 $A=(m+1)(1-m_{00})/2m$이면 $\hat{\alpha}_0$을 추정할 수 있을것이다. 즉,

$$\hat{\alpha}_{02}=\exp(\hat{\beta}_0^*)\cdot g_m\left[\frac{m+1}{2m}(1-m_{\infty})s^2\right] \tag{13}$$

3.2 추정량의 분산과 효율성 비교

식(9)에 의한 α_0의 추정량 $\hat{\alpha}_{01}$의 분산은

$$\begin{aligned}V(\hat{\alpha}_{01})&=E[\hat{\alpha}_{01}^2]-[E(\hat{\alpha}_{01})]^2\\&=\exp(2\beta_0)[\exp(m_{00}\sigma^2)-1]\end{aligned} \tag{14}$$

이며 식(13)에 의한 추정량 $\hat{\alpha}_{02}$의 분산은

$$V(\hat{\alpha}_{02})=\exp(2\beta_0)[\exp(2m_{00}\sigma^2)\cdot G_m\{(1-m_{00})\sigma^2/2\}-1] \tag{15}$$

이다. 식(15)에서

$$G_m(t)=\sum_{r=0}^{\infty}\Gamma(m/2)/\Gamma\{(m/2)+r\}\binom{m+2r-2}{r}t^r \tag{16}$$

로 정의되는 함수이며, Bradu & Mundlak (1970)에서 구할 수 있다. 따라서 이들 추정분산은 다음과 같이 구할 수 있다.

$$\begin{aligned}\hat{V}(\hat{\alpha}_{01})&=\exp(2\hat{\beta}_0)[\exp(m_{00}s^2)-1],\\\hat{V}(\hat{\alpha}_{02})&=\exp(2\hat{\beta}_0)[\exp(2m_{00}s^2)\cdot G_m\{(1-m_{00})s^2/2\}-1]\end{aligned} \tag{17}$$

또 식(8)에 의한 α_0의 추정량 $\tilde{\alpha}_0=\exp(\hat{\beta}_0^*)$는 편의 추정량이므로 $\tilde{\alpha}_0$의 평균제곱오차

(MSE)는 다음과 같다.

$$MSE(\tilde{\alpha}_0) = V(\tilde{\alpha}_0) + \{bias(\tilde{\alpha}_0)\}^2$$
$$= \exp(2\beta_0)[\exp\{(2m_{00}-1)\sigma^2\} - 2\exp\{(m_{00}-1)\sigma^2/2\} + 1] \quad (18)$$

따라서 식(18)을 이용한 $\hat{\alpha}_{01}$과 $\hat{\alpha}_{02}$의 상대효율은 다음과 같이 나타낼 수 있다. 즉

$$R.E.(\hat{\alpha}_{01}) = \frac{\exp\{(2m_{00}-1)\sigma^2\} - 2\exp\{(m_{00}-1)\sigma^2/2\} + 1}{\exp(m_{00}\sigma^2) - 1}$$

$$R.E.(\hat{\alpha}_{02}) = \frac{\exp\{(2m_{00}-1)\sigma^2\} - 2\exp\{(m_{00}-1)\sigma^2/2\} + 1}{\exp(2m_{00}\sigma^2)\cdot G_m\{(1-m_{00})\sigma^2/2\} - 1}$$

3.3 $y = E(y|\underline{x}_j)$ 의 추정

설명변수 $\underline{X}$의 범위안에서 주어진 p개의 설명변수를 $\underline{x}_j' = [1\ x_{1j}, \cdots, x_{pj}]$라고 할때 $\underline{x}_j'$에 의한 y의 예측치는 $\hat{y} = \underline{x}_j'\underline{\beta}$로 추정되지만 $\tilde{y}_j = \exp(\hat{y})$는 편의 추정치이므로 Y_j의 불편추정치는 다음과 같이 구해 진다.

$$E(\tilde{Y}_j) = E(\exp(\underline{x}_j'\hat{\underline{\beta}}))$$
$$= \exp\{\underline{x}_j'\underline{\beta} - (1-\delta)\sigma^2/2\} \quad (19)$$

이고, $\delta = \underline{x}_j'(\underline{Z}'\underline{Z})^{-1}\underline{x}_j$ 이다. 따라서 $\tilde{y}_j$에 편의량 $\exp\{(1-\delta)\sigma^2/2\}$ 의 추정량을 곱해주면 $\hat{Y}_j$의 불편추정량을 구할 수 있을 것이다. 즉 Y_j의 불편추정량을 $\hat{Y}_j$라 하면

$$\hat{Y}_j = \tilde{Y}_j \cdot \exp\{(1-\delta)s^2/2\}$$
$$= \exp\{\underline{x}_j'\hat{\underline{\beta}} + (1-\delta)s^2/2\} \quad (20)$$

과 같이 구해진다.

이와는 달리 식(12)에서와 같이 Finney의 g(·)함수식을 이용하면 B≠0에 대하여

$$E[\exp(\underline{x}_j'\hat{\underline{\beta}})\cdot g_m(Bs^2)] = \exp[\underline{x}_j'\underline{\beta} + \{mB/(m+1) + (\delta-1)/2\}\sigma^2] \quad (21)$$

이므로 $B = \{(m+1)/2m\}(1-\delta)$ 일때 Y_j의 불편추정량을 $\hat{Y}_{fj}$라 하면 다음과 같이 추정됨을 알 수 있다.

$$\hat{Y}_{fj} = \exp(\hat{y})\cdot g_m[(m+1)(1-\delta)s^2/2m] \quad (22)$$

Teekens & Koerts(1972)는 $E(y|\underline{x}_j)$의 새로운 추정량으로서 $\bar{y}_1 = \exp(\underline{x}_j'\hat{\underline{\beta}} + \xi s^2)$와 $\bar{y}_2 = \exp(\underline{x}_j'\hat{\underline{\beta}})\cdot g_m\{(m+1)/m\cdot\xi s^2\}$를 제안하였다. 여기서 ξ는 $MSE(y_{\xi_p}) = MSE\{\exp(\underline{x}_j'\hat{\underline{\beta}} + \xi s^2)\}$를 최소화하는 조건으로 $\xi = (1-3\delta)/2$ 이다. 이는 식(21)의 B를 $B' = (m+1)/m\cdot(1-3\delta)$로 나타낸것과 같은 결과이며 특히 이들은 δ의 값에 따라 $0 < \delta < 1/3$이면 $\bar{y}_2$

를, 그리고 $\delta \geq 1/3$이면 $\bar{y}_l$를 제안하였다. 그러나 여기서 식(20)과 식(22)는 불편추정량이므로 MSE를 생각할 필요는 없으며 $\hat{Y}_j$와 $\hat{Y}_{fj}$의 분산을 최소화하는 δ 값을 추정할 수 있으나 이는 확률변수이므로 추정치의 의미는 없다고 본다. 식(20)과 식(22)에 의한 예측치 $\hat{Y}_j$와 $\hat{Y}_{fj}$의 분산은 각각 다음과 같다. 즉

$$V(\hat{Y}_j) = \exp\{2\underline{x}_j'\underline{\beta} - (1-\delta)\sigma^2/2\}[\exp(2\delta\sigma^2) - \exp\{(1-\delta)\sigma/2\}] \tag{23}$$

$$V(\hat{Y}_{fj}) = \exp(2\underline{x}_j'\underline{\beta})[\exp\{(2\delta-1)\sigma^2\} \cdot G_m\{(1-\delta)\sigma^2/2\} - 1] \tag{24}$$

4. 적용사례

4.1 모수추정과 가설검정

승법모형의 모수추정문제를 실제 농업분야에서 수도(水稻)의 수량(收量)과 수량구성요소와의 관계에서 다루고 추정량의 효율성을 비교 검토하기 위해서 1994년도에 호남농업시험장에서 수행한 벼의 지방적응연락 시험결과(표 1)를 이용하였다.

[표 1]의 품종별 수량구성요소는 3반복의 평균치이며 수량은 편의상 100주(株)당 kg으로 환산하였다. 이 자료를 일반 다중선형회귀식 $y_j = \beta_0 + \sum_{i=1}^{4}\beta_i x_{ij} + \varepsilon_j$에 적합한 결과는 [표 2] 및 [표 3]과 같다.

[표 1]의 자료를 모형식(1)에 적합하기 위하여 대수변환 한 후에 식(6)에 적합시킨 결과는 [표 4] 및 [표 5]와 같다.

[표 2]와 [표 4]에서 보는바와 같이 일반 다중선형회귀모형식과 승법모형인 식(6)에 의한 수량과 수량구성요소와의 관계를 나타내는 통계량으로서 중상관계수는 각각 R^2=0.7938과 R2=0.8391로 비교적 높은 설명도를 보이고 있으나 모형식(6)에서 약간 높았다. 또 [표 3]과 [표 5]에서 보는바와 같이 이삭 수(X_1)에 대한 회귀계수는 유의성이 없으나 수당입수(X_2), 등숙율(X_3) 및 천립중(X_4)은 아주 높은 유의성을 보여 수량을 예측 또는 추정하는데 중요한 요소입을 보여주고 있다. 두 모형에 있어서 특히 $\hat{\beta}_0$과 $\hat{\beta}_3$에서 큰 차이를 보여 t-값에서도 높은 차이를 보였다.

수량구성요소와 수량과의 관계가 이론적으로는 승법모형이 다중선형모형보다 타당할것으로 생각되나 R^2 값이 생각했던 값보다 작게 나타난것은 아마도 수량구성요소의 표본오차에 기인된것으로 생각된다.

4.2 α_0의 불편추정

[표 1]의 자료를 식(6)에 적합할 때 $(\underline{Z}'\underline{Z})^{-1}$는 다음과 같다.

$$(\underline{Z}'\underline{Z})^{-1} = \begin{bmatrix} 310.29585829 & -15.78113839 & -30.11689410 & -18.43571046 & -15.96895784 \\ -15.78113839 & 4.45350234 & 0.82367067 & -1.18025690 & 1.69503859 \\ -30.11689410 & 0.82367067 & 4.52132389 & 0.99881721 & 0.99544757 \\ -18.43571046 & -1.18025690 & 0.99881721 & 7.38591688 & -5.16251583 \\ -15.96895784 & 1.69503859 & 0.99544757 & -5.16251538 & 9.72456875 \end{bmatrix}$$

Table 1. Yield compoments and yield of rice

Variety	No. of Panicles / hill (X_1)	No. of Grains / panicle (X_2)	% of ripened grains (X_3)	1000 grains weight (X_4)	Yield (kg /100hills)
대 청 벼	14.7	82	93.7	24.8	3.05
수원 394	15.7	80	90.7	21.7	3.19
수원 396	16.3	91	89.0	23.7	3.31
이리 399	16.3	93	91.3	20.6	3.66
동 진 벼	10.3	81	96.0	24.6	3.38
이리 402	16.7	78	86.3	25.7	3.55
밀양 117	16.0	89	92.0	21.1	3.09
수원 400	14.7	87	82.3	21.0	2.83
수원 402	14.7	86	92.7	22.8	3.10
수원 405	13.3	116	84.0	22.5	4.16
이리 404	16.3	87	92.7	21.0	3.52
이리 406	15.3	90	91.7	21.2	3.39
이리 407	16.0	91	89.7	22.4	3.79
밀양 122	16.0	85	95.7	22.5	3.52
밀양 123	16.3	93	88.3	21.6	3.37
수원 406	15.3	90	84.0	22.9	3.23
수원 408	15.3	106	84.0	19.8	3.30
수원 410	15.0	88	86.3	22.2	3.42
수원 413	14.3	112	80.3	22.7	4.03
수원 415	14.3	92	58.0	16.6	1.74
이리 410	16.7	85	93.7	21.9	3.35
이리 411	15.3	92	93.7	21.6	3.45
이리 412	15.0	94	92.0	22.6	3.38
이리 413	15.7	88	85.0	21.6	2.97
이리 415	16.7	78	93.0	21.0	3.07
계화 12	16.7	78	95.0	23.9	3.26
밀양 129	15.7	97	85.7	20.2	3.41
밀양 130	15.3	97	89.7	21.6	3.43
밀양 134	15.3	86	91.0	21.6	3.25
밀양 135	14.3	89	89.7	20.8	3.23

Table 2. Analysis of variance table for multiple regression model

SV	DF	SS	MS	F	Prob>F
Model	4	3.8724	0.9681	24.06	0.0001
Error	25	1.0061	0.0403		
Total	29	4.8785	-	R^2=0.7938	

Table 3. Estimates of parameters with their S.E. for testing hypothesis

Parameters	Estimates	SE	t-values	Prob > |t|
β_0	-5.6123	1.0236	-5.483	0.0001
β_1	0.0498	0.0311	1.598	0.1226
β_2	0.0352	0.0046	7.599	0.0001
β_3	0.0312	0.0067	4.631	0.0001
β_4	0.1016	0.0276	3.688	0.0011

Table 4. Analysis of variance when fitted to the equation (6)

SV	DF	SS	MS	F	Prob>F
Model	4	0.5130	0.12824	32.60	0.0001
Error	25	0.0983	0.00393		
Total	29	0.6113	-	R^2=0.8391	

Table 5. Estimates of parameters with their S.E. for testing hypothesis

Parameters	Estimates	SE	t-values	Prob > |t|
β_0	-10.4276	1.1048	-9.438	0.0001
β_1	0.1973	0.1324	1.491	0.1486
β_2	0.9894	0.1334	7.419	0.0001
β_3	0.9670	0.1705	5.673	0.0001
β_4	0.7452	0.1956	3.810	0.0008

α_0의 불편추정량은 식(9)와 식(13)에 의하여 구할 수 있다. 이들 식에서 m_{00}는 $(\underline{Z}'\underline{Z})^{-1}$의 첫번째 대각선 원소이므로 m_{00}=310.29585829이고 m은 오차의 자유도이므로 m=25이다.

식(10)은 다시

$$g_m(t)=\sum_{r=0}^{\infty}\Gamma(m/2)/\Gamma\{(m/2)+r\}\cdot(1/r!)\left[\frac{m^2}{2(m+1)}t\right]^r \qquad (25)$$

로 변형할 수 있으며(Bradu & Mundlak, 1970) 실제로 α_0의 추정치와 분산을 계산할 때 편리하게 쓸 수 있다. 따라서 α_0의 불편추정치와 분산은 [표 6]과 같다.

또 $\tilde{\alpha}_0$의 MSE는 식(18)에 의하여 MSE($\tilde{\alpha}_0$)=7.6628×10^{-9}로 추정되므로 $\hat{\alpha}_{01}$과 $\hat{\alpha}_{02}$이 높은 효율성을 보였고 또한 $\hat{\alpha}_{01}$이 $\hat{\alpha}_{02}$보다 효율면에서 5.2% 정도 우수한 추정량이라고 할 수 있다.

Table 6. Unbiased estimates and variances of α_0

α_0의 불편추정량	추정치	분산 $[\hat{V}(\hat{\alpha}_{0i})]$	$\hat{V}(\hat{\alpha}_{0i})/MSE(\tilde{\alpha}_0)$
$\hat{\alpha}_{01}$	exp(-11.03534136)	2.0906×10^{-9}	0.2728
$\hat{\alpha}_{02}$	exp(-11.04949675)	2.2003×10^{-9}	0.2871

4.3 $E(y|\underline{x})$의 예측

주어진 $\underline{x}_j'$에 의한 Y의 예측값 $\hat{Y}=\exp(\hat{y})=\exp(\underline{x}_j'\hat{\underline{\beta}})$은 편의 추정량이므로 식(20)과 식(22)에 의하여 불편예측량을 구할 수 있다. 밀양 122호의 경우 $\underline{X}_j'$=[16.0 85 95.7 22.5]이므로 $\underline{x}_j'$=[1 2.77259 4.44265 4.56122 3.11352]를 이용하면 $\delta=\underline{x}_j'(\underline{Z}'\underline{Z})^{-1}\underline{x}_j=0.068535362$이다. 따라서 식(20)과 식(22)에 의하여

$$\hat{Y}_j=\exp\{\underline{x}_j'\hat{\underline{\beta}}+(1-\delta)s^2/2\}=\exp(1.247715089)=3.4823769$$

$$\hat{Y}_{fj}=\exp(\underline{x}_j'\hat{\underline{\beta}})\cdot g_m[(m+1)/2m\cdot(1-\delta)s^2]=\{\exp(1.245884761)\}(1.001831861)=3.4823765 \qquad (26)$$

로 같은 결과를 얻는다. 이들 분산은 식(23)과 식(24)에 의하여 $\hat{V}(\hat{Y}_j)=0.028553842$, $\hat{V}(\hat{Y}_{fj})=0.03258397$로 추정되며 식(22)에 의한 예측값이 훨씬 효율적인 것으로 나타났다.

5. 결 론

벼의 수량과 수량구성요소와의 관계를 맺는 연구에서 일반적으로 선형중회귀식을 이용하여왔으나 실제 자료에 식(1)의 승법모형을 적합시킨 결과 모수의 추정치에는 별다른 상대적인 변화없이 보다 높은 결정계수(R^2)얻을 수 있었다. 다만, 승법모형에서 α_0의 최소제곱 추정량은 편의 추정량이라는 결점은 있으나 식(9)와 식(13)에 의하여 불편추정량을 구할 수 있었고, 특히 식(26)에서 보는바와 같이 $\hat{y}_j$과 $\hat{y}_{fj}$ 는 같은것으로 나타났는데 이는 표본수가 커짐에 따라 $g_m(m/(m+1)\cdot t)=e^t$의 관계에 의한 것으로 생각된다. 또한 식(20) 또는 식(22)에 의한 불편추정치 $\hat{Y}_j$과 관측값 Y_j와의 상관계수는 $R^2(y_j,\hat{y}_j)$=0.8780으로 승법모형에 의한 R^2보다 약 4.6% 개선된 것으로 나타났다.

본 연구에서 승법모형식에 적합할 때 R^2 값이 비교적 낮게 나타난 것은 아마도 수량구성요소를 조사하기위한 표본을 추출할 때 소량의 표본을 조사하므로서 나타나는 표본오차가 크기때문인것으로 생각된다. 따라서 수량을 대표할 수 있는 표본을 충분히 조사하면 보다 적합도가 높은 결과를 기대할 수 있으므로 승법모형의 이용이 보다 바람직하다고 판단된다.

사사 본 논문을 위하여 자료를 제공해 주신 호남농업시험장장 조수연박사께 감사드립니다.

참 고 문 헌

Bradu, D. and Mundlak, Y. (1970). Estimation in lognormal linear models. *Journal of American Statistical Association*, Vol. 65, 198 - 211.

Evans, I.G. and Shaban, S.A. (1974). A note an estimation in lognormal models. *Journal of American Statistical Association*, Vol. 69, 779 - 781.

Evans, I.G. and Shaban, S.A. (1976). Point estimation in multiplicative models. *Econometrica*, Vol 44, No. 3, 467 - 473.

Finney, D.J. (1941). On the distribution of a variable whose logarithm is normally distributed. *Journal of the Royal Statistical Society*, Supplement, Vol. 7, No. 2, 155 - 161.

Golderberger, A.S. (1968). The interpretation and estimation of Cobb-Douglas functions. *Econometrica*, Vol. 35, 464 - 471.

Heien, D.M. (1968). A note on log-linear regression. *Journal of American Statistical Association*, Vol. 63, 1034 - 1038.

Teekens, R. and Koerts, J. (1972). Some statistical implications of the log transformation of multiplicative models. *Econometrica*, Vol. 40, 793 - 819.

Thoni, H. (1969). A table for estimating the mean of g lognormal distribution. *Journal of American Statistical Association*, Vol. 64, 632 - 636.

Parameter Estimation in the Multiplicative Models

Suk-Hwan Chang[3]

Abstract The parameters in the multiplicative model $Y_j = \alpha_0 \prod_{k=1}^{p} X_{kj}^{\beta_k} v_j$ are usually estimated by the least squares method after logarithmic transformation, and the least square estimator of α_0 is known to be biased, i.e., $E(\exp(\hat{\beta}_0)) \neq \alpha_0$.

In the present study the unbaised estimators of α_0 are examined (1) by modifying the least squares estimator and (2) by applying the Finney's results. The variances are also compared. In addition it has been observed that multiplicative model can be used to express the relationship beetween rice yield and yield components.

Key words : Multiplicative model, Parameter, Cobb-Douglas

[3] Keimyung University, Department of Statistics, Taegu, Korea 704-701