多變量 正規分布에서의 選拔効果(I)

—遺傳偏差의 比率에 對하여—

申 漢 豊*

表現型 變數 Y가 遺傳變數 X와 環境變數 E로 表示되고 X와 E가 相互獨立이며 각각 다음과 같은 正規分布를 한다고 하자.

$$X \sim N(\mu,\ \sigma^2),\quad E \sim N(0, \omega^2)$$

대체로 $Y \geq y$이거나 $Y \leq y$인 형태일때 遺傳 및 育種的 選拔은 $Y=X+E$의 형태로 나타난다.

롭슨〔3〕은 選拔을 반복하였을때 遺傳變數 X의 平均期待値와 遺傳變數 X의 條件附分布의 영향을 연구하였고 이와같은 一變量分布의 경우 選拔의 効果는 遺傳力과 全分散에 대한 遺傳分散의 比에 달려 있다 하였다.

이러한 選拔模型을 p-次元 空間에 적용하면 遺傳偏差의 比率을 求할 수 있다.

遺傳統計的 理論

一變量分布인 경우 表現型變數 Y는 다음 式과 같다.

$$Y \sim N(\mu, \sigma^2+\omega^2)$$

여기서 Y^*를 $Y^*=\mu+E$로 定義하면 遺傳分散은 $\sigma^2=0$로 되고, X를 μ로써 대치하면 $Y^* \sim N(\mu, \omega^2)$와 같이 된다.

$Y \geq y$가 되는 選拔을 하면 遺傳偏差가 正의 方向 즉 +가 되는데 이는 分布의 上位部位 (upper tails of the distribution)를 뜻하며 이를 數式化하면

$$\begin{aligned} D(y) &= Pr\{Y \geq y\} - Pr\{Y^* \geq y\} \\ &= \{1-\phi[(y-\mu)\ (\sigma^2+\omega^2)^{-\frac{1}{2}}]\} - \{1-\phi[(y-\mu)/\omega]\} \\ &= \phi[\mu(1-h^2)^{-\frac{1}{2}}] - \phi(\mu), \end{aligned}$$

여기에서 $\mu=(y-\mu)(\sigma^2+\omega^2)^{-\frac{1}{2}}$, $h^2=\sigma^2(\sigma^2+\omega^2)^{-1}$, 그리고 ϕ는 累積標準正規分布(cumulative standard normal distribution)이다.

---

* 韓國科學院 技術社會硏究室 硏究員

같은 形態로 $Y \leq y$가 되는 選拔을 하면 遺傳偏差가 負의 方向 즉 一가 되는데

$$H(y)=Pr\{Y \leq y\}-Pr\{Y^* \leq y\}$$

$$=\phi(\mu)-\phi[\mu(1-h^2)^{-\frac{1}{2}}]=-D(y)$$

多變量인 경우 p-次元 벡타 Y는 遺傳變數 X와 環境變數 E의 合으로되는 分布를 하며 X와 E는 相互獨立이다. 이를 數式化하면

$$X \sim N(\mu, \Sigma), \quad E \sim N(0, \Omega),$$

$$Y=X+E \sim N(\mu, \Sigma+\Omega)$$

遺傳變異가 없다고 가정하면 Y^*는 다음과 같다.

$$Y^*=\mu+E \sim N(\mu, \Omega),$$

選拔部位 $R(Y)$에 대한 遺傳偏差의 比率은 $D(R)=Pr\{Y \in R(Y)\}-Pr\{Y^* \in R(Y^*)\}$ $R(Y)=(Y|a'Y \leq k)$인 選拔指數 a'에 대하여 選拔하면

$$a'Y \sim N[a'\mu, \ a'(\Sigma+\Omega)a]\text{와}$$

$$a'Y^* \sim N(a,'\mu \ a'\Omega a)\text{가 되므로}$$

$$D(R)=Pr\{a'Y \geq k\}-rP\{a'Y^* \geq k\}$$

$$=\phi\{u[a'(\Sigma+\Omega)a/(a'\Omega a)]^{\frac{1}{2}}\}-\phi(u),$$

여기서 $u=(k-a'\mu)[a'(\Sigma+\Omega)a]^{-\frac{1}{2}}$이 되며 이것을 一變量分布로 만들 수 있다.

positive definite symmetric matrix M으로 定義되는 elliptic region에 대한 選拔은

$$R(Y)=[Y|(Y-\mu)'M(Y-\mu) \geq k, \ Y \geq \mu],$$

여기서 $Y \geq \mu$는 component-wise inequality이고, non-singular matrices인 G와 H가 存在하여 다음과 같이 된다.

$$G'(\Sigma+\Omega)G=I, \ G'MG=\Lambda$$

$$H'\Omega H=I, \ H'MH=\Gamma$$

여기서 I는 恒等行列(identity matrix)이고, Λ와 Γ는 對角行列(diagonal matrices 對角項(diagonal element)인 λ_i와 ν_i를 가지고 있는 것으로 한다. U_i와 V_i의 成分을 갖

는 標準 正規 벡타(standard normal vector)로 變形하면

$$U=G'(Y-\mu),\quad V=H'(Y^*-\mu)\text{가 된다.}$$

이에 대한 遺傳偏差의 選拔部位는

$$D(R)=Pr\{Y\in R(Y)\}-Pr\{Y^*\in R(Y^*)\}$$
$$=2^{-p}[Pr\{U'\Lambda U\geq k\}-Pr\{V'\Lambda V\geq k\}]$$
$$=2^{-p}[Pr\{\Sigma U_i^2\lambda_i\geq k\}-Pr\{\Sigma V_i^2\gamma_i\geq k\}].$$

上記式의 두개의 quadratic forms는 루벤 [4]과 콧즈 [2]가 論議한 바와 같은 分布에 따른다고 할 수 있을 것이다.

R_0部位가 특별히 다음과 같이 定義될 때, 즉

$$R_0(Y)=[Y\,|\,\phi(Y;\mu,\Omega)\leq\phi(Y;\mu,\ \Sigma+\Omega)],$$

여기서 ϕ는 平均(mean)과 共分散 行列(covariance matrix)를 갖는 正規 p-次元密度(normal p-dimensional density)를 나타내므로 다음 式이 成立된다.

$$R_0(Y)=\{Y\,|\,|\Omega|^{-\frac{1}{2}p}\exp[-\frac{1}{2}(Y-\mu)'\Omega^{-1}(Y-\mu)]$$
$$\leq|\Sigma+\Omega|^{-\frac{1}{2}p}\exp[-\frac{1}{2}(Y-\mu)'(\Sigma+\Omega)^{-1}(Y-\mu)]\}$$
$$=\{Y\,|\,(Y-\mu)'[\Omega^{-1}-(\Sigma+\Omega)^{-1}](Y-\mu)$$
$$\geq p\ln(|\Sigma+\Omega|/|\Omega|)\}$$

上記式의 non-singular matrices인 Σ와 Ω에 대하여 다음과 같은 C 行列이 존재한다.

$$C'(\Sigma+\Omega)C=I,\quad C\Sigma C=\Lambda,\quad C'\Omega C=I-\Lambda,$$

여기서 Λ는 一變量인 경우에 遺傳力과 유사한 λ_i의 要素를 갖는 對角行列(diagonal matrix)이다. 이式을 $U=C'(Y-\mu)$인 標準正規(standard normal)로 變形하면 다음과 같다.

$$Pr\{Y\in R_0(Y)\}=Pr\{U'[(I-\Lambda)^{-1}-I]U\geq q\}$$
$$=Pr\{\sum_i U_i^2\lambda_i/(1-\lambda_i)\geq q\}$$

여기서 $q=p\ln(|\Sigma+\Omega|/|\Omega|=-p\Sigma\ln(1-\lambda_i)$이다. $(1-\lambda_i)^{-\frac{1}{2}}$을 要素로 갖는 W를 對角行列

(diagonal matrix)이라하고 V가 $V=W'C'(Y^*-\mu)$라는 標準化된 벡타(standardized vector) 라면

$$Pr\{Y^*\in R_0(Y^*)\}=Pr\{V'\Lambda V\geq q\}=Pr\{\Sigma V_i^2\lambda_i\geq q\}$$

이 되고 R_0部位의 遺傳偏差의 比率은 다음과 같이 전개된다.

$$D(R_0)=Pr\{\Sigma U_i^2\lambda_i/(1-\lambda_i)\geq q\}-Pr\{\Sigma U_i^2\lambda_i\geq q\}$$

여기서 U_i^2은 自由度가 1인 χ^2分布이고 相互獨立이다. 실제적으로 계산을 할 경우 다음 式의 解를 求하여야 한다.

$$|\Sigma+\Omega-\lambda\Omega|=0$$

다음例는 申[5]이 사탕옥수수 실험의 一部 資料中 乾物量과 쥬스純度를 각각 Y_1 과 Y_2 로 하였을 때 遺傳分散과 環境共分散의 行列의 推定値이다.

實際的 應用

$$\Sigma=\begin{pmatrix} 0.105 & -0.099 \\ -0.099 & 0.210 \end{pmatrix} \qquad \Omega=\begin{pmatrix} 0.806 & -0.344 \\ -0.344 & 7.840 \end{pmatrix}$$

방정식의 根은 $|\Sigma+\Omega-\lambda\Omega|=0$ 으로 解를 求하여

$\lambda_1=0.01448$, $\lambda_2=0.11844$, $q=0.2814$ 로 되고 그라드와 솔로몬[1] 表를 使用하여 Y 値에 대한 R_0를 求하면

$$Pr\{Y\in R_0(Y)\}=Pr\{0.01470U_1^2+0.13436U_2^2\geq 0.2814\}=0.16,$$

$$Pr\{Y^*\in R_0(Y^*)\}=Pr\{0.01448U_1^2+0.11844U_2^2\geq 0.2814\}=0.14,$$

따라서

$D(R_0,\ Y\geq\mu)=(0.16-0.14)/4=0.005$가 되고 $R_0(Y_1)$과 $R_0(Y_2)$를 선발하면

$$D[R_0(Y_1)]=\phi(0.96)-\phi(1.01)=0.010$$

$$D[R_0(Y_2)]=\phi(0.86)-\phi(0.87)=0.003$$

이 된다.

REFERENCES

〔1〕 Grad, A. and Solomon, H. "Distribution of Quadratic Forms and Some Applications," *Annals of Mathematical Statistics* 36(1955), 464—77.

〔2〕 Kotz, S. *et al.*, "Distribution Quadratic Forms: Ⅰ, Central Case Ⅱ, Non-central Case," *Annals of Mathematical Statistics* 38(1967), 823—48.

〔3〕 Robson, D., *Note on Repeated Selection in the Normal Case,* Cornell Univ. Biometrics Unit Memo. Bu.—171—M. (1964).

〔4〕 Ruben, H., "Probability Content of Regions under Spherical Normal Distributions: Ⅲ, the Bivariate Normal Integral," *Annals of Mathematical Statistics* 32 (1961), 171—86.

〔5〕 Shin, H. P., *Gene Action in the Inheritance of Agronomic Traits in Intervarietal Diallel Crosses and Relative Imporance of Gene Effects for Quantitative Characters,* Unpublished Ph. D. Dissertation. Univ. of Hawaii, 1972.

SUMMARY

Genetic Selection Problems under Multivariate Normal Distribution

Han Poong Shin*

This paper discusses several problems in genetic selection under a multivariate normal distribution. Primary forcus of this paper is a proportion of genetic deviates and a distribution of the genetic variate after selection.

* Researcher, The Korea Advanced Institute of Science