原 著

豚において選抜と集団の大きさが分散成分および
育種価の推定に及ぼす影響

佐 藤 正 寛

(独)農業・食品産業技術総合研究機構, 畜産草地研究所, つくば市, 305-0901

(2008 年 11 月 14 日受付, 2008 年 12 月 8 日受理)

要 約 選抜および集団の大きさが REML 法による分散成分および育種価の推定値に及ぼす影響について検討した。分散成分の推定に用いた記録は, モンテ・カルロ法によるコンピュータシミュレーションにより発生させた。雄 1 頭に雌 10 頭を交配し, 一腹から雌雄各 2 頭を育成した後, 選抜と交配を繰り返す世代重複のない 5 世代分の記録を発生させた。繁殖に供した雄の頭数は世代あたり 1~100 頭の 100 通りとした。形質の遺伝率は 0.25 または 0.50 の 2 通りとし, 無作為および表型価による 2 通りの選抜を想定した。世代の効果を母数効果とし, 各条件下で 100 反復の継代記録を発生させ, REML 法により変量効果における分散推定値の 100 反復の平均およびその平均平方誤差の平方根 (SMSE) を算出した。また, 分散の推定値を用いて, 各個体の育種価および母数効果の推定値を算出した。無作為選抜集団における相加的遺伝分散および誤差分散の SMSE は, 表型選抜集団に比べて僅かに小さくなった。集団のサイズが大きくなるとともに, 相加的遺伝分散および誤差分散の SMSE は減少した。母数効果の SMSE は表型選抜集団よりも無作為選抜集団のほうが小さくなった。真の育種価と推定育種価との相関係数は, 無作為選抜集団が表型選抜集団に比べて大きくなった。

緒 言

BLUP 法により種豚候補の遺伝的能力を推定するためには, あらかじめ当該集団の遺伝的パラメーターを推定しておく必要がある。近年, 遺伝的パラメーターの推定には REML 法が広く用いられている。わが国では, 系統造成を行っている試験研究機関や民間ブリーダーから全国規模での種豚能力評価に至るまで, 様々な大きさの集団に

ついて遺伝的パラメーターが推定されている。しかし, 集団の大きさと遺伝的パラメーターの推定値との関係を明確にした例は少ない。また, 選抜集団では遺伝的パラメーターの推定値に偏りの生じる場合が報告されている¹⁾。さらに, 推定値の平均平方誤差の平方根 (SMSE) は, 選抜集団と非選抜集団とでは異なることが知られている¹⁻⁴⁾。

そこで本研究では, 豚の系統造成における選抜形質や全国種豚能力評価において用いられている

Effects of Selection and Population Size on Estimates of Variance Components and Breeding Values in Pigs

M. SATOH

National Institute of Livestock and Grassland Science, Tsukuba-shi 305-0901, Japan

連絡者: 佐藤正寛 (E-mail: hereford@affrc.go.jp Tel. 029-838-8625)

生産形質に焦点をあて、集団の大きさや選抜の有無などの条件を変えたときの分散成分の推定精度および育種価の推定精度について調べた。

方 法

モンテ・カルロ法によるコンピュータシミュレーションにより、継代記録を発生させた。シミュレーションでは、雄 1 頭に対し雌 10 頭を交配し、各雌一腹から生まれた雄 2 頭雌 2 頭を育成し、次世代を無作為または表型価によって選抜するものとした。したがって、雄および雌の選抜率はそれぞれ 5% および 50% である。これを世代の重複なしに 5 世代繰り返した。繁殖集団の大きさは各世代で一定とし、世代あたりの種雄頭数は 1 頭から 100 頭（種雌頭数は 10 頭から 1000 頭）の 100 通りとした。基礎集団 (G0) は記録を持たないものとし、それ以外の世代 (G1~G5) ではすべての個体が 1 形質の 1 記録を持つものとした。基礎集団の表型分散は 1 とし、1 日平均増体重や背脂肪厚等の生産形質を想定し、形質の遺伝率は 0.25 または 0.50 の 2 通りとした。世代の効果を母数効果とし、その大きさは平均 0、分散 0.75 の一様分布によって発生させた値を各世代に無作為に割り当てた。選抜の有無 (2 通り)、集団の大きさ (100 通り) および遺伝率 (2 通り) の 400 通りの組合せにおいて、各条件下で 100 反復のデータセットを無限遺伝子座モデルによって発生させた。すなわち、 i 世代に属する j 番目の個体の記録 y_{ij} を、

$$y_{ij} = F_i + u_j + e_j, \quad u_j \sim N(0, \sigma_a^2), \quad e_j \sim N(0, \sigma_e^2)$$

とした。ここで、 F_i は i 番目の世代の効果（母数効果）、 u_j は j 番目の個体の育種価、 e_j は j 番目の個体の無作為誤差で、それぞれ平均 0、分散 σ_a^2 （相加的遺伝分散）および σ_e^2 （誤差分散）の正規分布に従うものとした。また、G1 以降の u_j の発生には、遺伝子の分離により引き起こされる変動（Mendelian sampling）部分において個体 j の雄親および雌親の近交係数の影響を考慮した⁵⁾。したがって、 u_j は、

$$u_j = \frac{u_{sj} + u_{dj}}{2} + \left(1 - \frac{F_{sj} + F_{dj}}{2}\right)^{1/2} \cdot \varepsilon,$$

$$\varepsilon \sim N\left(0, \frac{\sigma_a^2}{2}\right)$$

で表される。ここで、 u_{sj} 、 u_{dj} はそれぞれ個体 j の雄親および雌親の育種価、 F_{sj} および F_{dj} はそれらの近交係数である。

それぞれのデータごとに分散成分を推定するとともに、その推定値を用いて母数効果の BLUE および育種価の BLUP (EBV) を求めた。次に、100 反復のデータセットごとに σ_a^2 および σ_e^2 の推定値の平均およびその SMSE を算出した。また、最終世代における EBV について、データセットごとに真値 (TBV) との相関係数およびその標準偏差 (SD) を算出した。母数効果の BLUE についてもデータセットごとに SMSE を算出した。

分散成分の推定にはプログラム VCE5⁶⁾、育種価および母数効果の推定にはプログラム MTDFREML⁷⁾を用いた。

結 果

選抜方法、集団のサイズおよび遺伝率を変えたときの σ_a^2 および σ_e^2 における推定値の反復による平均（それぞれ、 σ_a^2 および σ_e^2 とする）およびそれらの SMSE を図 1 に示した。各条件下における σ_a^2 および σ_e^2 は真値の周辺に分布したものの、種雄頭数が少ない（集団のサイズが小さい）とき、推定値の平均は真値とは異なる場合があった。分散成分の SMSE は集団のサイズの増加とともに急速に減少し、種雄頭数が 10~20 頭を超えると、SMSE の減少量は緩やかになった。分散成分の推定値の反復平均において、選抜法の違いによる影響はみられなかった。しかし、分散成分の SMSE は、無作為選抜集団よりも表型選抜集団において大きくなる傾向がみられ、遺伝率が 0.5 のとき ($\sigma_a^2 = \sigma_e^2 = 0.5$)、 σ_a^2 の SMSE は σ_e^2 のそれよりも大きくなる傾向にあった。また、 σ_a^2 の SMSE は遺伝率の低い形質のほうが高い形質よりも小さくなった。

遺伝率の推定値およびその SMSE を図 2 に示した。分散成分の推定値と同様、SMSE は集団のサイズの増加とともに急速に減少し、種雄頭数が 10~20 頭を超えると、SMSE の減少は緩やかになった。遺伝率の SMSE は、無作為選抜集団よりも表型選抜集団において大きくなる傾向がみられた。表型選抜集団では遺伝率の低い形質における

図 1. 分散成分の推定値（相加的遺伝分散：左，誤差分散：右）の平均（上）とその平均平方誤差の平方根（下）

Fig. 1. Estimated additive genetic (left side) and error (right side) variances (above) and their square roots of mean square errors (below)

SMSE が小さくなる傾向にあったが，無作為選抜集団では遺伝率の違いによる影響はみられなかった。

母数効果の SMSE を図 3 に示した。種雄頭数の増加とともに，母数効果の推定値の SMSE は減少した。また，表型選抜集団における SMSE は，無作為選抜集団のそれよりも大きくなる傾向がみられた。

世代効果における SMSE の世代に対する回帰係数を表 1 に示した。集団のサイズが小さい場合，表型選抜集団において回帰係数の 0 からの偏差が統計的に有意になるものがみられた。有意差のみられた回帰係数はいずれも正の値を示した。

集団のサイズが大きくなると，回帰係数の 0 からの偏差に有意性は認められなかった。

シミュレーションによって発生させた最終世代における TBV と EBV との相関係数およびその SD を図 4 に示した。集団のサイズの増大に伴い相関係数は高まり，その SD は減少する傾向にあった。特に，種雄頭数が 10 頭から 20 頭程度まで増加したときの相関係数の増加量および SD の減少量は大きなものであった。無作為選抜は表型選抜に比べて相関係数が高くなった。しかし，その SD には選抜方法による顕著な差はみられなかった。遺伝率の高い形質はそれが低い形質に比べ，相関係数の SD は小さくなる傾向にあった。

図 2. 遺伝率の推定値 (上) とその平均平方誤差の平方根 (下)

Fig. 2. Heritability estimates and their square roots of mean square errors (below)

考 察

REML 法は分散成分の推定法として最も広く用いられている方法の 1 つである。REML 法による分散成分の推定値はパラメーター空間内に存在し、推定のための情報量が多ければ分散成分の推定精度は高くなる等の利点がある^{8,9)}。しかし、豚の系統造成等のように集団のサイズが小さく、育種改良の期間が短い場合には、少ないデータから分散成分を推定する必要がある。佐藤ら¹⁰⁾は系統造成と同等のサイズからなる集団における分散

図 3. 母数効果の推定値における平均平方誤差の平方根

Fig. 3. Square roots of mean square errors of fixed effect estimates

成分の推定精度について検討している。しかし、分散成分の推定精度において、選抜や集団のサイズによる影響には不明確点が多い。

REML 法による分散成分の推定値は、選抜下においても不偏であることが理論的に知られている^{11,12)}。JENSEN と MAO¹³⁾ や SCHENKEL と SCHAEFFER²⁾ は選抜下においても REML 法による分散成分の推定値は不偏であることをシミュレーション実験によって示している。しかし、本研究の結果から、REML 法による σ_a^2 や σ_e^2 は、世代あたりの種雄頭数が 10 頭 (全頭数 = 種雌 100 頭 $\times$ 4 頭/腹 $\times$ 5 世代 = 2000 頭) を下回る場合には真値からの偏りがみられ、種雄頭数が 5 頭 (全頭数 = 1000 頭) 以下の場合には特に大きな偏りがみられた (図 1)。したがって、分散成分を推定する場合、記録を有する個体数が 1000 頭以下の場合には、推定値に偏りの生じる可能性を勘案する必要がある。

PIERAMATI と VAN VLECK³⁾ および VAN TASSELL ら⁴⁾ は、平均のみを母数効果としたモデルにおいて REML 推定値における選抜の影響を調べた結果、選抜集団における分散成分の SMSE は無作為選抜集団のそれよりも小さくなったと報告している。しかし、SCHENKEL と SCHAEFFER²⁾ は同世代のグループを母数効果としたモデルにおいて、選

表 1. 異なる集団の大きさにおける世代効果（母数効果）の平均平方誤差平方根（SMSE）の世代に対する線形回帰係数

Table 1. Linear regressions of square roots of mean square errors (SMSE) of fixed generation effect on generation with different population sizes

Heritability	Selection	Number of breeding males per generation									
		1-20		21-40		41-60		61-80		81-100	
0.25	Random	-0.016	NS	-0.022	NS	-0.007	NS	0.002	NS	0.002	NS
	Phenotype	0.152	**	0.105	*	0.004	NS	0.000	NS	-0.014	NS
0.50	Random	-0.011	NS	0.012	NS	-0.001	NS	-0.003	NS	-0.001	NS
	Phenotype	0.161	**	0.086	NS	0.016	NS	0.001	NS	-0.013	NS

NS : not significant, * : $P < 0.05$, ** : $P < 0.01$.

抜集団は無作為選抜集団よりも分散成分の SMSE が大きくなったと報告しており, SATOH ら¹⁾も同様の結果を得ている。このように, 分散成分の推定値における SMSE には選抜の有無による一貫した傾向はみられない。しかし本研究では, SCHENKEL と SCHAEFFER²⁾や SATOH ら¹⁾の結果と同様, 表型選抜における分散成分の SMSE は無作為選抜のそれに比べて大きくなる傾向がみられた。また, 種雄頭数が少ない範囲では, 分散成分の SMSE は種雄頭数の増加とともに急激に減少した。したがって, 記録を持つ頭数が少ない場合, 分散成分の推定値の真値からの偏りは大きくなる傾向にあり, 特に選抜集団における推定精度の低下が懸念された。また, 分散成分の推定精度だけではなく, 母数効果の推定精度にも同様の傾向がみられたことから, 分散成分の推定には当該集団における選抜の有無が影響していると考えられる。

本研究では世代の効果を母数効果として与えているため, 選抜による遺伝的改良の効果と世代効果との推定値の交絡の可能性が考えられる。もし, このような交絡があるならば, 世代の増加とともに SMSE も増加すると考えられる。実際, 表 1 に示したように, 集団のサイズが小さい場合には, 表型選抜集団において世代効果の SMSE の回帰係数に有意差の認められる場合があった。一方, 集団のサイズが大きい場合には, そのような有意差は認められなかった。したがって, 系統造

成や 1 農場規模の比較的小さな集団において, 選抜形質における遺伝的改良量を EBV によって推定する場合, 実際の改良量よりも過小評価する可能性がある。

選抜候補個体の情報だけで推定した育種価の正確度は, 遺伝率が 0.25 および 0.50 のとき, それぞれ 0.500 および 0.707 となる。これらの値に対する本研究の TBV と EBV の相関係数は, 遺伝率の低い形質で特に大きくなった。このことは, 遺伝率が低い形質ほど血縁個体の情報が利用され, EBV における推定精度の増加量が大きくなるためである。たとえば, 選抜候補個体との血縁係数が 1/4 以上の血縁個体（選抜候補個体とその全・半きょうだい, 両親とその全きょうだい, 祖父母）の情報を用いた家系選抜指数¹⁴⁾によって選抜の正確度を求めると, 遺伝率が 0.25 および 0.50 のとき, それぞれ 0.664 および 0.788 となる（ただし, 近親交配はないものとし, 血縁個体の頭数は本研究に用いた交配比およびきょうだい数に基づく）。これらの値は集団のサイズが十分大きいときの値（図 4）とほぼ一致する。一方, 表型選抜集団では無作為選抜集団よりも TBV と EBV との相関係数は小さくなった。この原因の 1 つとして, 表型選抜集団における σ_a^2 , σ_e^2 および母数効果の SMSE が無作為選抜集団のそれよりも大きくなったことが考えられる。しかし, 佐藤ら¹⁰⁾は, 真値とは異なる遺伝率による EBV の精度は真値を用いたときのそれよりも低くなるものの, その

図 4. 最終世代における真の育種価と推定育種価との相関係数(上)とその標準偏差(下)

Fig. 4. Correlation coefficients (above) between true and estimated breeding values of the trait in the last generation and standard deviations (below)

差は非常に小さいものであったと報告している。このように 1 形質に対する選抜では、分散成分の推定精度が育種価の推定精度に与える影響は小さいと考えられる。一方、表型選抜集団は無作為選抜集団に比べて、相加的遺伝分散の減少量が大きくなることが知られている (Bulmer 効果)¹⁵⁾。したがって、Bulmer 効果による相加的遺伝分散の減少が TBV と EBV との相関係数に影響を与えている可能性があると考えられる。

本研究では遺伝率が 0.25 と 0.50 の 2 通りの形質を仮定した。しかし、わが国における全国規模での種豚能力評価では、生産頭数や 3 週齢一腹体重などの繁殖形質も能力評価の対象形質となっている。繁殖形質では、記録を持つ個体数が少ないと σ_a^2 の推定精度が低くなるため、遺伝率の推定値が 0 になることがある¹⁶⁾。また、記録を持つ個体数が多くても、選抜集団では分散成分の推定値に偏りの生じることがある¹⁾。そのため、このような形質では 2 産次以降の記録を利用することで、より正確に分散成分や育種価を推定する試みがなされている。さらに、このような形質を含む複数形質における遺伝的パラメータの推定精度についての報告例はない。したがって、今後は限性形質を含む複数形質における分散成分の推定精度について検討する必要があると考える。

文 献

- 1) SATOH, M., C. HICKS, K. ISHII and T. FURUKAWA : Choice of statistical model for estimating genetic parameters and direct and maternal genetic effects in swine, *J. Anim. Breed. Genet.*, **119**, 285-296, 2002.
- 2) SCHENKEL, F.S. and L.R. SCHAEFFER : Effects of nonrandom parental selection on estimation of variance components, *J. Anim. Breed. Genet.*, **117**, 225-239, 2000.
- 3) PIERAMATI, C. and L.D. VAN VLECK : Effects of genetic groups on estimates of additive genetic variance, *J. Anim. Sci.*, **71**, 66-70, 1993.
- 4) VAN TASSELL, C.P., G. CASELLA and E.J. POLLAK : Effects of selection on estimates of variance components using Gibbs sampling and restricted maximum likelihood, *J. Dairy Sci.*, **78**, 678-692, 1995.
- 5) BULMER, M.G. : Normal distribution theory : In *The Mathematical Theory of Quantitative Genetics*. Oxford Univ. press, New York, 121-143, 1980.
- 6) BOLDMAN, K., L.A. KRIESE, L.D. VAN VLECK and S.D. KACHMEN : A manual for use of MTDFREML, 1-115, USDA, Washington DC, 1993.

- 7) NEWMAIER, A. and E. GROENEVELD : REML estimation of covariance in sparse linear models, *Genet. Sel. Evol.*, **30**, 3-26, 1998.
- 8) SCHAEFFER, L.R. : Variance component estimation course notes, 62-72, Univ. New England, Armidale, 1998.
- 9) 南美穂子 : 制限付き最尤法 (REML 推定法), *応用統計学*, **25**, 73-78, 1996.
- 10) 佐藤正寛・古川 力・石井和雄 : 豚系統造成において個体数が単一形質の REML 法による分散成分および育種価の推定に及ぼす影響, *日豚会誌*, **36**, 130-135, 1999.
- 11) HENDERSON, C.R. : Recent developments in variance and covariance estimation, *J. Anim. Sci.*, **63**, 208-216, 1986.
- 12) SEARLE, S.R. : Variance components - some history and a summary of estimation methods, *J. Anim. Breed. Genet.*, **106**, 1-29, 1989.
- 13) JENSEN, J. and I.L. MAO : Estimation of genetic parameters using sampled data from populations undergoing selection, *J. Dairy Sci.*, **74**, 3544-3551, 1991.
- 14) 佐藤正寛 : 血縁情報を取り入れた選抜指数を算出するプログラムの開発, *日豚会誌*, **40**, 11-20, 2003.
- 15) BULMER, M.G. : Selection : In *The Mathematical Theory of Quantitative Genetics*. Oxford Univ. press, New York, 144-163, 1980.
- 16) 佐藤正寛 : 豚系統造成における産子数の遺伝的パラメータの推定精度, *日豚会誌*, **46**, 1-6, 2009.

Effects of Selection and Population Size on Estimates of Variance Components and Breeding Values in Pigs

Masahiro SATOH

National Institute of Livestock and Grassland Science, Tsukuba-shi 305-0901, Japan

The effects of selection and population size on estimates of variance components and breeding values using REML estimation were assessed by stochastic computer simulation. One hundred closed breeding herds of one through 100 sires per generation were simulated, assuming six generations including a base population without overlap. A boar was mated to 10 gilts randomly and each mating produced two males and two females as candidates for selection. Two different heritabilities of 0.25 and 0.50 per trait were used. Selection was either at random or based on the phenotypic performance of the individual. Generation constants were generated as a fixed effect. One hundred replicates were simulated for each combination of heritability, population size, and selection scheme. For each replicate, variance component estimates and their square roots of mean square errors (SMSE) were calculated, and breeding values and generation effects were estimated using the estimates of additive genetic and error variances. SMSE of additive genetic and error variance estimates were slightly smaller from randomly selected populations than from phenotypically selected populations. SMSE of additive genetic and error variance estimates decreased with the increase in population size. SMSE of estimates of fixed generation effect from randomly selected populations were smaller than from phenotypically selected populations. For randomly selected populations, correlations between true and estimated breeding values were higher and their SMSE were slightly smaller than phenotypically selected populations.

Jpn. J. Swine Science, **46**, 1 : 17-24

Key words : variance component estimates, breeding value estimates, REML, selection, population size