

Genetikai paraméterek becslése random regresszióval

Márkus Szilárd¹ – Némcová Eva² –
Fazekas István³ – Komlósi István¹

¹Debreceni Egyetem Agrár- és Műszaki Tudományok Centruma,
Mezőgazdaságtudományi Kar, Debrecen

²Research Institute of Animal Production, Prága-Uhříněves,
Csehország

³Debreceni Egyetem Informatikai Kar,
Alkalmazott Matematika és Valószínűségszámítás Tanszék,
Debrecen
szmarkus@agr.unideb.hu

ÖSSZEFOGLALÁS

A random regressziós modellekkel történő tenyésztérbecslés egyik legfontosabb része a variancia komponensek becslése. Mivel ez a művelet igényli a legtöbb erőforrást, számos kutatás foglalkozik a különböző becslési eljárások megvalósításával. A variancia komponensek becslésére használható a csökkentett maximum likelihood módszer (REML), a Gibbs-féle mintavételezés és az \mathcal{R} módszer. A variancia komponensek ismerete szükséges az öröklődhetőségi érték és az ismételhetőség kiszámításához.

Közleményünk célja a variancia komponensek becslése a magyarországi holstein-fríz tejtermelési adatokból random regressziós ismételhetőségi modellel, valamint az additív genetikai és állandó környezeti variancia, továbbá az öröklődhetőségi és ismételhetőségi érték változásának vizsgálata a laktáció során.

Kulcsszavak: random regresszió, variancia komponensek, öröklődhetőség

SUMMARY

One of the most important part of the genetic evaluation using a random regression model is the estimation of variance components. This is the topic of many papers because the large computational costs. We can use restricted maximum likelihood (REML), Gibbs sampling and \mathcal{R} method for the estimation of genetic parameters. The variance components are necessary to calculate the heritabilities and repeatabilities.

The aim of our paper is to estimate the variance components using a random regression repeatability model from test day data set of Hungarian Holstein-Friesian dairy cows and to analyse the change of additive genetic and permanent environmental variance, heritability and repeatability over lactation.

Keywords: random regression, variance components, heritability

BEVEZETÉS

Napjainkban a regressziós modelleknek egyre fontosabb szerep jut számos állatfaj egyedének tenyésztérbecslésében. Ezek a modellek, befejeési nap (BN) modell néven, a tejhasznosítású szarvasmarhák tenyésztérbecslésében a legelterjedtebbek (Jensen, 2001; Szyda és Liu, 1999). Az INTERBULL számos tagországában használják a holstein-fríz és más tejhasznosítású szarvasmarhák tenyésztérbecslésében a BN modellt.

A BN modellek a 305 napos laktációs termelésen alapuló modellekkel szemben pontosabb becslést eredményeznek, mert beépíthetők azok a hatások, amelyek a laktáció egy részében befolyásolják a tejtermelést. A BN modellek teljes- és részlaktációs tenyésztérbecslésére is alkalmasak, sőt meghatározható velük a két tenyésztérbecslés közötti korreláció is. A BN modellek további előnye, hogy alkalmasak a perzisztencia kiszámítására is. Hátrányként jelentkezik rendkívüli számológépigényük, így megvalósításukhoz nagy kapacitású számítástechnikai háttér szükséges.

A BN modelleknek két fő változata ismert: a fix regressziós- és a random regressziós modell. A fix regresszióval a fix hatású tényezőket regresszióval közelítjük, így kezelési csoportonként különböző alakú laktációs görbéket kapunk. Egy adott egyedhez tartozó laktációs görbe azonban alakjában nem tér el az azonos kezelési csoportba tartozó egyedekétől. Random regresszióval viszont minden egyed laktációjára egyedi laktációs görbe illeszthető. Ezzel lehetőség nyílik a laktációs görbe szerinti szelekcióra.

A random regressziós modellekkel végzett tenyésztérbecslés egyik legfontosabb feltétele a variancia komponensek és a genetikai paraméterek becslése. A variancia komponensek becslésének három leggyakrabban használt módja a csökkentett maximum likelihood módszer (restricted maximum likelihood, REML), a Gibbs-féle mintavételezés (Gibbs sampling) és az \mathcal{R} módszer (\mathcal{R} method). A REML módszer nagy előnye, hogy a kapott eredmények mindig a paraméterterbe esnek, hátránya viszont a nagyméretű problémáknál jelentkező erőforrásigény, valamint hogy programozási szempontból viszonylag nehéz megvalósítani. A Gibbs-féle mintavételezés előnye a könnyű megvalósíthatóság programozási szempontból, de a lassú konvergencia miatt rendkívül nagy futásidőre kell számítani. Az \mathcal{R} módszer nagyméretű modellekre is jól alkalmazható, de viszonylagosan nagy hibát eredményezhet (Misztal, 2000). Meyer (1997), valamint Meyer és Hill (1997) közleményeikben a REML elméleti háttérét, Jamrozik és mtsai (1997) a Gibbs mintavételezést, Druet és mtsai (2001), valamint Reverter és mtsai (1994) pedig az \mathcal{R} módszert használták a variancia komponensek becslésére.

Vizsgálataink célja a holstein-fríz fajta tejtermelésével kapcsolatos variancia komponensek, valamint öröklődhetőségi és ismételhetőségi értékek becslése REML módszerrel a befejési nap adatok alapján random regressziós modellel, továbbá szeretnénk ismertetni az öröklődhetőségi érték változását a laktáció során.

ANYAG ÉS MÓDSZER

Kutatásaink során a Mezőgazdasági Szakigazgatási Hivatal Központ (MgSzH) és jogelődje, az Országos Mezőgazdasági Minősítő Intézet (OMMI) által működtetett központi adatbázis adatait használtuk. A vizsgálat alapjául 62 511, 1999 és 2005 között ellett első, második és harmadik laktációs holstein-fríz fajtájú egyed befejési adatai szolgáltak. Ezek megfeleltek a következő szűrési feltételeknek: az egyedek legalább 75%-os holstein-fríz génhányaddal rendelkeztek, első ellési életkoruk legalább 21 hónap volt, az adott laktációban az elléstől a sikeres termékenyítésig eltelt napok, azaz nyitott napok (days open) száma, valamint - az első laktációt követően - az előző laktációban megfigyelt nyitott napok száma 21 és 240 nap közé esett. Ezen felül csak olyan egyedek adatait vettük figyelembe, amelyek teljes laktációt teljesítettek, mind a négy nagyszülejük ismert, s volt legalább négy, ugyancsak teljes laktációs adattal rendelkező lánytestvérük, valamint a tenyészetben az adott befejési napon legalább öt másik egyed is fejtek.

A variancia komponenseket REML módszerrel becsültük. A számítások elvégzéséhez a REMLF90 szoftvert (Misztal, 2002) használtuk. A Misztal (2000), valamint Meyer és Hill (1997) által leírt számolási módszer alapegyenlete a következő formában írható:

$$y_{ij}(t) = F + \sum_{n=1}^{n_p} \alpha_n x_{jn}(t) + \sum_{n=1}^{n_p} \gamma_n x_{jn}(t) + e_{ij}(t)$$

Az egyenletben y_{ij} a j-edik tehén napi termelése, amely tehén az i-edik populációhoz tartozik, F jelöli a fix hatásokat (telep×befejési nap, a laktáció sorszama Wilmlink függvényel modellezve), x_{jn} a laktációs görbe n-edik görbe paramétere, t a tehén adott laktációjában eltöltött napjainak számát jelöli (days in milk, DIM), α_n az additív genetikai hatás n-edik regressziós együtthatója, γ_n az állandó környezeti hatás n-edik regressziós együtthatója, e_{ij} a véletlen hiba. A random hatások közelítésére szintén a Wilmlink függvényt használtuk.

EREDMÉNYEK

A befejési napokon becsült additív genetikai variancia értéke 17,41 kg² és 25,89 kg² között változik. Az 1. ábrán látható, hogy a laktáció elejétől fokozatosan csökken a laktációs csúcs eléréséig. A laktáció csúcsa körül magasabb értéket vesz fel. A laktáció utolsó szakaszában jelentősen emelkedik.

Az állandó környezeti variancia értéke 5,19 kg² és 17,17 kg² között változik. A laktáció első

harmadában közel állandó, majd a laktáció második harmadától fokozatosan emelkedik (1. ábra).

1. ábra: Az additív genetikai- (a) és az állandó környezeti variancia (b) változása a laktáció során

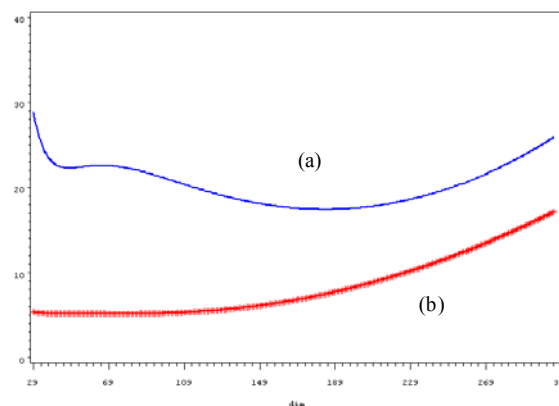


Figure 1: The change of additive genetic (a) and permanent environmental variance (b) over the lactation

Az öröklődhetőségi érték (h^2) a laktáció első szakaszában közel állandó, 0,15 érték körül mozog, csak a laktációs csúcs környékén emelkedik (0,16). A laktáció második felében fokozatosan emelkedik 0,18 értékig (2. ábra).

Az ismételhetőségi érték a h^2 -értékhez hasonló trendet követ 0,18 és 0,26 értékekkel (2. ábra).

2. ábra: Az öröklődhetőség (a) és az ismételhetőség (b) változása a laktáció során

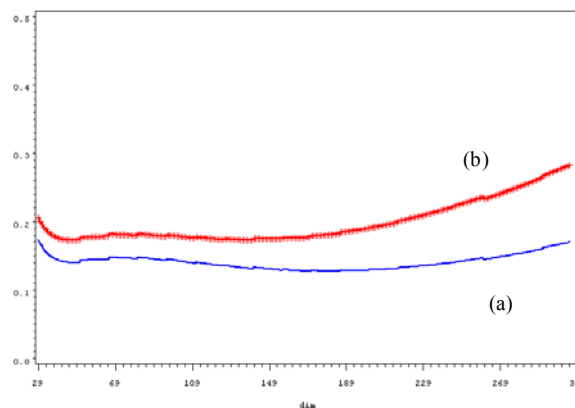


Figure 2: The change of heritability (a) and repeatability (b) over the lactation

MEGBESZÉLÉS

Az additív genetikai és állandó környezeti variancia változása a laktáció során a nemzetközi szakirodalomban leírtaknak megfelelő trendet követ (Olori és mtsai, 1999; van der Werf és mtsai, 1998).

Az öröklődhetőségi értékek az Olori és mtsai (1999) által leírtakhoz hasonló módon változnak a laktáció során, bár az értékek végig alacsonyabb szinten maradnak.

A laktáció első napjaiban, laktációs csúcstól körüli napokon, valamint a laktáció utolsó szakaszában megfigyelhető additív genetikai variancia, öröklődhetőségi érték és ismételhetőségi érték magasabb szintje azt mutatja, hogy ezekben a szakaszokban mért teljesítmény jól öröklődik, erre

végzett szelekció eredményesebb lehet, mint a laktáció más szakaszaiban mutatott teljesítményre végzett szelekció. A random regressziós modellekkel végzett becslések eszközt szolgáltatnak ezen tulajdonságok további fejlesztésére.

IRODALOM

- Druet, T.-Misztal, I.-Duangjinda, M.-Reverter, A.-Gengler, N. (2001): Estimation of genetic covariances with Method \mathcal{R} . *Journal of Animal Science* 79, 605-615.
- Jamrozik, J.-Kistemaker, G. J.-Dekkers, J. C. M.-Schaeffer, L. R. (1997): Comparison of Possible Covariates for Use in a Random Regression Model for Analyses of Test Day Yields. *Journal of Dairy Science* 80, 2550-2556.
- Jensen, J. (2001): Genetic Evaluation of Dairy Cattle Using Test-Day Models. *J. Dairy Sci.* 84, 2803-2812.
- Meyer, K. (1997): An "average information" Restricted Maximum Likelihood algorithm for estimating reduced rank genetic covariance matrices or covariance functions for animal models with equal design matrices. *Genetic Selection and Evolution* 29, 97-116.
- Meyer, K.-Hill, W. G. (1997): Estimation of genetic and phenotypic covariance functions for longitudinal or 'repeated' records by Restricted Maximum Likelihood. *Livestock Production Science* 47, 185-200.
- Misztal, I. (2000): *Computational Techniques in Animal Breeding*. University of Georgia, Athens.
- Misztal, I. (2002): *REMLf90 Manual*. [Online] www.nce.ads.uga.edu/~ignacy/numpub/blupf90/docs/remlf90.pdf
- Olori, V. E.-Hill, W. G.-Mc Guirk, B. J.-Brotherstone, S. (1999): Estimating variance components for test day milk records by restricted maximum likelihood with a random regression animal model. *Livestock Production Science* 61, 53-63.
- Reverter, A.-Golden, B. L.-Bourdon, R. M.-Brinks, J. S. (1994): Method \mathcal{R} Variance Components Procedure: Application on the Simple Breeding Value Model. *Journal of Animal Science* 72, 2247-2253.
- Szyda, J.-Liu, Z. (1999): Modelling test day data from dairy cattle. *J. Appl. Genet.*, 40, 103-116.
- Van der Werf, J. H. J.-Goddard, M. E.-Meyer, K. (1998): The Use of Covariance Functions and Random Regressions for Genetic Evaluation of Milk Production Based on Test Day Records. *Journal of Dairy Science* 81, 3300-3308.