



HAL
open science

Genetic heteroscedastic models for ordinal traits: application to sheep litter size

Samira Fathallah, Loys Bodin, Ingrid David

► To cite this version:

Samira Fathallah, Loys Bodin, Ingrid David. Genetic heteroscedastic models for ordinal traits: application to sheep litter size. *Genetics Selection Evolution*, 2016, 48 (1), pp.30. 10.1186/s12711-016-0202-4 . hal-01341364

HAL Id: hal-01341364

<https://hal.science/hal-01341364>

Submitted on 4 Jul 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

RESEARCH ARTICLE

Open Access



Genetic heteroscedastic models for ordinal traits: application to sheep litter size

Samira Fathallah, Loys Bodin* and Ingrid David

Abstract

Background: Classical genetic canalization models, which accommodate the mean and variance of a trait separately, provide a flexible approach to take heteroscedasticity for continuous traits into account. However, this model is not appropriate for discrete traits. The aim of this work was to propose heteroscedastic threshold models suitable for the genetic analysis of ordinal data.

Methods: In order to first fit the mean and variance of ordinal traits separately, we extended the classical threshold model (TM) for discrete data by introducing non-genetic and genetic factors of heterogeneity on the variance of its underlying variable, which leads to a homothetic threshold model HTM and its alternative parameterization HTM' in which the thresholds of different individuals are linked by a homothetic-translation. Relaxing the constraint between the thresholds led us to propose an independent threshold model ITM that was more flexible than HTM' but required the estimation of more parameters. TM, HTM and ITM models were applied to study 19,671 records on litter size in Romane sheep.

Results: Both HTM and ITM were able to disentangle the link between the mean and variance that holds in the classical homoscedastic threshold model. The results obtained for the litter size of Romane ewes showed that the data was best fitted with HTM compared to ITM and TM. The correlations between the observed and predicted variances were equal to 0.6 and 0.2 for HTM and TM, respectively. These analyses showed the existence of a genetic component for the heterogeneity of litter size in sheep that was taken into account in HTM.

Conclusions: HTM is the most suitable model to study the variability of litter size in sheep. It accommodates both the mean and variance separately while requiring the estimation of only a few parameters.

Background

Robustness can be viewed as the ability to maintain a stable phenotype regardless of the environmental conditions. In line with this approach, Scheiner and Lyman [1] proposed a model which considers that the expression of a trait is controlled by two sets of genes. One set controls the level of the performance and the other set controls the environmental variability of the trait. Various authors developed heteroscedastic models (i.e. canalization models) to jointly estimate the influence of various factors on both the mean and the variability of a trait [2–5].

These models were implemented with frequentist [6, 7] or Bayesian methods [8, 9], and applied to a large variety

of continuous traits (birth weight of piglets, birth weight of rabbits, weight of chickens, etc.).

The methods described above mainly target continuous traits. Theoretically, these linear models do not analyze discrete traits adequately because of their non-normal distributions. However, some previous empirical studies showed that the performance of linear and non-linear models was similar in terms of goodness-of-fit and predictive ability [10, 11] when analyzing the level of discrete livestock data. However, it is not known whether these properties are still true when dealing with the variability of discrete data.

The main objective of this research was to propose new models that take the heterogeneous individual variability of ordinal traits into account using derived threshold models. The models are described consecutively in the “Methods” section. “Heteroscedastic extensions of the

*Correspondence: loys.bodin@toulouse.inra.fr
GenPhySE, Université de Toulouse, INRA, INPT, INP-ENVT, Castanet-Tolosan, France

threshold model” are presented in the first subsection, and a “Numerical application” of these models to litter size in Romane sheep is presented in the second subsection.

Methods

Heteroscedastic models for ordinal traits

Homothetic threshold models for ordinal traits

A simple way of studying environmental sensitivity of an ordinal trait is to combine the threshold model (TM) [12] (appropriate for the study of ordinal traits) and the canalization model (appropriate for continuous traits) [13]. Thus, heteroscedasticity is considered under the liability threshold model and we call this new model the homothetic threshold model or HTM:

$$\mathbf{t} = (0, 1, \tau_3, \dots, \tau_{n-1}); \quad \mathbf{l}|\mathbf{u}, \mathbf{v} \sim N(1\mu + \mathbf{u}, f^2(\eta, \mathbf{v})),$$

where \mathbf{l} is the normally distributed liability and \mathbf{t} is a set of thresholds, which transform this liability into the observed discrete variable y in n categories, μ is the overall mean, \mathbf{u} the vector of additive genetic effects that affect the mean of the liability, η is the overall mean and \mathbf{v} is the vector of additive genetic effects that affect the variability of the liability. Several functions f have been proposed in the literature (exponential, linear or square root functions [3]). The vectors of genetic values are assumed to be Gaussian:

$$\mathbf{G}' = \begin{bmatrix} \sigma_{u'}^2 & 0 & \sigma_{u',v'}(\mu'_2 - \mu'_1) & \sigma_{u',v'}(\mu'_3 - \mu'_1) & \sigma_{u',v'}(\mu'_{n-1} - \mu'_1) \\ 0 & 0 & 0 & 0 & 0 \\ & (\mu'_2 - \mu'_1)^2 \sigma_{v'}^2 & (\mu'_2 - \mu'_1)(\mu'_3 - \mu'_1) \sigma_{v'}^2 & (\mu'_2 - \mu'_1)(\mu'_{n-1} - \mu'_1) \sigma_{v'}^2 & (\mu'_2 - \mu'_1)(\mu'_{n-1} - \mu'_1) \sigma_{v'}^2 \\ & & (\mu'_3 - \mu'_1)^2 \sigma_{v'}^2 & (\mu'_3 - \mu'_1)(\mu'_{n-1} - \mu'_1) \sigma_{v'}^2 & (\mu'_3 - \mu'_1)(\mu'_{n-1} - \mu'_1) \sigma_{v'}^2 \\ & & & (\mu'_{n-1} - \mu'_1)^2 \sigma_{v'}^2 & (\mu'_{n-1} - \mu'_1)^2 \sigma_{v'}^2 \end{bmatrix}$$

$$\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \mathbf{G} \otimes \mathbf{A}\right),$$

where \otimes denotes the Kronecker product, \mathbf{A} is the known additive genetic relationship matrix and \mathbf{G} the genetic matrix $\begin{bmatrix} \sigma_u^2 & \rho\sigma_u\sigma_v \\ \rho\sigma_u\sigma_v & \sigma_v^2 \end{bmatrix}$, where σ_u^2 and σ_v^2 are the variances of the additive genetic effects that affect the mean and the variability of the liability, and ρ their coefficient of correlation.

A fully equivalent parameterization of HTM can be written. This parameterization or HTM' model is reached by fixing the parameters of the liability and considering that the matrix of thresholds \mathbf{T}' is affected by environmental and genetic effects i.e.:

$$\begin{aligned} \mathbf{T}' &= \begin{pmatrix} \tau'_{11} & \tau'_{21} & \dots & \tau'_{(n-1)1} \\ \tau'_{1i} & \tau'_{2i} & \dots & \tau'_{(n-1)i} \\ \vdots & \vdots & & \vdots \\ \tau'_{1N} & \tau'_{2N} & \dots & \tau'_{(n-1)N} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{t}'_1 = \mu' + u'_1 \mathbf{1}^T + \mathbf{v}'_1 \\ \mathbf{t}'_i = \mu' + u'_i \mathbf{1}^T + \mathbf{v}'_i \\ \vdots \\ \mathbf{t}'_N = \mu' + u'_N \mathbf{1}^T + \mathbf{v}'_N \end{pmatrix} \\ &= (\tau'_1, \tau'_2, \tau'_3, \dots, \tau'_{(n-1)N}); \quad \mathbf{l} \sim N(0, 1). \end{aligned}$$

Under this parameterization, the vector of thresholds \mathbf{t}'_i is specific to each individual i , with $\mathbf{t}'_i = \mu' + u'_i \mathbf{1}^T + \mathbf{v}'_i$ where μ' is an overall set of $n - 1$ thresholds, u'_i a subject-specific deviation from this overall set and \mathbf{v}'_i a vector of size $n - 1$ that reflects the specific deviation of the subject threshold from this overall set with:

$$\begin{bmatrix} u' \\ \mathbf{v}'_1 \\ \mathbf{v}'_2 \\ \mathbf{v}'_3 \\ \vdots \\ \mathbf{v}'_{n-1} \end{bmatrix} \sim N(0, \mathbf{G}' \otimes \mathbf{A}),$$

where

Under this parameterization, the correlation between τ'_{k_1} and τ'_{k_2} is set to 1 and the transformation of \mathbf{t}'_i to $\mathbf{t}'_{i'}$ (from individuals i to i') is not a translation as in TM but a homothetic translation, which means that the ratio $R = (\tau'_{k_3i} - \tau'_{k_1i}) / (\tau'_{k_2i} - \tau'_{k_1i})$ is constant for the whole population i.e.:

$$\begin{aligned} &(\tau'_{k_3i} - \tau'_{k_1i}) / (\tau'_{k_2i} - \tau'_{k_1i}) \\ &= (\tau'_{k_3i'} - \tau'_{k_1i'}) / (\tau'_{k_2i'} - \tau'_{k_1i'}), \quad \forall k_1, k_2, k_3 \\ &= (\mu'_{k_3} - \mu'_{k_1}) / (\mu'_{k_2} - \mu'_{k_1}). \end{aligned}$$

HTM and HTM' are equivalent models with the same number of parameters ($n + 1$) being estimated. The relationship between \mathbf{t}'_i and \mathbf{t} is:

$$\mathbf{t}'_i = \frac{\mathbf{t} - (\mu + u_i)\mathbf{1}^T}{f(\eta, v_i)}$$

The estimations obtained with HTM' can be transformed into the estimations that would have been obtained with HTM. Overall, in the case of a linear function f , the approach produces:

$$\sigma_v^2 = \frac{\sigma_v'^2}{(\mu'_2 - \mu'_1)^2},$$

whereas for the other parameter we have:

$$\sigma_u^2 = \frac{1}{(\mu'_2 - \mu'_1)^2} \times \left((\mu'_1)^2 \sigma_v'^2 + \sigma_u'^2 + \sigma_v'^2 \sigma_u'^2 + (\sigma_{v',u'})^2 - 2\mu'_1 \sigma_{v',u'} \right).$$

Finally, the genetic correlation ρ is:

$$\rho = \frac{1}{(\mu'_2 - \mu'_1)^2} \left(\frac{\sigma_{v',u'} - \mu'_1 \sigma_v'^2}{\sigma_u \sigma_v} \right).$$

These formulas are detailed in "Appendix 1". The heritability of the trait on the liability scale is estimated by the ratio of the animal variance that acts on the mean to the total variance of the underlying variable i.e.:

$$h_1^2 = \frac{\sigma_u^2}{\sigma_u^2 + f^2(\eta, \sigma_v)}$$

For instance in the case of a linear function f :

$$h_1^2 = \frac{\sigma_u^2}{\sigma_u^2 + \eta^2 + \sigma_v^2}$$

Extension of the homothetic threshold model into the independent threshold model

A straightforward extension of the above HTM' model can also be considered and consists in releasing the constraint of the homothetic-translation link between the thresholds of the individuals. The model therefore becomes an independent threshold model ITM, i.e.:

$$\mathbf{T} = \begin{pmatrix} \tau_{11} & \tau_{21} & \dots & \tau_{(n-1)1} \\ \tau_{1i} & \tau_{2i} & \dots & \tau_{(n-1)i} \\ \vdots & \vdots & \ddots & \vdots \\ \tau_{1N} & \tau_{2N} & \dots & \tau_{(n-1)N} \end{pmatrix} = \begin{pmatrix} \mathbf{t}_1 = \mu + \mathbf{v}_1 \\ \mathbf{t}_i = \mu + \mathbf{v}_i \\ \vdots \\ \mathbf{t}_N = \mu + \mathbf{v}_N \end{pmatrix} \\ = (\boldsymbol{\tau}_1, \boldsymbol{\tau}_2, \boldsymbol{\tau}_3, \dots, \boldsymbol{\tau}_{(n-1)}); \quad \mathbf{l} \sim N(0, 1),$$

where $\mathbf{t}_i = \mu + \mathbf{v}_i$ with μ being an overall set of $n - 1$ thresholds, and \mathbf{v}_i a vector of size $n - 1$, which reflects the specific deviation of the subject threshold from this overall set with:

$$\begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_j \\ \vdots \\ \mathbf{v}_{n-1} \end{bmatrix} \sim N \left[0, \begin{bmatrix} \sigma_{v_1}^2 & \sigma_{v_1 v_j} & \dots & \sigma_{v_1 v_{n-1}} \\ & \sigma_{v_j}^2 & \dots & \sigma_{v_j v_{n-1}} \\ & & \ddots & \vdots \\ & & & \sigma_{v_{n-1}}^2 \end{bmatrix} \otimes \mathbf{A} \right].$$

The total number of parameters in this model is $(n + 2)(n - 1)/2$. It should be noted that ITM cannot be parameterized like HTM using a fixed set of thresholds and a model on the liability l . In this model, fixing the thresholds to modify the parameterization would imply that the underlying variable would no longer be Gaussian.

Due to the lack of constraints between the thresholds (except for the obvious relationship $-\infty < \tau_1 < \tau_2 < \dots < \tau_{n-1} < \infty$), ITM fits a much larger range of situations than a homothetic threshold model, which assumes a strong constraint either on the homothetic translation relationship between thresholds (HTM') or on the normality of the liability (HTM).

Interestingly, we found that another parameterization of ITM is a constrained multiple-trait model MTM (Appendix 2). Indeed, each observation of the ordinal trait can be considered as the result of n exclusive binary traits:

$$1_k(y) := \begin{cases} 1 & \text{if } y = k/k = 1 \dots n \\ 0 & \text{otherwise} \end{cases}$$

The observed variable y can thus be transformed into n indicator functions which can be analyzed using a multivariate binary threshold model. Only $n - 1$ indicator functions need to be considered in this MTM since the value of one of the indicator functions is fixed conditionally to the others. Each sub-model $k, k \in [1, n - 1]$ can be written as MTM_k:

$$\tau_k | u_k; \quad l_k \sim N(0, 1).$$

According to this model, the probability that an individual might achieve the observation k or that the k^{th} indicator function is equal to 1 is:

$$\pi_k = P(1_k(y) = 1) = P(l_k \leq \tau_k | u_k),$$

and the $n - 1$ probabilities π_k should satisfy the relation $\sum_{k=1}^{n-1} \pi_k < 1$. This constraint must be added to the multivariate binary threshold model.

As for ITM, there are $(n + 2)(n - 1)/2$ parameters to estimate i.e. $n - 1$ thresholds and $n(n - 1)/2$ variance covariance parameters.

The relationship between the thresholds of MTM and ITM is as follows:

$$\tau_{ik}^{\text{MTM}} = \Phi^{-1} \left(\Phi \left(\tau_{ik}^{\text{ITM}} \right) - \Phi \left(\tau_{i(k-1)}^{\text{ITM}} \right) \right).$$

Under the latter models, one cannot really consider that the expression of the ordinal trait is controlled by two sets of genes. The combination of elementary genetic values for each individual results in the heteroscedasticity of the global trait. In this model, the heritability can be calculated for each class as:

$$h_k^2 = \frac{\sigma_{u_k}^2}{\sigma_{u_k}^2 + 1}.$$

The different parameterizations of the models are summarized in Table 1. HTM' and ITM can be fitted using the ASReml software.

Numerical application

Data

These models were applied to study litter size in Romane sheep. All data (litter size, environmental variation factors and pedigrees over ten generations) were extracted from the national database for genetic evaluation and research managed by the Institut de l'Élevage (French Livestock Institute) and the Centre de Traitement de l'Information Génétique (Genetic Information Processing Center, Jouy-en-Josas, France). Data consisted of records on litter size of 11,073 multiparous ewes from 23 flocks with at least four lambings from 2000 to 2013. The mean prolificacy (2.13) in this sample was high and consequently there was a high global variability in litter size.

The main features of this dataset are in Table 2. To avoid extreme case problems [14], litter sizes of more than four lambs were pooled into a single class for all analyses, which resulted in four classes of litter size i.e. [1, 2, 3, 4+].

The fixed effects included in the analysis were selected to be as similar as possible to those found in others studies [15, 16]. This was done by comparing the likelihood-ratio of all the possible linear mixed models which included the following fixed effects: age of ewe (nine levels), time interval between lambing in months (seven levels), parity (seven levels), age of the ewe at first lambing in months (nine levels), season of lambing (two levels), as well as the random effects: flock and ewe effect. For this step of the analysis, the ewe effect was considered as random but the genetic relationships between individuals were ignored.

Estimation of parameters and comparison of methods

The three threshold models TM, HTM' and ITM described above were applied to this data. The classical TM, used as a basis for comparison, included the previously selected fixed effects and the ewe as both a random genetic effect (u^*) and a random permanent environmental effect (p^*) (i.e. $\mathbf{t}_i^* = \mathbf{x}_i\boldsymbol{\beta} + u_i^*\mathbf{1}^T + p_i^*\mathbf{1}^T$). The homothetic threshold model HTM' considered that fixed, genetic and permanent effects were common for all thresholds and threshold-specific random genetic effects with constrained genetic (co)variance to ensure

Table 1 Main characteristics of the different models

Models	Number of parameters	Thresholds	Underlying
Population			
<i>Homoscedastic</i>			
TM	n	$\mathbf{t}^* = \{\mathbf{t}_1^*, \mathbf{t}_2^*, \mathbf{t}_3^*, \dots, \mathbf{t}_{n-1}^*\}$	$Y \sim N(0, 1)$
<i>Heteroscedastic</i>			
HTM	$n + 2$	$\mathbf{t} = \{0, 1, \mathbf{t}_3, \dots, \mathbf{t}_{n-1}\}$	$Y \sim N(\mu + u, f^2(\eta, \mathbf{v}))$
HTM'	$n + 2$	$\mathbf{t}' = \{\mathbf{t}'_1, \mathbf{t}'_2, \mathbf{t}'_3, \dots, \mathbf{t}'_{n-1}\}$	$Y \sim N(0, 1)$
ITM	$(n + 2)(n - 1)/2$	$\mathbf{t}'' = \{\mathbf{t}''_1, \mathbf{t}''_2, \mathbf{t}''_3, \dots, \mathbf{t}''_{n-1}\}$	$Y \sim N(0, 1)$
Animal i			
<i>Homoscedastic</i>			
TM	n	$\mathbf{t}_i^* = \{\mathbf{t}_{1i}^*, \mathbf{t}_{2i}^*, \mathbf{t}_{3i}^*, \dots, \mathbf{t}_{(n-1)i}^*\}$ $\mathbf{t}_j^* = \mathbf{t}_i + \mu_i + u_i$	$y_i = \varepsilon_i$
<i>Heteroscedastic</i>			
HTM	$n + 2$	$\mathbf{t}_i = \{0, 1, \mathbf{t}_{3i}, \dots, \mathbf{t}_{(n-1)i}\}$	$y_i = \mu_i + u_i + \varepsilon_i$
HTM'	$n + 2$	$\mathbf{t}'_i = \{\mathbf{t}'_{1i}, \mathbf{t}'_{2i}, \mathbf{t}'_{3i}, \dots, \mathbf{t}'_{(n-1)i}\}$ $\mathbf{t}'_j = \mathbf{t}'_i + \mu'_i + u'_i + v'_i$	$y_i = \varepsilon_i$
ITM	$(n + 2)(n - 1)/2$	$\mathbf{t}''_i = \{\mathbf{t}''_{1i}, \mathbf{t}''_{2i}, \mathbf{t}''_{3i}, \dots, \mathbf{t}''_{(n-1)i}\}$ $\mathbf{t}''_j = \mathbf{t}''_i + \mu''_i + u''_i$	$y_i = \varepsilon_i$

Table 2 Features of Romane sheep

Item	
Number of records	19,671
Number of animals	11,073
Number of sires	1096
Number of dams	6550
Mean	2.13
<i>Distribution of litter size (%)</i>	
Single	17.70
Twin	55.01
Triplet	24.18
Quadruplet and +	3.11

homothetic translation (i.e. $\mathbf{t}'_i = \mathbf{x}_i\boldsymbol{\beta} + u'_i\mathbf{1}^T + p'_i\mathbf{1}^T + \mathbf{v}'_i$). Finally, an additional extension resulting in the independent threshold model ITM was fitted and it considered that the fixed and permanent effects were common for all thresholds and included a genetic effect that was specific to each threshold with no constrained genetic (co)variance (i.e. $\mathbf{t}''_i = \mathbf{x}_i\boldsymbol{\beta} + p''_i\mathbf{1}^T + \mathbf{v}''_i$).

ASReml software [17] was used to implement all the models.

Models were compared using the likelihood ratio test for nested models (TM vs. HTM', LRT df = 2; TM vs. ITM, LRT df = 5) and the Bayesian Information Criterion (BIC) for non-nested models (BIC = $\ln(N - p) * c - 2\ln L$, HTM' vs. ITM, where L is the restricted maximum likelihood of the model, N the number of observations, p and c the number of fixed effects and covariance parameters) [18]. The models were also compared based on their ability to predict new records as follows: 75 % of records were used to estimate parameters that were used to predict the remaining 25 %. Five replicates of this design were sampled at random. To evaluate the predictive ability of the models, MSEP (mean square error of prediction) statistic was computed as:

$$MSEP = \frac{1}{M} \sum_{j=1}^M \left(\frac{1}{m} \sum_{i=1}^m (y_{ij} - \hat{y}_{ij})^2 \right),$$

where y and \hat{y} correspond to the observed and predicted litter size, m is the number of data in the validation subset and M is the subset number.

For ewes with observations, we compared the raw means and variances of their litter sizes to the corresponding predicted values on the observable scale obtained with the different models. The raw mean and variance were calculated for each individual from the observed data. The predicted values were calculated as:

$$\hat{y}_i = \sum_{k=1}^4 k\hat{\pi}_{i,k},$$

and

$$\hat{\sigma}_i^2 = \sum_{k=1}^4 k^2\hat{\pi}_{i,k} - \left(\sum_{k=1}^4 k\hat{\pi}_{i,k} \right)^2,$$

where $\hat{\pi}_{i,k}$ was the probability that the litter size of a ewe i was k . These probabilities were calculated using $\hat{\pi}_{i,k} = \Phi(\hat{\tau}_{i,k}) - \Phi(\hat{\tau}_{i,k-1})$ with as above $\tau_0 = -\infty$ and $\tau_4 = +\infty$ ($\hat{\tau}_{i,k}$ corresponds to $\hat{\tau}_{i,k}^*$, $\hat{\tau}'_{i,k}$ or $\hat{\tau}''_{i,k}$ depending on the model). To measure the ability of the models to break the link between $\hat{y}_i = t$ and $\hat{\sigma}_i^2 = z$, we used the following criterion:

$$B = \frac{E(Var(z|t))}{Var(z)},$$

which ranges from 0 to 1. The criterion B was 0 if $E(Var(z|t)) = 0$, which means that a single value of the mean has a single value of the variance. The criterion B was 1 if $E(Var(z|t)) = 1$, which means that a single value of the mean has different variance values. Because the density associated to the variance $Var(z|t)$ is unknown, data were ordered in ascending order of litter size and were discretized into s subgroups, and the variance of y in each subgroup $\hat{\sigma}_m^2$ was calculated in order to calculate:

$$E(Var(z|t)) = E\left(\hat{\sigma}_{y_i}^2 \middle| \hat{y}_i\right) = \sum_{m=1}^s \hat{\sigma}_m^2/s.$$

The subgroup number was defined in order to get a sufficient amount of data in each subgroup to obtain a good estimation of the subgroup variance. The intensity of the dispersion of each model was also assessed by calculating the correlation, either for the mean or the variance, between the observed values and the predicted values on the observable scale.

Results

Comparison of models

The values of the criteria (LRT, BIC, MSE and MSEP) used to compare the adjustment quality of the models are in Tables 3 and 4.

The results showed that the homoscedastic model should be rejected and that the models that account for genetic heterogeneity of litter size distribution should be favored. Among the heteroscedastic models, BIC values were lower with HTM' than with ITM. MSE estimates of the mean were similar for the three models, whereas MSE estimates of the variance differed and were higher with TM than with HTM' (0.13 with TM vs. 0.10 with HTM'). The differences between MSEP values are small, which is explained by the fact that this parameter can estimate prediction error for the level of the trait only and not for its variability. To evaluate the ability of the models to predict

Table 3 Values obtained from likelihood-ratio, BIC tests, B criteria, and mean squared error (MSE) between observed and fitted values in the different models

Models/tests	Likelihood-ratio (LRT)	BIC test	B	MSE of mean	MSE of variance
TM	TM versus HTM': >>11 (Reject H0 P value <0.001)	54,005	0.013	0.07	0.13
HTM'	TM versus ITM: >>16 (Reject H0 P value <0.001)	49,517	0.51	0.07	0.10
ITM		50,663	0.42	0.07	0.10

TM threshold model, HTM' homothetic threshold model, ITM independent threshold model, B B criteria

Table 4 Values obtained from mean squared error of prediction (MSEP) and B criteria on the validation subset in the different models

Models/tests	MSEP	B criteria
Threshold model (TM)	0.473	0.005
Homothetic threshold model (HTM')	0.471	0.363
Independent threshold model (ITM)	0.481	0.371

variance, we calculated the B criteria on the validation subset by discretizing the whole data into 200 adjacent subgroups. The results are in Table 4. Based on the B criteria, HTM' and ITM were more efficient than TM (0.36 for HTM', 0.37 for ITM vs. 0.005 for TM).

Homoscedastic threshold model used as reference (TM)

With TM, the genetic additive variance of the liability σ_u^2 was equal to 0.079 and the variance of the permanent environmental σ_p^2 effect was equal to 0.034, which resulted in a heritability h^2 of 0.07. The correlation between the predicted and observed means was very high ($r = 0.80$) (Table 5), but it was low between the predicted and observed variances ($r = 0.20$). In spite of the curvilinearity between the predicted means and variances, the correlation between these variables was high ($r = 0.99$) and close to 1, and the B parameter was equal to 0.013 (Table 3), which reveals the one-to-one relationship between these variables when this homoscedastic threshold model was used as shown on the scatterplot of Fig. 1a.

Table 5 Correlations between the predicted and observed means as well as the predicted observed variance with the different models

Model	Predicted mean versus observed mean	Predicted variance versus observed variance
TM	0.80	0.21
HTM'	0.79	0.61
ITM	0.78	0.59

TM threshold model, HTM' homothetic threshold model, ITM independent threshold model

Homothetic threshold model (HTM')

The genetic additive variances obtained with HTM' are in Table 6. By design, the correlation between the genetic values associated to thresholds (v) was equal to 1 in this model and the variance $\sigma_{v,1}^2$ was equal to 0. Moreover, and as expected due to the homothetic relationship, we observed an increase of $\sigma_{v,3}^2$ compared to $\sigma_{v,2}^2$. The variances were equal to 0.04, 0.05 and 0.20 for $\sigma_{u,1}^2$, $\sigma_{v,2}^2$, and $\sigma_{v,3}^2$, respectively. The variance of the permanent environmental effect was equal to 0.015.

By transforming the results of the HTM' parameterization to those of HTM, we could estimate the variance of the genetic effects that affect the mean of the liability as well as the correlation between these effects. The values found for the mean of the liability, the scaling function and the correlation between effects were $\sigma_u^2 = 0.069$, $\sigma_v^2 = 0.020$ and $\rho = -0.40$, respectively. The variance of the permanent effect was $\sigma_p^2 = 0.015$ whereas, by design, the mean of the residual variance was $\eta^2 = 1$. Since the model fitted for the environmental variance did not include fixed effects but only a general mean and a genetic effect, the heritability of the liability mean could be estimated independently of a given environment and was as follows:

$$h_l^2 = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_p^2 + \eta^2 + \sigma_v^2} = 0.062.$$

The correlation between observed and predicted means was high ($r = 0.79$) and not different from the value obtained with TM. This contrasts with the correlation between predicted and observed variances ($r = 0.61$). Although the prediction of the variances was still lower than that of the means, it was much higher than with TM. HTM' could therefore break the link between the predicted means and variances and the B parameter was equal to 0.51 (Table 3) even if the correlation between these variables was still high ($r = 0.88$). The scatterplot in Fig. 1b shows this variability of the predicted variance for a given predicted mean.

In this homothetic translation model, the gap Δ_{1-2} between the two first thresholds is an indirect measure of its variability for each individual. As an example, we

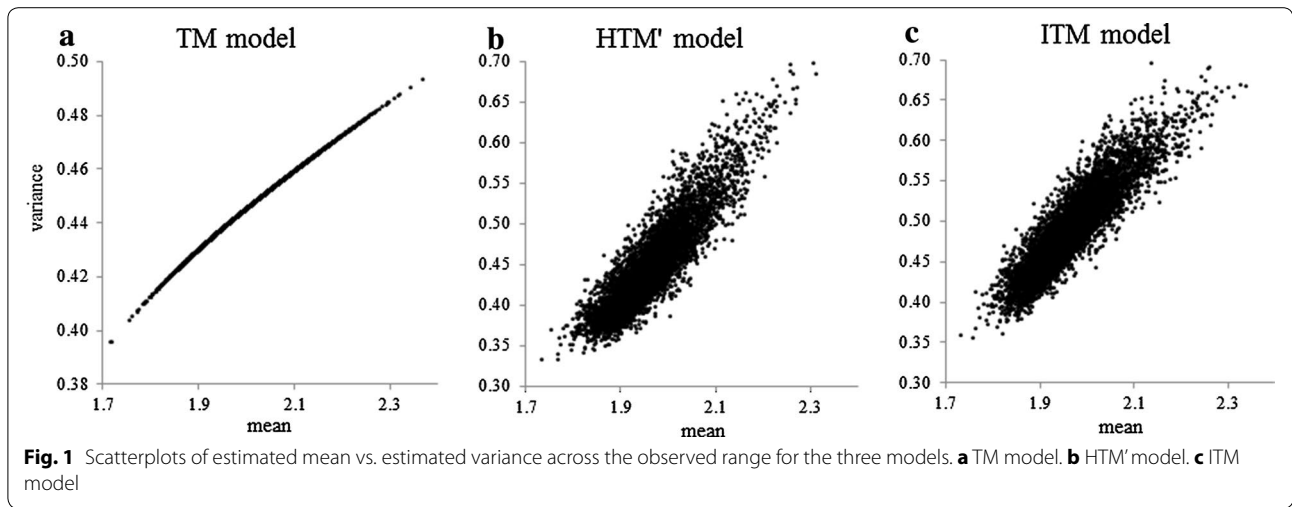


Table 6 Genetic variance and correlation between thresholds for Romane sheep using the HTM' model

Thresholds/genetic variance	$\sigma_{v',1}^2$	$\sigma_{v',2}^2$	$\sigma_{v',3}^2$
	0.04	0.13	0.13
		0.05	1
			0.20

estimated the distribution of the litter size of two individuals having the same average litter size but different variances (Table 7). For these individuals, the values of the mean and variance estimated with HTM' are close to the observed values of the mean and variance. The ratio is constant for the two individuals, and the difference in percentage of grand-offspring born as quadruplets or more can be large (0 vs. 6 %).

Independent threshold model (ITM)

The genetic variances and correlations obtained with ITM are in Table 8. Genetic variances were small, but

larger for the second and third thresholds than for the first threshold. Genetic correlations were moderate ($\rho > 0.4$) and it was highest between v''_2 and v''_3 . Using the transformation formulas described in "Appendix 2", it was possible to estimate the genetic variance associated with twin lambing ($\sigma_{ls=2}^2 = 0.14$), the corresponding heritability ($h_{ls=2}^2 = 0.128$) as well as the genetic correlation between the first and the second classes of litter size ($\rho = 0.44$). The correlations between the predicted and observed means ($r = 0.78$) and between the predicted and observed variances ($r = 0.59$) were close to those obtained with HTM (Table 8). The B parameter ($B = 0.42$, Table 3), the correlation between the predicted means and variances ($r = 0.91$) and the scatter plot (Fig. 1c) were also similar to those obtained with HTM'. The Spearman rank correlations between the thresholds of HTM' and ITM were high (Table 9).

Discussion

A genetic heteroscedastic model was previously proposed [6] to analyze categorical variables and break down

Table 7 Observed values (prolificacy, variance) and estimated values with HTM' [thresholds, different gap between the first thresholds (Δ_{1-2}), ratio (R), prolificacy, variance, and percentage of the simple (% LS 1), double (% LS 2), triplet (% LS 3) and quadruplet (% LS 4)] for two individuals

Observed values			Estimated values with HTM' model										
Ind	Prolificacy	Variance	τ'_1	τ'_2	τ'_3	Δ_{1-2}	R	Prolificacy	Variance	% LS1	% LS2	% LS3	% LS4+
1	2.00	0.50	-0.63	0.92	2.16	1.55	1.8	2.09	0.54	16	53	25	6
2	2.00	0.33	-0.89	0.86	2.28	1.75	1.8	2.02	0.38	21	64	15	0

Table 8 Genetic variance and correlation between thresholds for Romane sheep using the ITM model

Thresholds/genetic variance	$\sigma_{v''.1}^2$	$\sigma_{v''.2}^2$	$\sigma_{v''.3}^2$
	0.035	0.631	0.430
		0.138	0.921
			0.176

Table 9 Spearman rank correlations between the HTM' and ITM models

First threshold	0.988
Second threshold	0.989
Third threshold	0.988

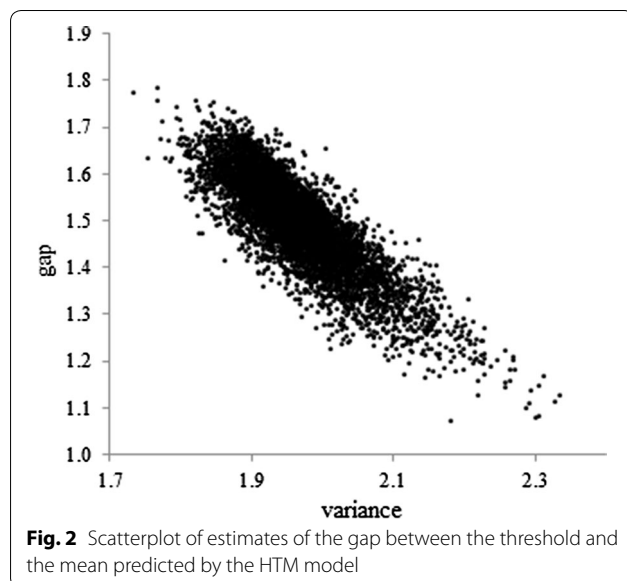
the link between mean and variance that holds with the classical homoscedastic threshold model. This model originated from the combination of two base models: the homoscedastic threshold model for categorical data [19] and the structural heteroscedastic model for continuous variables [2]. In this paper, we studied further extensions of this model by analyzing the relationship that exists between categories and in particular by releasing the main constraint.

In this paper, we propose two kinds of heteroscedastic models, both with different parameterizations. The first type (HTM/HTM') was fully derived from the canalization model proposed for continuous traits [2] with which it was easy to draw comparisons. The second type consisted in relaxing the relationship that exists between thresholds (translation in TM or homothetic translation

in HTM) and led to a model that considers each threshold independently i.e. ITM. We demonstrated that this model was equivalent to a constrained multiple-trait model MTM on each class of the ordinal trait. This second type of model is much more flexible than HTM and its application in selection would be of particular interest because it allows for each class to be controlled through a multiple selection index [20] with specific environmental effects and economic weights for each class. Modeling each class of the trait to control its distribution or modeling only the most important class are trivial issues and have already been considered [21], however the correct multiple trait model must include a constraint on the sum of the probabilities of the $n - 1$ binary variables. In spite of their potential interest, the use of this kind of model (ITM or MTM) for genetic evaluation and selection would be limited since they require a larger number of parameters to be estimated than HTM or HTM' and would probably show convergence problems when the number of classes exceeds three. Furthermore, their biological interpretation is difficult; for instance we could not link the parameters of ITM or MTM to the biological model of Schneider and Lyman [22], which assumes the existence of different genes that control the mean and the variance of a trait.

We compared TM, HTM' and ITM through the analysis of data on litter size from both Romane and Rouge de l'Ouest sheep (the latter results are not shown). The low heritability of the liability to litter size obtained with an homoscedastic threshold model (i.e. TM) agreed with previous studies that used TM [23–25]. However, LRT, BIC, and MSE parameters showed that this homoscedastic model was less successful than both heteroscedastic models. Consequently, the distribution of sheep litter size shows heteroscedasticity of genetic origin as revealed by both the homothetic HTM' and independent threshold ITM models.

Taking the genetic heteroscedasticity in HTM' and the subsequent transformation of the parameters to those of HTM into account did not substantially change the heritability of the liability to litter size compared to TM. However, with HTM, the additive genetic variance of the variability of the trait was smaller than the corresponding variance of its mean. This is in line with results in the literature that showed that when additive genetic variances of the mean and the variability of a trait were jointly estimated, their values were smaller for the variability than for the mean [6, 9, 26–31]. This seems to be a common rule regardless of the methodology used for such estimations, the trait or the species. In contrast, the genetic correlation between the mean of a trait and its variability is highly variable and covers the whole range of its variation interval. For instance, in a mouse population, genetic



correlations of -0.93 and 0.97 were reported for litter size and individual birth weight, respectively [29]. In our study, we found a negative genetic correlation between the mean and the variability of sheep litter size with HTM' ($\rho = -0.40$). As highlighted by Yang et al. [32], this parameter might be affected by an artefact due to the scale of measurement or skewness of the data, although with the threshold model no such skewness affected this correlation. However, a study on another French sheep breed (Rouge de l'Ouest) for which the mean prolificacy (prolificacy = 1.80; LS3+ = 12.5 %) is lower than in Romane sheep (results not shown) reported a similar value i.e. -0.25 .

Relaxing the homothetic constraint between the thresholds of HTM' did not improve the fit of the model or the predictive ability of the data as might be expected, but instead increased the BIC values due to the larger number of parameters to estimate in ITM. Moreover, although the constraint on the genetic correlation between thresholds was relaxed in ITM, the correlation between the second and third thresholds remained close to 1 and the correlation between the first and second thresholds was high; in addition, the Spearman rank correlations between the thresholds of HTM' and ITM were close to 1 (Table 9). This suggests that the true nature of the link between the thresholds is almost a homothetic translation in sheep and that the distributions of litter size classes are not independent but linked by a strong common law. It also means that the concept of sensitivity thresholds, which transform individual Gaussian variables that differ in mean and variance into the observed categories, might be relevant.

Nonetheless in our study, only genetic effects were considered as threshold-specific in ITM in order to compare the different models using LRT or BIC criteria. Of course, threshold-specific fixed effects and permanent effects could be included in ITM. In the same way, threshold-specific fixed effects and permanent effects could be included in HTM'. Unfortunately, HTM with a threshold-specific permanent effect did not converge and threshold-specific fixed factors with proportional effects cannot be implemented in ASReml. It may be possible to estimate fixed effects in the residual variance using an iterative process that tests the influence of each effect on the variance of the model. We did not test such algorithms on our data and it would probably be a slow and difficult process. Another solution would be to modify existing software for canalizing selection of a continuous trait in order to take into account the discrete characteristics of the data.

Using ITM, Martin et al. [33] showed that the polymorphism of a single gene affected not only the mean prolificacy in Lacaune sheep, but also modified its

variability by changing the gap between the thresholds in a homothetic way. A superiority of the homothetic model (HTM'), as well as a similar genetic variability of the liability to litter size were found for Rouge de l'Ouest sheep (results not shown), thus covering the whole range of prolificacy in breeds for which canalization might be of interest. By design, the gap between the first two thresholds estimated for each individual in HTM' is inversely proportional to the individual residual variance of HTM (Fig. 2). This criterion has been used to compare the variability of about 50 French sheep breeds [34].

Conclusions

This study shows that we can model the variabilities (mean and variance) of a discrete trait using a suitable model, namely the HTM' model. Thus, the canalization of discrete traits is possible and the gap between the first two thresholds estimated for each individual in this model could be used to estimate the breeding value for the variability of litter size independently from that of the mean.

Authors' contributions

SF, ID and LB jointly designed the study and discussed the results. All authors read and approved the final manuscript.

Acknowledgements

The authors would like to thank the CORAM for the funding and Helene Munteguy for linguistic revision and referees for their helpful contributions.

Competing interests

The authors declare that they have no competing interests.

Appendix 1

In this section we describe how we calculated the parameters σ_u^2 , σ_v^2 and ρ_{uv} of a HTM model based on the HTM' estimates. Let t'_k be the k^{th} threshold of the HTM' model with $\mathbf{t}'_i = \mu'_i + u'_i \mathbf{1}^T + \mathbf{v}'_i$, a simple variable transformation gives:

$$\begin{cases} t'_2 = \frac{t_2 - (\mu + u)}{f(\eta + v)} \\ t'_1 = \frac{t_1 - (\mu + u)}{f(\eta + v)} \end{cases}$$

Given that $t_1 = 0$ and $t_2 = 1$ in HTM and $t'_1 = \mu'_1 + u'$ and $t'_2 = \mu'_2 + u' + v'_2$ in HTM', we obtain:

$$\begin{cases} f(\eta, v) = \frac{1}{\mu'_2 - \mu'_1 + v'_2} \\ = \frac{1}{\mu'_2 - \mu'_1} \left(\frac{1}{1 + v'_2 / (\mu'_2 - \mu'_1)} \right) \\ u = -\mu - f(\eta, v) (\mu'_1 + u') \end{cases}$$

Considering that the ratio $v'_2 / (\mu'_2 - \mu'_1)$ is small in comparison to 1, we applied Taylor's expansion to the previous equation:

$$\Rightarrow \begin{cases} f(\eta, v) = \frac{1}{\mu'_2 - \mu'_1} - \frac{v'_2}{(\mu'_2 - \mu'_1)^2} \\ u = -\mu - \frac{1}{\mu'_2 - \mu'_1} \left(1 - \frac{v'_2}{\mu'_2 - \mu'_1}\right) (\mu'_1 + u') \end{cases}$$

Thus, In the case of a linear function f and given that $v'_2 = (\mu'_2 - \mu'_1)v'$, we obtain:

$$\Rightarrow \begin{cases} v = C_1 - \frac{v'}{\mu'_2 - \mu'_1} \\ u = C_2 + v' \left(\frac{\mu'_1}{\mu'_2 - \mu'_1}\right) - u' \left(\frac{1}{\mu'_2 - \mu'_1}\right) + u' v' \left(\frac{1}{\mu'_2 - \mu'_1}\right) \end{cases},$$

where $C_1 = \frac{1}{\mu'_2 - \mu'_1} - \eta$ and $C_2 = -\mu - \frac{\mu'_1}{\mu'_2 - \mu'_1}$ are constant terms.

Therefore, given that the variance of the product of two centered random variables X and Y is $\sigma_{X*Y}^2 = \sigma_X^2 \sigma_Y^2 + (\sigma_{X,Y})^2$ and that $\sigma_{XY,Y} = 0$, we obtain:

$$\begin{cases} \sigma_v^2 = \frac{\sigma_{v'}^2}{(\mu'_2 - \mu'_1)^2} \\ \sigma_u^2 = \frac{1}{(\mu'_2 - \mu'_1)^2} \left((\mu'_1)^2 \sigma_{v'}^2 + \sigma_{u'}^2 + \sigma_{v'}^2 \sigma_{u'}^2 + (\sigma_{v',u'})^2 - 2\mu'_1 \sigma_{v',u'} \right) \end{cases}$$

The correlation between v and u is: $\rho_{uv} = \frac{\sigma_{uv}}{\sigma_u \sigma_v}$, with:

$$\begin{aligned} \sigma_{vu} &= \text{cov} \left(-\frac{v'}{\mu'_2 - \mu'_1}, C_2 + v' \left(\frac{\mu'_1}{\mu'_2 - \mu'_1} \right) \right. \\ &\quad \left. - u' \left(\frac{1}{\mu'_2 - \mu'_1} \right) + u' v' \left(\frac{1}{\mu'_2 - \mu'_1} \right) \right) \\ &= \frac{1}{(\mu'_2 - \mu'_1)^2} (\sigma_{v',u'} - \mu'_1 \sigma_{v'}^2). \end{aligned}$$

Appendix 2

In this section, we describe how we calculated the genetic variance (u''^{**}) in the MTM model based on the genetic variances (u'_1, u'_2) of the ITM model.

The probability that the trait be of class 1 (2, or 3, etc.) is the same for both models, thus:

$$\begin{aligned} &\begin{cases} P(y = 1)_{ITM} = P(y = 1)_{MTM} \\ P(y = 2)_{ITM} = P(y = 2)_{MTM} \end{cases}, \\ &\Rightarrow \begin{cases} L(t''_1) = L(t''^{**}_1) \\ L(t''_2) - L(t''_1) = L(t''^{**}_2) \end{cases}, \\ &\Rightarrow \begin{cases} \frac{1}{1+e^{-t''_1}} = \frac{1}{1+e^{-t''^{**}_1}} \\ \frac{1}{1+e^{-t''_2}} - \frac{1}{1+e^{-t''_1}} = \frac{1}{1+e^{-t''^{**}_2}} \end{cases}, \end{aligned}$$

where L is the Logit distribution function.

$$\begin{aligned} &\Rightarrow \begin{cases} e^{-t''^{**}_1} = e^{-t''_1} \\ e^{-t''^{**}_2} = \frac{1}{\frac{1+2e^{-t''_2} + e^{-(t''_2+t''_1)}}{e^{-t''_2} - e^{-t''_1}}} \end{cases} \\ &\Rightarrow \begin{cases} t''^{**}_1 = t''_1 \\ t''^{**}_2 = \ln \left(\frac{1+2e^{-t''_2} + e^{-(t''_2+t''_1)}}{e^{-t''_2} - e^{-t''_1}} \right) \end{cases} \end{aligned}$$

Thus, the variance $\sigma_{u''^{**}}^2$ is:

$$\sigma_{u''^{**}}^2 = \sigma^2 \ln \left(\frac{1+2e^{-t''_2} + e^{-(t''_2+t''_1)}}{e^{-t''_2} - e^{-t''_1}} \right).$$

We see that, from the parameters of the ITM model, we can deduce the parameters of the MTM model.

Received: 6 February 2015 Accepted: 14 March 2016

Published online: 01 April 2016

References

- Scheiner SM, Lyman RF. The genetics of phenotypic plasticity I. Heritability. *J Evol Biol.* 1989;2:95–107.
- SanCristobal-Gaudy M, Elsen JM, Bodin L, Chevalet C. Prediction of the response to a selection for canalisation of a continuous trait in animal breeding. *Genet Sel Evol.* 1998;30:423–51.
- Garcia M, David I, Garreau H, Ibañez-Escriche N, Mallard J, Masson JP, et al. Comparisons of three models for canalising selection or genetic robustness. In: Proceedings of the 60th annual meeting of the European association for animal production: 24–27 Aug 2009; Barcelona; 2009.
- Hill WG, Zhang XS. Effects on phenotypic variability of directional selection arising through genetic differences in residual variability. *Genet Res.* 2004;83:121–32.
- Mulder HA, Bijma P, Hill WG. Prediction of breeding values and selection responses with genetic heterogeneity of environmental variance. *Genetics.* 2007;175:1895–910.
- SanCristobal-Gaudy M, Bodin L, Elsen JM, Chevalet C. Genetic components of litter size variability in sheep. *Genet Sel Evol.* 2001;33:249–71.
- Rönnegård L, Felleki M, Fikse F, Mulder HA, Strandberg E. Genetic heterogeneity of residual variance—estimation of variance components using double hierarchical generalized linear models. *Genet Sel Evol.* 2010;42:8.
- Sorensen DA, Andersen S, Gianola D, Korsgaard I. Bayesian inference in threshold models using Gibbs sampling. *Genet Sel Evol.* 1995;27:229–49.
- Ibáñez-Escriche N, Garcia M, Sorensen D. GSEVM v.2: MCMC software to analyze genetically structured environmental variance models. *J Anim Breed Genet.* 2010;127:249–51.
- Perez-Cabal MA, Campos G, Vazquez AI, Gianola D, Rosa GJ, Weigel KA, et al. Genetic evaluation of susceptibility to clinical mastitis in Spanish Holstein cows. *J Dairy Sci.* 2009;92:3472–80.
- Vazquez AI, Gianola D, Bates D, Weigel KA, Heringstad B. Assessment of Poisson, logit, and linear models for genetic analysis of clinical mastitis in Norwegian Red cows. *J Dairy Sci.* 2009;92:739–48.
- Gianola D. Theory and analysis of threshold characters. *J Anim Sci.* 1982;54:1079–96.
- Ibáñez-Escriche N, Garcia M, Sorensen D. GSEVM v. 3: MCMC software to analyze genetically structured environmental variance models. *J Anim Breed Genet.* 2010;127:249–51.

14. Misztal I, Gianola D, Foulley JL. Computing aspects of a nonlinear method of sire evaluation for categorical data. *J Dairy Sci.* 1989;72:1557–68.
15. Janssens S, Vandepitte W, Bodin L. Genetic parameters for litter size in sheep: natural versus hormone-induced oestrus. *Genet Sel Evol.* 2004;36:43–562.
16. Jousseins C, Trocmé E, Delmas D, Ducourtieux C, Mouret J, Dugat J, et al. Les coûts de production en élevage ovin allaitant dans le Sud-ouest—Campagne 2010. Réseaux d'élevage pour le conseil et la prospective, collection référence. Paris: Institut de l'Élevage; 2012.
17. Gilmour AR, Gogel BJ, Cullis BR, Thompson R. ASReml user guide release 3.0. Hemel Hempstead: VSN International Ltd; 2009.
18. Schwarz G. Estimating the dimension of a model. *Ann Stat.* 1978;6:461–4.
19. Foulley JL, Gianola D. Statistical analysis of ordered categorical data via a structural heteroskedastic threshold model. *Genet Sel Evol.* 1996;28:249–73.
20. Hazel LN. The genetic basis for constructing selection indexes. *Genetics.* 1943;28:476–90.
21. Amer PR, Bodin L. Quantitative genetic selection for twinning rate in ewes. *Proc N Z Soc Anim Prod.* 2006;66:429–33.
22. Scheiner SM, Lyman RF. The genetics of phenotypic plasticity. II. Response to selection. *J Evol Biol.* 1991;4:23–50.
23. Matos CA, Thomas DL, Gianola D, Tempelman RJ, Young LD. Genetic analysis of discrete reproductive traits in sheep using linear and nonlinear models: I. Estimation of genetic parameters. *J Anim Sci.* 1997;75:76–87.
24. Janssens S, Vandepitte W, Bodin L. Genetic parameters for litter size in sheep: natural versus hormone-induced oestrus. *Genet Sel Evol.* 2004;36:543–62.
25. Matika O, van Wyk JB, Erasmus GJ, Baker RL. Genetic parameter estimates in Sabi sheep. *Livest Prod Sci.* 2003;79:17–28.
26. Garreau H, Bolet G, Larzul C, Robert-Granié C, Saleil G, SanCristobal M, et al. Results of four generations of a canalising selection for rabbit birth weight. *Livest Sci.* 2008;119:55–62.
27. Gutiérrez JP, Nieto B, Piqueras P, Ibáñez N, Salgado C. Genetic parameters for canalisation analysis of litter size and litter weight traits at birth in mice. *Genet Sel Evol.* 2006;38:445–62.
28. Ros M, Sorensen D, Waagepetersen R, Dupont-Nivet M, SanCristobal M, Bonnet JC, et al. Evidence for genetic control of adult weight plasticity in the snail *Helix aspersa*. *Genetics.* 2004;168:2089–97.
29. Rowe SJ, White IM, Avendaño S, Hill WG. Genetic heterogeneity of residual variance in broiler chickens. *Genet Sel Evol.* 2006;38:617–35.
30. Sorensen D, Waagepetersen R. Normal linear models with genetically structured residual variance heterogeneity: a case study. *Genet Res.* 2003;82:207–22.
31. Ibáñez-Escriche N, Moreno A, Nieto B, Piqueras P, Salgado C, Gutiérrez JP. Genetic parameters related to environmental variability of weight traits in a selection experiment for weight gain in mice; signs of correlated canalised response. *Genet Sel Evol.* 2008;40:279–93.
32. Yang Y, Christensen OF, Sorensen D. Analysis of a genetically structured variance heterogeneity model using the Box–Cox transformation. *Genet Res (Camb).* 2011;93:33–46.
33. Martin P, Raoul J, Bodin L. Effects of the *FecL* major gene in the Lacaune meat sheep population. *Genet Sel Evol.* 2014;46:48.
34. Bodin L, Elsen JM. Variability of litter size of French sheep breeds following natural or induced ovulation. *Anim Sci.* 1989;48:535–41.

Submit your next manuscript to BioMed Central
and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

