

Supplementary Information

Throughout this document, when the mean and variance of the Normal distribution probability and cumulative density functions are not mentioned, they are assumed equal to 0 and 1, respectively for the mean and variance. Hence, by default, $f_{\mathcal{N}}(\ell)$ and $\Phi_{\mathcal{N}}(\ell)$ correspond to a standard Normal distribution.

A Proof for analytical solutions for binary traits

Equivalence between a Binomial/Probit GLMM and the threshold model We consider the probability of the observed data z to be 1, knowing that the latent trait value is ℓ (hence from the GLMM perspective):

$$P(z = 1) = \Phi_{\mathcal{N}}(\ell) = P(X < \ell) = P(0 < \ell - X), \quad (\text{S1})$$

where X is a random variable following a standard Normal distribution. Let ϵ be a variable such that $\epsilon = -X$, then ϵ also follows a standard Normal distribution. This allows us to rewrite the above equation as:

$$P(z = 1) = P(0 < \ell + \epsilon), \quad (\text{S2})$$

which is the probability defined from the threshold model perspective. The addition of the ϵ term, which has a variance of 1, explains the appearance of the so-called “link variance”. Note that the exact same reasoning apply to the logit link function, which inverse function is the cumulative distribution function of a logistic distribution of location 0 and scale 1. Thus, the “link variance” associated with a logit link is $\pi^3/2$.

This equivalence, and the fact that the liability and expected data scales are different is illustrated in Fig. S1.

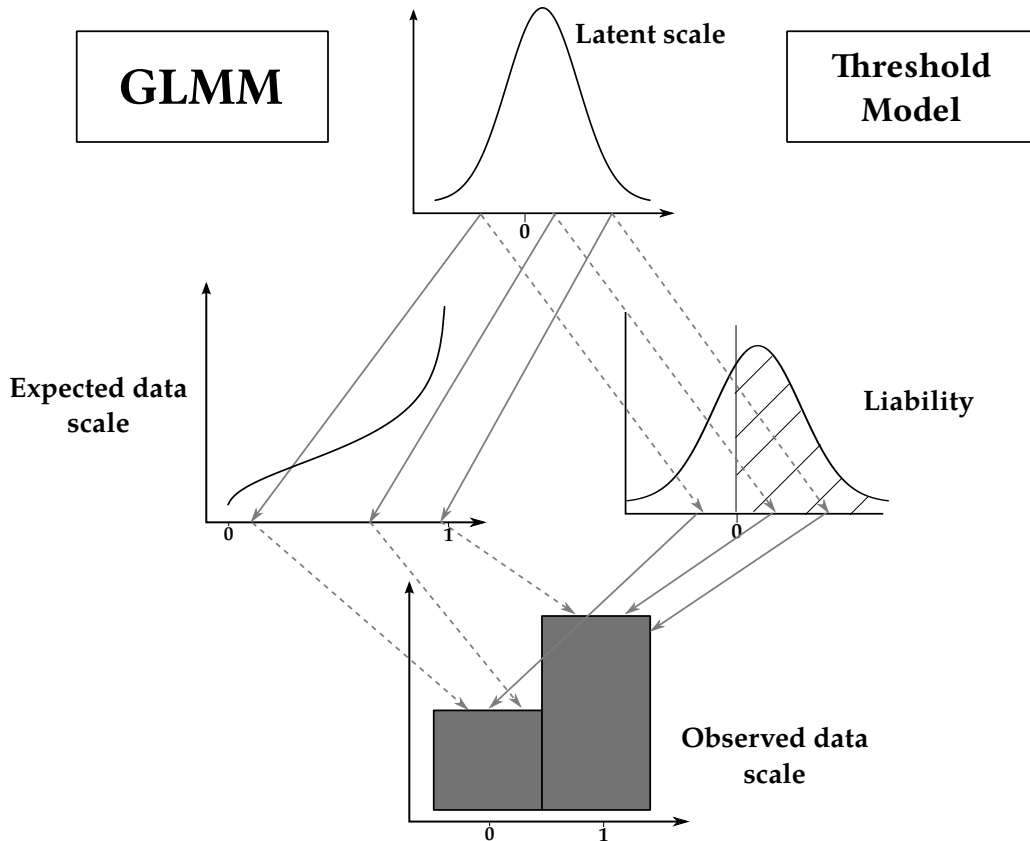


Figure S1: Comparison between a Binomial/probit model (left) and a threshold model (right). GLMMs are characterised by a probit-link transformation followed by a Binomial distribution, whereas the threshold model uses a threshold on the liability scale to directly define the observed binary values. Deterministic relationships are denoted using grey plain arrows, whereas stochastic relationships are denoted using grey dashed arrows.

Observed phenotypic mean Let p be the average phenotype (as well as the probability of 1, which is an equivalent definition). Note that because of the equivalence between the GLMM and the threshold model:

$$p = \int \Phi_{\mathcal{N}}(\ell) f_{\mathcal{N}}(\ell, \mu, V_{\text{P}}) d\ell = 1 - \Phi_{\mathcal{N}}(0, \mu, V_{\text{P}} + 1), \quad (\text{S3})$$

which provides a quick analytical way to compute p . If fixed effects are included in the model, p is simply this calculation averaged over the elements of $\mathbf{X}\hat{\mathbf{b}}$:

$$p = 1 - \frac{1}{N} \sum_i \Phi_{\mathcal{N}}(0, \mu + (\mathbf{X}\hat{\mathbf{b}})_i, V_{\text{P}} + 1). \quad (\text{S4})$$

Observed phenotypic variance The observed variance of the phenotype only depends on the mean p and is $p(1-p)$. This is a property of a binomial distribution with only one trial and is consistent with Eq. ??:

$$\begin{aligned} V_{\text{P,obs}} &= \int (\Phi_{\mathcal{N}}(\ell) - p)^2 f_{\mathcal{N}}(\ell, \mu, V_{\text{P}}) d\ell + \int v(\ell) f_{\mathcal{N}}(\ell, \mu, V_{\text{P}}) d\ell, \\ &= \int \Phi_{\mathcal{N}}(\ell)^2 f_{\mathcal{N}}(\ell, \mu, V_{\text{P}}) d\ell - p^2 + \int \Phi_{\mathcal{N}}(\ell) f_{\mathcal{N}}(\ell, \mu, V_{\text{P}}) d\ell - \int \Phi_{\mathcal{N}}(\ell)^2 f_{\mathcal{N}}(\ell, \mu, V_{\text{P}}) d\ell, \\ &= p - p^2, \\ &= p(1-p). \end{aligned} \quad (\text{S5})$$

This very simple analytical solution allows to easily compute $V_{\text{P,obs}}$. Of course, when fixed effects are included, p should be computed using Eq. S4.

Consistency with Dempster & Lerner equation Using Dempster and Lerner (1950) equation and the threshold model framework, one would compute the observed-scale heritability as:

$$h_{\text{obs, DL}}^2 = \frac{t^2}{p(1-p)} \frac{V_{\text{A}}}{V_{\text{P}} + 1} \quad (\text{S6})$$

Note that the $V_{\text{P}} + 1$ arise because of the addition of the so-called ‘‘link variance’’ (Nakagawa and Schielzeth, 2010). The term t is the probability density of a standard normal distribution evaluated at the p th quantile. Hence (standardising the latent distribution):

$$t = f_{\mathcal{N}}(\Phi^{-1}(p)) = f_{\mathcal{N}}\left(\frac{\mu}{\sqrt{V_{\text{P}} + 1}}\right) \quad (\text{S7})$$

On the other hand, using our framework, one would compute the observed heritability by computing Ψ as in Eq. ??, and combine it with Eqs. ??&??:

$$h_{\text{obs}}^2 = \frac{\Psi^2 V_{\text{A}}}{p(1-p)} \quad (\text{S8})$$

By comparing Eqs. S6&S8, one can see the identity holds if, and only if:

$$\Psi = \frac{t}{\sqrt{V_{\text{P}} + 1}} = f_{\mathcal{N}}(0, \mu, V_{\text{P}} + 1) \quad (\text{S9})$$

In order to prove this identity, let us compute the ratio between the two:

$$\begin{aligned} \Psi / f_{\mathcal{N}}(0, \mu, V_{\text{P}} + 1) &= \frac{\int f_{\mathcal{N}}(x) f_{\mathcal{N}}(x, \mu, V_{\text{P}}) dx}{f_{\mathcal{N}}(0, \mu, V_{\text{P}} + 1)} \\ &= \frac{\sqrt{\frac{V_{\text{P}} + 1}{V_{\text{P}}}} \frac{1}{\sqrt{2\pi}} \int \exp\left[-\frac{1}{2} \left(x^2 + \frac{(x - \mu)^2}{V_{\text{P}}} - \frac{\mu^2}{V_{\text{P}} + 1}\right)\right] dx}{\frac{1}{\sqrt{2\pi}} \int \exp\left[-\frac{1}{2} \left(x^2 + \frac{(x - \mu)^2}{V_{\text{P}}}\right)\right] dx} \\ &= \frac{\sqrt{\frac{V_{\text{P}} + 1}{V_{\text{P}}}} \frac{1}{\sqrt{2\pi}} \int \exp\left[-\frac{1}{2} \left(\frac{(x + x V_{\text{P}} - \mu)^2}{V_{\text{P}}(V_{\text{P}} + 1)}\right)\right] dx}{\frac{1}{\sqrt{2\pi}} \int \exp\left[-\frac{1}{2} \left(\frac{(x - (\mu/(V_{\text{P}} + 1)))^2}{V_{\text{P}}/(V_{\text{P}} + 1)}\right)\right] dx} \\ &= \int f_{\mathcal{N}}\left(x, \frac{\mu}{V_{\text{P}} + 1}, \frac{V_{\text{P}}}{V_{\text{P}} + 1}\right) dx \\ &= 1 \end{aligned} \quad (\text{S10})$$

Note that, if fixed effects are included in the model, Ψ can be computed by averaging over them, as in Eq. S4:

$$\Psi = \frac{1}{N} \sum_i f_{\mathcal{N}}(0, \mu + (\mathbf{X}\hat{\mathbf{b}})_i, V_{\text{P}} + 1) \quad (\text{S11})$$

B Proof for analytical solutions for Poisson traits

Observed phenotypic mean The observed phenotypic mean, hereafter noted λ , is the quantity (in absence of fixed effects):

$$\lambda = \int \exp(\ell) f_{\mathcal{N}}(\ell, \mu, V_{\text{P}}) d\ell = \exp(\mu + V_{\text{P}}/2) \quad (\text{S12})$$

after simplifications in Mathematica. Again, when fixed effects are included, it suffices to average over them:

$$\lambda = \frac{1}{N} \sum_i \exp(\mu + (\mathbf{X}\hat{\mathbf{b}})_i + V_{\text{P}}/2) = \frac{1}{N} \sum_i \lambda_i, \quad (\text{S13})$$

where λ_i is the observed mean conditional to $(\mathbf{X}\hat{\mathbf{b}})_i$.

Expected-scale variance In absence of fixed effects, the expected-scale variance is defined as follow (using Keonig's formula):

$$V_{\text{P,exp}} = \int \exp(\ell)^2 f_{\mathcal{N}}(\ell, \mu, V_{\text{P}}) d\ell - \lambda^2, \quad (\text{S14})$$

which, using Mathematica, simplifies into:

$$V_{\text{P,exp}} = \lambda^2 [\exp(V_{\text{P}}) - 1]. \quad (\text{S15})$$

When fixed effects, are included in the model, the formula does not simplify as much, because of the averaging:

$$V_{\text{P,exp}} = \exp(V_{\text{P}}) \frac{1}{N} \sum_i \exp(2\mu + 2(\mathbf{X}\hat{\mathbf{b}})_i + V_{\text{P}}) - \lambda^2 \quad (\text{S16})$$

Note that λ should be computed according to Eq. S15 in that case. Let us define the parameter Λ as

$$\Lambda = \frac{1}{N} \sum_i \exp(2\mu + 2(\mathbf{X}\hat{\mathbf{b}})_i + V_{\text{P}}) = \frac{1}{N} \sum_i \lambda_i^2, \quad (\text{S17})$$

so that $V_{\text{P,exp}} = \Lambda \exp(V_{\text{P}}) - \lambda^2$. The parameter Λ is the average of the squared values of λ_i . Unfortunately, $\Lambda \neq \lambda^2$, hence there is no further simplification.

Distribution variance Because the variance of a Poisson distribution is equal to the mean, the distribution variance reduces to λ :

$$V_{\text{dist}} = \int \exp(\ell) f_{\mathcal{N}}(\ell, \mu, V_{\text{P}}) d\ell = \lambda \quad (\text{S18})$$

When fixed effects are included in the model, λ should be computed as in Eq. S15.

Observed-scale additive variance In order to compute the observed-scale additive variance, we need the parameter Ψ defined in Eq. ?? in the main text. Again, because the derivative of an exponential is an exponential, the calculation reduces to λ :

$$\Psi = \int \exp(\ell) f_{\mathcal{N}}(\ell, \mu, V_{\text{P}}) d\ell = \lambda \quad (\text{S19})$$

Once again, when fixed effects are included in the model, λ should be computed as in Eq. S15.

The observed-scale additive variance is then computed as $\Psi^2 V_{\text{A}} = \lambda^2 V_{\text{A}}$.

Observed-scale heritability In presence of fixed effects in the model, the heritability is consequently the "simple" following ratio:

$$h_{\text{obs}}^2 = \frac{\lambda^2 V_{\text{A}}}{\Lambda \exp(V_{\text{P}}) - \lambda^2 + \lambda} \quad (\text{S20})$$

Negative-Binomial distribution All the above results can be extended to the Negative-Binomial distribution, which, compared to the Poisson distribution, includes an overdispersion parameter (hereby noted θ). The new parameter θ only affects the distribution variance such as:

$$h_{\text{obs}}^2 = \frac{\lambda^2 V_{\text{A}}}{\Lambda \exp(V_{\text{P}}) - \lambda^2 + \lambda + \exp(2(\mu + V_{\text{P}}))/\theta} \quad (\text{S21})$$

C Calculation of the Price-Robertson identity

The Price-Robertson identity (Robertson, 1966; Price, 1970) is based on the covariance between the (latent) additive genetic values and the expected fitness. This covariance can be computed using Koenig’s formula for covariance:

$$\Delta\mu = \text{cov}(a, W_{\text{exp}}) = E[a \times W_{\text{exp}}] - E[a] E[W_{\text{exp}}] = E[a \times W_{\text{exp}}], \quad (\text{S22})$$

since, by definition, $E[a] = 0$. Hence we obtain the following complex calculation:

$$\text{cov}(a, W_{\text{exp}}) = \iint a(W_{\text{exp}}(\ell))f(\ell|a)f(a)d\ell da = \iint aW_{\text{exp}}(\ell)f_{\mathcal{N}}(\ell, \mu + a, V_{\text{RE}} + V_{\text{O}})f_{\mathcal{N}}(a, 0, V_{\text{A}})d\ell da. \quad (\text{S23})$$

Double-integration can be difficult to solve, but we can simplify further the computation by noticing that:

$$\begin{aligned} \text{cov}(a, W_{\text{exp}}) &= \iint aW_{\text{exp}}(\ell)f_{\mathcal{N}}(\ell, \mu + a, V_{\text{RE}} + V_{\text{O}})f_{\mathcal{N}}(a, 0, V_{\text{A}})d\ell da, \\ &= \int af_{\mathcal{N}}(a, 0, V_{\text{A}}) \int W_{\text{exp}}(\ell)f_{\mathcal{N}}(\ell, \mu + a, V_{\text{RE}} + V_{\text{O}})d\ell da. \end{aligned} \quad (\text{S24})$$

Noting that

$$E[W_{\text{exp}}|a] = \int W_{\text{exp}}(\ell)f_{\mathcal{N}}(\ell, \mu + a, V_{\text{RE}} + V_{\text{O}})d\ell, \quad (\text{S25})$$

we can simplify Eq. S24 into

$$\Delta\mu = \text{cov}(a, W_{\text{exp}}) = \int aE[W_{\text{exp}}|a]f_{\mathcal{N}}(a, 0, V_{\text{A}})da. \quad (\text{S26})$$

This computation of the expected response on the latent scale is the one implemented in the QGGLMM R package.

D Code for the R package QGglimm

The up-to-date code of the package can be found at the following GitHub repository:

<https://github.com/devillemereuil/qgglmm>

The package can be downloaded and installed in R from the same repository. It requires the packages `mvtnorm` and `R2Cuba`.

References

- Dempster, E. R. and Lerner, I. M. (1950). Heritability of Threshold Characters. *Genetics*, 35(2):212–236.
- Nakagawa, S. and Schielzeth, H. (2010). Repeatability for Gaussian and non Gaussian data: a practical guide for biologists. *Biological Reviews*, 85(4):935–956.
- Price, G. R. (1970). Selection and covariance. *Nature*, 227:520–521.
- Robertson, A. (1966). A mathematical model of the culling process in dairy cattle. *Animal Production*, 8:95–108.