

Non-Gaussian random effects in nonlinear mixed effects models in Pumas®

Mohamed Tarek¹, Vijay Ivaturi¹, Andreas Noack^{1,3}

¹Pumas-AI Inc., ³Julia Computing, Inc.

Abstract

Laplace approximation and its various approximations are some of the most commonly used techniques in pharmaceutical nonlinear mixed effects modelling to marginalize the random effects. Classically, only Gaussian random effects have been used with transformations applied after the fact. However recently Pumas®-NLME™ introduced the ability to use non-Gaussian random effects in a nonlinear mixed effects model. In this study, the use of non-Gaussian random effects in nonlinear mixed effects (NLME) models in Pumas [1] will be mathematically justified and its practical importance will be highlighted using 2 experiments.

Laplace Method

Nonlinear mixed effects model

- θ : deterministic population-level parameters known as fixed effects
- η : random per subject parameters known as random effects (assume a single subject)
- z : covariates of the subject, e.g. age, sex, etc.
- y : vector response variable given θ, η, z and the model
- Marginal probability:

$$p(y|z, \theta) = \int_{D_\eta} p(y, \eta|z, \theta) d\eta = \int_{D_\eta} p(y|\eta, z, \theta) p(\eta|\theta) d\eta$$

Laplace method - general

$$\int_D f(x) dx \approx f(x_0) \sqrt{(2\pi)^N / | -H |}$$

- f : a positive scalar-valued function of x
- x : vector of N integration variables
- x_0 : global maximizer of f or $\log f$ such that $\frac{d(\log f)}{dx} = 0$ and $\frac{d^2(\log f)}{dx^2}$ is negative definite
- H : $\frac{d^2(\log f)}{dx^2}$ at x_0

Laplace method - NLME

$$\int p(y, \eta) d\eta \approx p(y, \eta_0) \sqrt{(2\pi)^N / | -H_\eta |}$$

where:

- η : integration variables - random effects
- $p(y, \eta)$: joint probability of the observations y and the random effects η
- η_0 : empirical Bayes estimate
- H_η is $d^2(\log p)/d\eta^2$ at η_0

If the domain is constrained, we can also do variable substitution:

$$\int_D p(y, \eta) d\eta = \int_D p(y, \eta(\xi)) d\eta(\xi) = \int p(y, \eta(\xi)) |d\eta/d\xi| d\xi$$

where

- ξ : unconstrained integration variables
- $\eta(\xi)$: transformation function from the unconstrained integration variables to the random effects
- $|d\eta/d\xi|$: determinant of the Jacobian of the transformation $d\eta/d\xi$

Conclusion and Future Work

LA and FOCE are generic techniques that can be used to marginalize non-Gaussian random effects. Non-Gaussian random effects are necessary for maximal modelling flexibility. The unique properties of some non-Gaussian distributions can be used to handle certain data pathologies.

References

1. Chris Rackauckas, Yingbo Ma, Andreas Noack, Vaibhav Dixit, Patrick Kofod Mogensen, Simon Byrne, Shubham Maddhashiya, José Bayoán Santiago Calderón, Joakim Nyberg, Jogarao V.S. Gobburu, Vijay Ivaturi (2020). Accelerated Predictive Healthcare Analytics with Pumas, a High Performance Pharmaceutical Modeling and Simulation Platform. *bioRxiv* 2020.11.28.402297; doi: <https://doi.org/10.1101/2020.11.28.402297>

Laplace Method

Convergence

The modes are the same

$$\log f(x_1) < \log f(x_2) \\ f(x_1) < f(x_2)$$

Ratios are not

$$\log f(x_1) - \log f(x_2) = c \quad f(x_1)/f(x_2) = e^c \\ n \log f(x_1) - n \log f(x_2) = nc \quad \frac{e^{nf(x_1)}}{e^{nf(x_2)}} = e^{nc} \\ \frac{n \log f(x_1)}{n \log f(x_2)} = 1 + \frac{c}{\log f(x_2)}$$

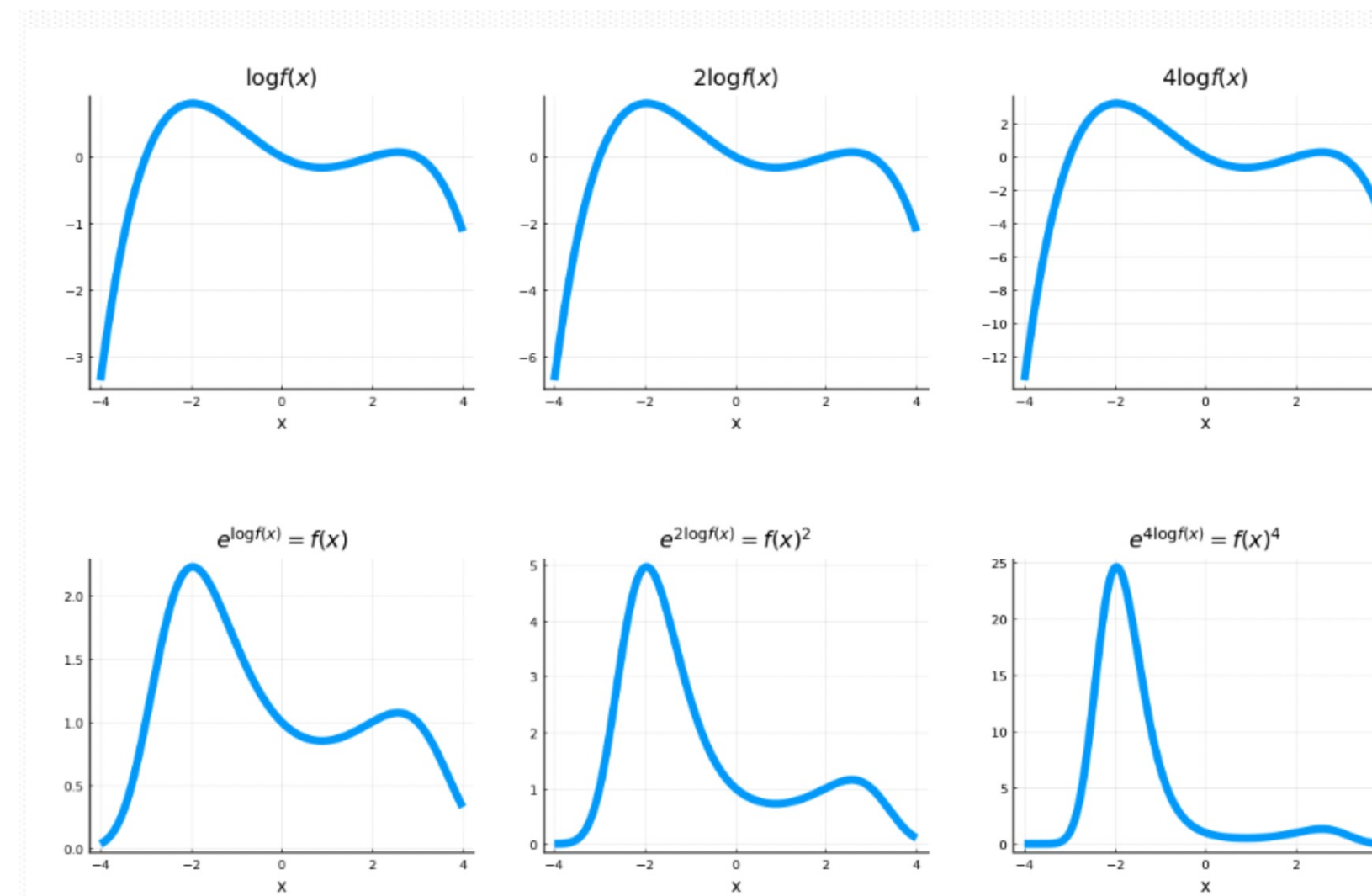
$$\eta \sim N(0, \Omega)$$

Gaussian additive error model:

$$y_i \sim N(\mu_i(\eta), \sigma^2) \\ \log p(\eta, y) = \sum_i -\frac{1}{2} \log(2\pi\Omega) - \frac{\eta^2}{2\Omega} - \frac{1}{2} \log(2\pi\sigma^2) - \frac{(y_i - \mu_i(\eta))^2}{2\sigma^2}$$

Gaussian proportional error model:

$$y_i \sim N(\mu_i(\eta), (\sigma\mu_i(\eta))^2) \\ \log p(\eta, y) = \sum_i -\frac{1}{2} \log(2\pi\Omega) - \frac{\eta^2}{2\Omega} - \frac{1}{2} \log(2\pi\mu_i(\eta)^2\sigma^2) - \frac{(y_i - \mu_i(\eta))^2}{2\mu_i(\eta)^2\sigma^2}$$



Generalized FOCE

Chain rule for Hessians:

$$(f \cdot g)''(x) = \frac{dg^T}{dx} \frac{d^2f}{dg^2} \frac{dg}{dx} + \sum_i \frac{\partial f}{\partial g_i} \frac{d^2g_i}{dx^2}$$

- f : scalar-valued function
- $\frac{d^2f}{dg^2}$: Hessian of f at $g(x)$
- g : vector-valued function with N outputs
- $\frac{\partial f}{\partial g_i}$: i^{th} component of the gradient of f at $g(x)$
- $\frac{dg}{dx}$: Jacobian of g at x
- $\frac{d^2g_i}{dx^2}$: the Hessian of g_i (i^{th} output of g) at x

If $g_i(x)$ is linear in x for all i :

$$\frac{d^2g_i}{dx^2} = 0 \\ (f \cdot g)''(x) = \frac{dg^T}{dx} \frac{d^2f}{dg^2} \frac{dg}{dx}$$

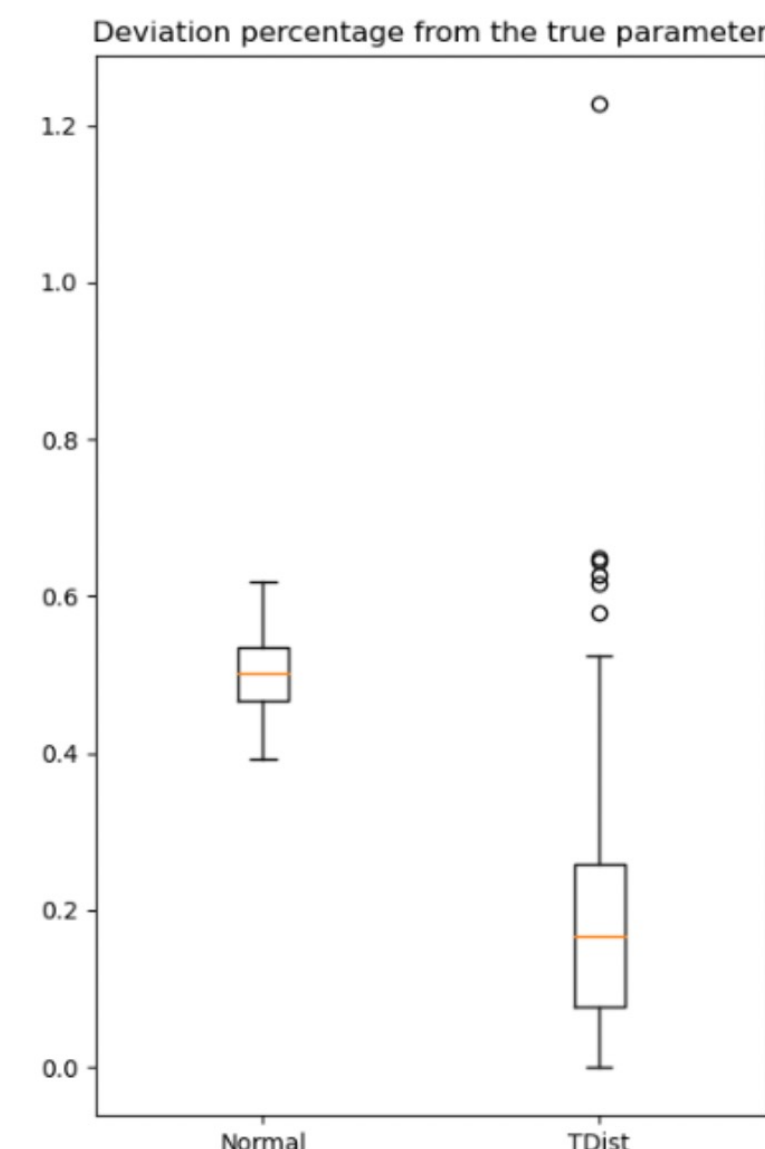
Syntax and Experiments

Syntax

```
model_bioav = @model begin
  @param begin
    thetaKa ∈ RealDomain(lower=0.0, init=0.2)
    thetaCL ∈ RealDomain(lower=0.0, init=0.1)
    thetaV ∈ RealDomain(lower=0.0, init=1.0)
    alphaBioav ∈ RealDomain(lower=1.0, upper=10.0, init=2.75)
    betaBioav ∈ RealDomain(lower=1.0, upper=10.0, init=2.75)
    sigma ∈ RealDomain(lower=0.0, init=0.1)
  end
  @random begin
    eta ~ MvNormal(Matrix{Float64}(0.1*I, 2, 2))
    etaBioav ~ Beta(alphaBioav, betaBioav)
  end
  @pre begin
    Ka = thetaKa
    CL = thetaCL*exp(eta[1])
    Vc = thetaV*exp(eta[2])
    bioav = (Depot=etaBioav,)
  end
  @dynamics Depots1Central1
  @derived begin
    cp = @. Central / Vc
    dv = @. Normal(cp, cp*sigma)
  end
end
```

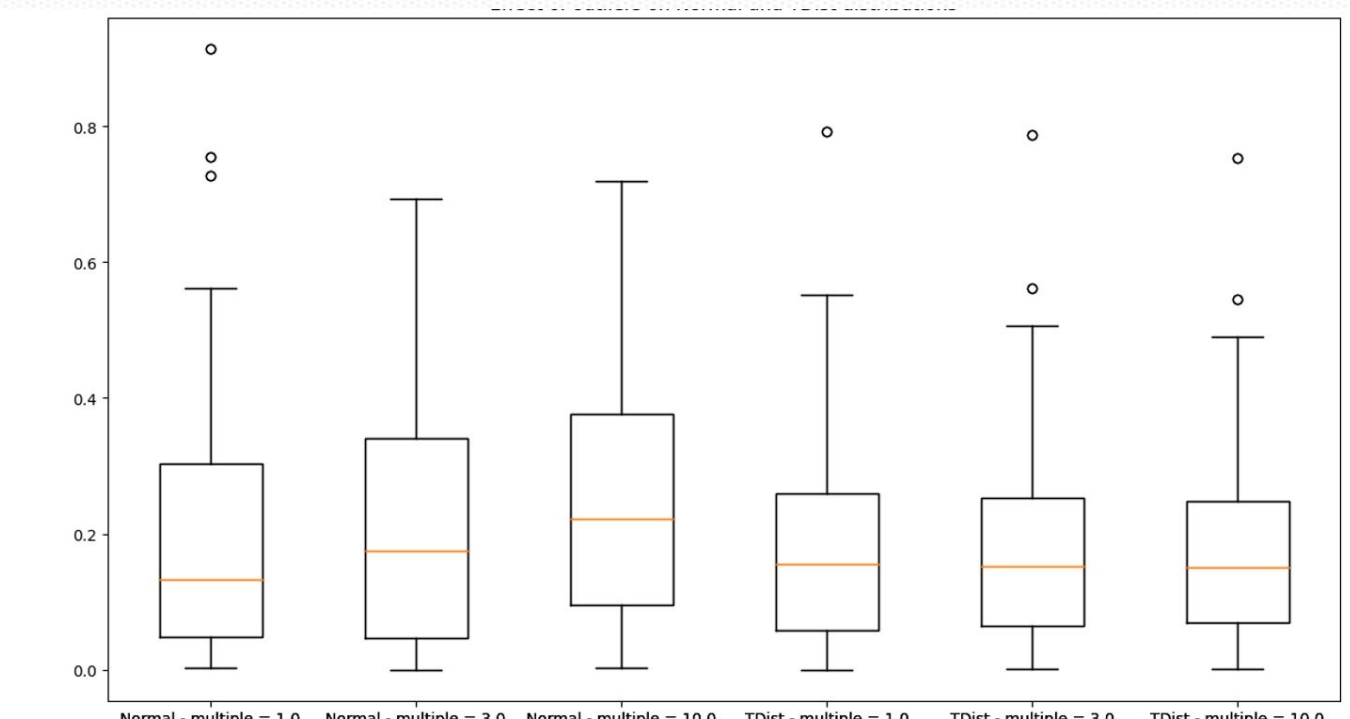
Experiment 1

- Sample N random effects from TDist – one for each subject
- Sample observations for each subject
- Fit a model with a Normal random effect to the sampled data
- Calculate the percentage deviation of the estimated population parameter from the true population parameter
- Repeat many times and plot a box plot of the deviation percentages



Experiment 2

- Sample $N - 1$ random effect – one for each subject
- Choose the N th random effect to be the maximum sampled random effect value multiplied by a multiple
- Sample observations for each subject
- Fit the model to the data and check the deviation percentage from the true population parameter
- Repeat many times and plot a box plot of the deviation percentages



```
Arcsine | Beta | BetaPrime | Biweight
Chi | Chisq | Cosine | Epanechnikov
Erlang | Exponential | FDist | Frchet
Gamma | GeneralizedExtremeValue
GeneralizedPareto | Gumbel | InverseGamma
InverseGaussian | Kolmogorov | Logistic
LogitNormal | LogNormal | MvLogNormal
NormalCanon | PGeneralizedGaussian
Rayleigh | Semicircle | SymTriangularDist
TDist | Triweight | Uniform | Weibull
```

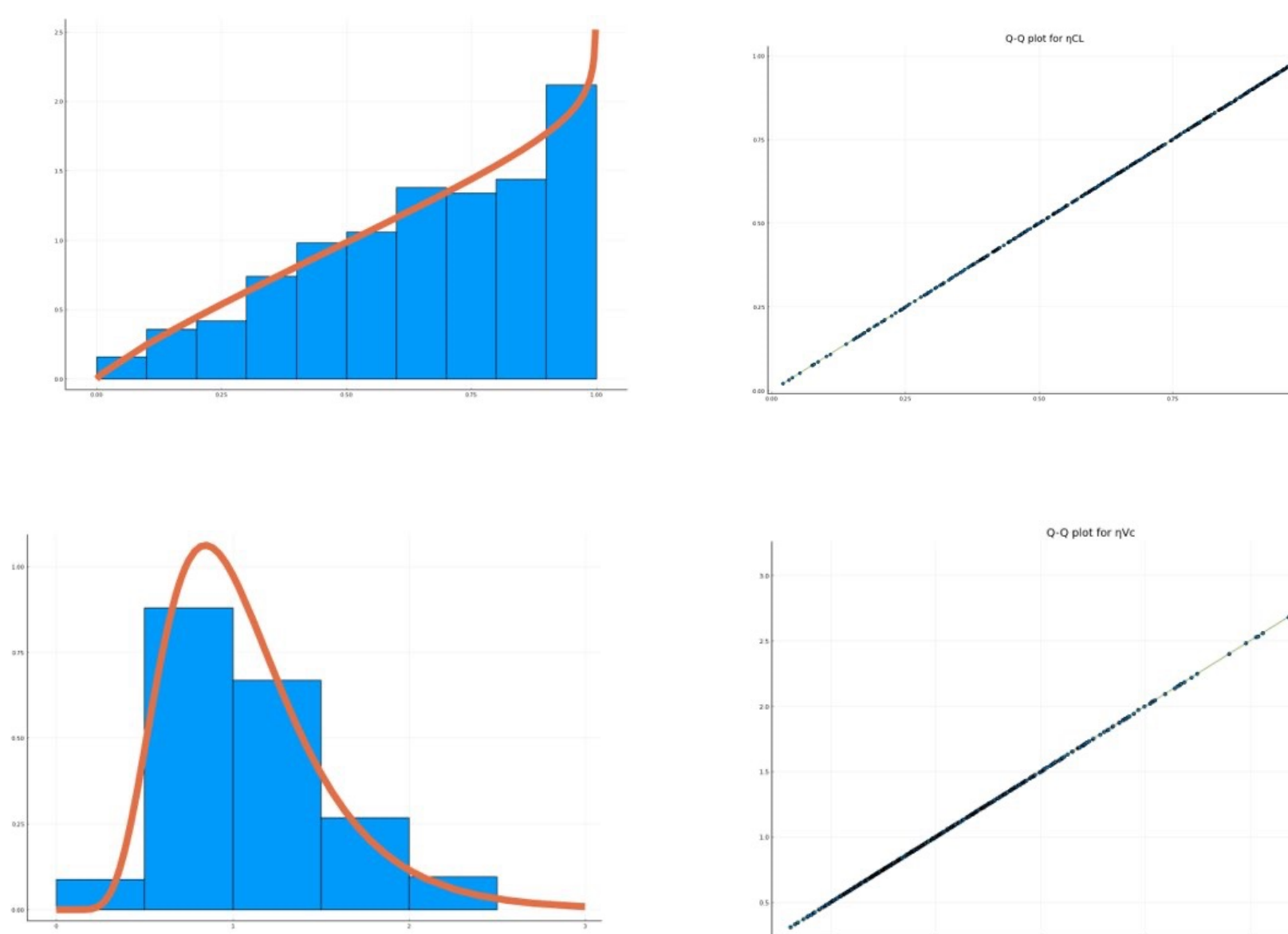
EBE Diagnostic

- Let $\{\eta_1, \eta_2, \dots, \eta_N\}$ be N independent samples from the random effect's distribution, one for each subject.
- For each subject i , let D observations be generated by the given model and $\eta = \eta_i$.
- Let the log joint probability of η and the D observations of subject i be $l_i(\eta; D)$.
- Let the model be identifiable.

then

$$\eta_i = \arg \max_{D \rightarrow \infty} \lim l_i(\eta; D)$$

Therefore, the EBEs will have the same distribution as the random effects.



```
mdl = @model begin
  @param begin
    theta ∈ VectorDomain(1, init=[0.5])
    alpha ∈ RealDomain(lower=0.0, init=1.3, upper=10.0)
    beta ∈ RealDomain(lower=0.0, init=3.3, upper=10.0)
    omega ∈ RealDomain(lower=0.01, init=0.1)
    sigma ∈ RealDomain(lower=0.01, upper=1.0, init=0.1)
  end
  @random begin
    etaCL ~ Beta(alpha, beta)
    etaVc ~ LogNormal(theta, omega)
  end
  @pre begin
    CL = theta[1] * etaCL
    Vc = etaVc
  end
  @dynamics Central1
  @derived begin
    conc := @. Central / Vc
    dv ~ @. Normal(conc, sigma)
  end
end
```