

# A Model Family for Hierarchical Data With Combined Normal and Conjugate Random Effects

**Geert Molenberghs**

Interuniversity Institute for Biostatistics and statistical Bioinformatics (I-BioStat)

Universiteit Hasselt & Katholieke Universiteit Leuven, Belgium

`geert.molenberghs@uhasselt.be`

`www.censtat.uhasselt.be`



Dallas, Texas

SAS Global Forum 2015

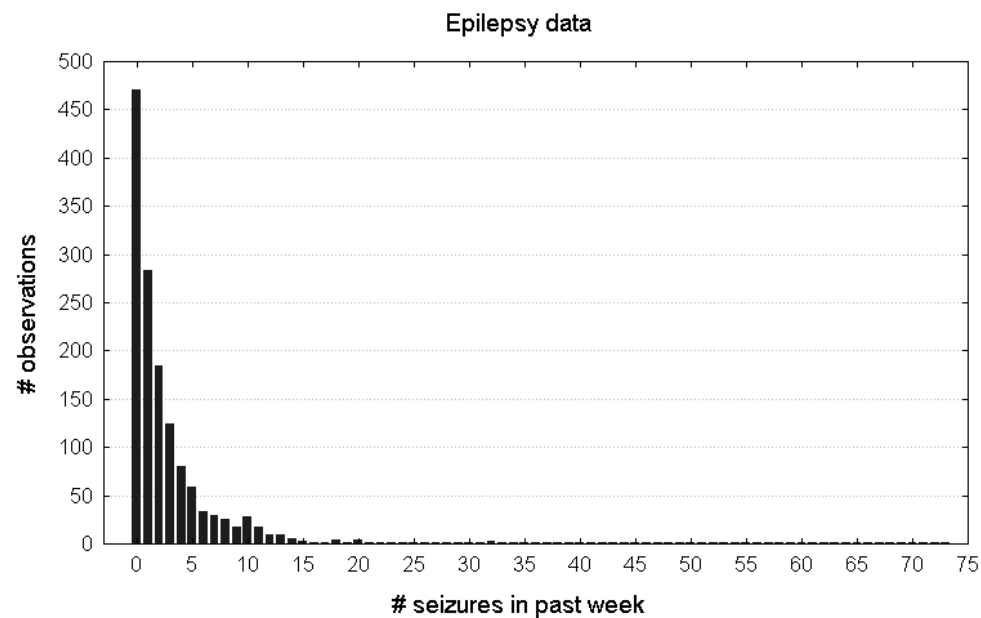
SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. © indicates USA registration.

Other brand and product names are trademarks of their respective companies.

# The Epilepsy Data

---

- Randomized, double-blind, parallel group multi-center study
- placebo (45)  $\longleftrightarrow$  new anti-epileptic drug (AED; 44)
- in combination with one or two other AED's
- 12-week run-in period
- 16 weeks of follow up (some until week 27)
- outcome of interest: is the number of epileptic seizures experienced during the last week
- research question: reduction in  $\#$  seizures by new therapy

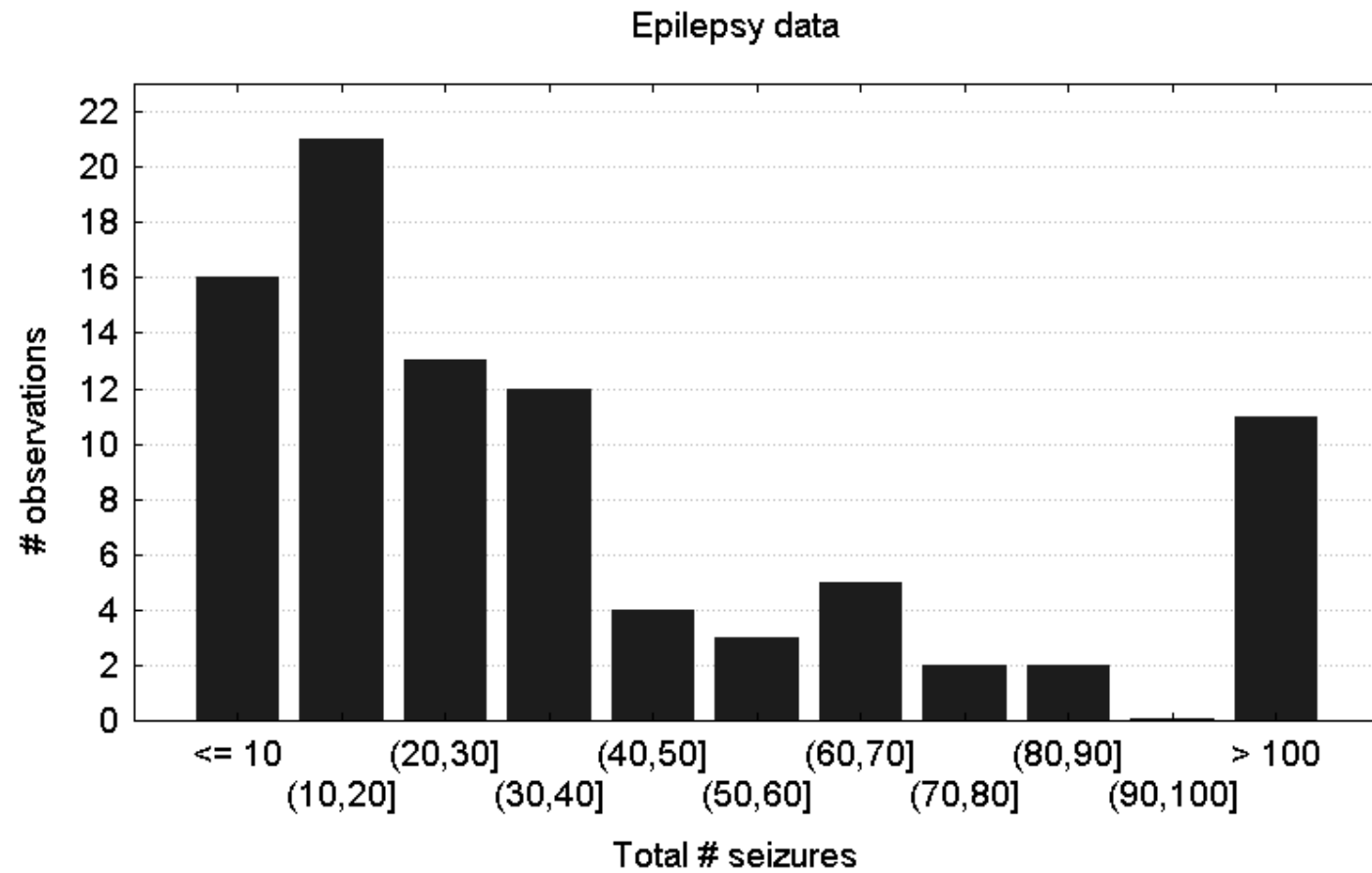


- Let  $Y_i$  now be the total number of seizures for subject  $i$ :

$$Y_i = \sum_{j=1}^{n_i} Y_{ij}$$

where  $Y_{ij}$  was the original (longitudinally measured) weekly outcome.

- Histogram:



- Number of measurements at some time points:

Week	# Observations		
	Placebo	Treatment	Total
1	45	44	89
5	42	42	84
10	41	40	81
15	40	38	78
16	40	37	77
17	18	17	35
20	2	8	10
27	0	3	3

- As these sums are not taken over an equal number of visits for all subjects, the above histogram is not a 'fair' one as it does not account for differences in  $n_i$  for this.
- We will therefore use the following Poisson model:

$$Y_i \sim \text{Poisson}(\lambda_i)$$

$$\ln(\lambda_i/n_i) = \mathbf{x}_i' \boldsymbol{\beta}$$

- The regression model is equivalent to

$$\lambda_i = n_i \exp(\mathbf{x}_i' \boldsymbol{\beta}) = \exp(\mathbf{x}_i' \boldsymbol{\beta} + \ln n_i)$$

- Since  $n_i$  is the number of weeks for which the number of seizures was recorded for subject  $i$ ,  $\exp(\mathbf{x}_i' \boldsymbol{\beta})$  is the average number of seizures per week
- $\ln n_i$  is called an offset in the above model
- Covariates in  $\mathbf{x}_i$ : treatment & baseline seizure rate

# Onychomycosis (Toenail) Data

---

*De Backer, De Keyser, De Vroey, Lesaffre (British Journal of Dermatology 1996)*

- **T**oenail **D**ermatophyte **O**nychomycosis: Common toenail infection, difficult to treat, affecting more than 2% of population
- Classical treatments with antifungal compounds need to be administered until the whole nail has grown out healthy
- New compounds have been developed that reduce treatment to 3 months
- Randomized, double-blind, parallel group, multicenter study for the comparison of two such new compounds (*A* and *B*) for oral treatment

- Design:

- ▷  $2 \times 189$  patients randomized, 36 centers
- ▷ 48 weeks of total follow up (12 months)
- ▷ 12 weeks of treatment (3 months)
- ▷ Measurements at months 0, 1, 2, 3, 6, 9, 12

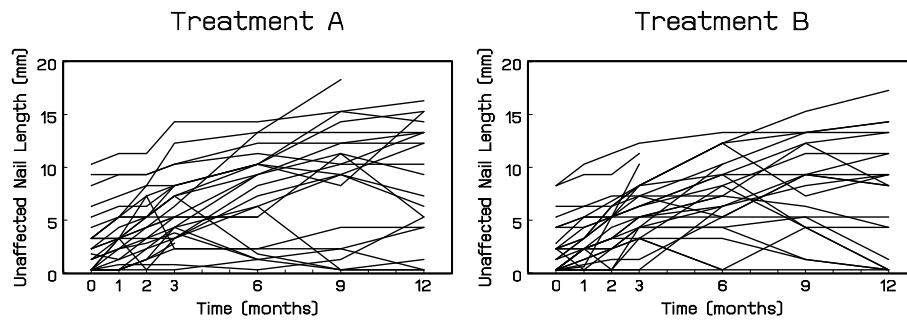
- General research question:

Are both treatments equally effective for the treatment of TDO?



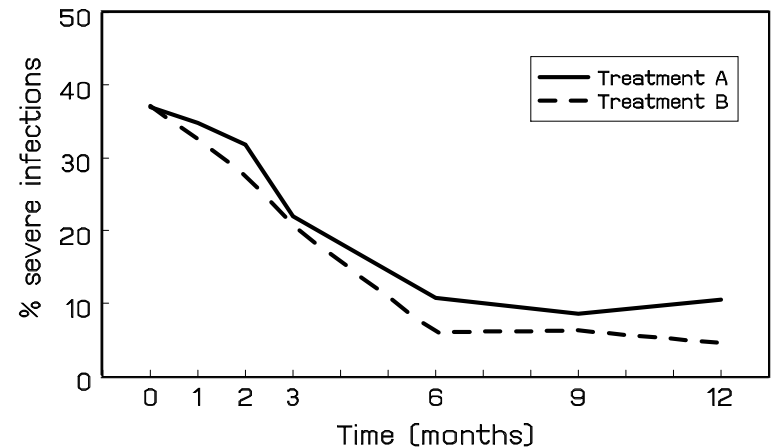
## Unaffected nail length (mm)?

- As response is related to toe size, we restrict to patients with big toenail as target nail  $\implies$  150 and 148 subjects
- 30 randomly selected profiles, in each group:



## Severity relative to treatment of TDO?

Toenail data



Complication: Dropout (24%):

# Recurrent Asthma Attacks in Children

---

*Duchateau and Janssen (2007)*

- Asthma is occurring more frequently in very young children (6–24 months)
- **Prevention trial:** New application of existing anti-allergic drug for high-risk children
- placebo       $\longleftrightarrow$       drug
- Usually more than one event       $\longrightarrow$       diary
- Recorded:
  - ▷ Start and end of risk period
  - ▷ Occurrence of asthma event

# Methodological Context

Study type	Outcome type			
	$N(\mu, \sigma^2)$	<del><math>N(\mu, \sigma^2)</math></del>		
$Y_i$	<b>Effect</b>	<b>Param</b>	<b>Effect</b>	<b>Param</b>
	Mean	$\mu$	Mean	$\mu$
	Var.	$\sigma^2$	Var.	$\mu$ (& $\theta_i$ )
	Corr.	—	Corr.	$(\theta_i)$
$(Y_{i1}, \dots, Y_{in_i})$	<b>Effect</b>	<b>Param</b>	<b>Effect</b>	<b>Param</b>
	Mean	$\mu_i$	Mean	$\mu_i$
	Var.	$b_i$ & $\Sigma_i$	Var.	$b_i$ (& $\theta_i$ )
	Corr.	$b_i$ & $\Sigma_i$	Corr.	$b_i$ (& $\theta_i$ )

# The Generalized Linear Model

---

- Suppose a sample  $Y_1, \dots, Y_N$  of independent observations is available
- All  $Y_i$  have densities  $f(y_i|\theta_i, \phi)$  which belong to the exponential family:

$$f(y|\theta_i, \phi) = \exp \{ \phi^{-1} [y\theta_i - \psi(\theta_i)] + c(y, \phi) \}$$

- $\theta_i$  the natural parameter
- Linear predictor:  $\theta_i = \mathbf{x}_i' \boldsymbol{\beta}$
- $\theta$  is the scale parameter (overdispersion parameter)
- Mean–variance relationship:

$$\text{Var}(Y) = \phi \psi'' [\psi'^{-1}(\mu)] = \phi v(\mu)$$

- The function  $v(\mu)$  is called the variance function.
- The function  $\psi'^{-1}$  which expresses  $\theta$  as function of  $\mu$  is called the link function.
- $\psi'$  is *the inverse link function*

# Example: The Normal Model

---

- Model:

$$Y \sim N(\mu, \sigma^2)$$

- Density function:

$$\begin{aligned} f(y|\theta, \phi) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{\sigma^2}(y - \mu)^2\right\} \\ &= \exp\left\{\frac{1}{\sigma^2}\left(y\mu - \frac{\mu^2}{2}\right) + \left(\frac{\ln(2\pi\sigma^2)}{2} - \frac{y^2}{2\sigma^2}\right)\right\} \end{aligned}$$

- Exponential family:

- ▷  $\theta = \mu$

- ▷  $\phi = \sigma^2$

- ▷  $\psi(\theta) = \theta^2/2$

- ▷  $c(y, \phi) = \frac{\ln(2\pi\phi)}{2} - \frac{y^2}{2\phi}$

- Mean and variance function:

- ▷  $\mu = \theta$

- ▷  $v(\mu) = 1$

- Note that, under this normal model, the mean and variance are not related:

$$\phi v(\mu) = \sigma^2$$

- The link function is here the identity function:  $\theta = \mu$

# Example: The Bernoulli Model

---

- Model:

$$Y \sim \text{Bernoulli}(\pi)$$

- Density function:

$$\begin{aligned} f(y|\theta, \phi) &= \pi^y(1 - \pi)^{1-y} \\ &= \exp \{y \ln \pi + (1 - y) \ln(1 - \pi)\} \\ &= \exp \left\{ y \ln \left( \frac{\pi}{1 - \pi} \right) + \ln(1 - \pi) \right\} \end{aligned}$$



- Exponential family:

- ▷  $\theta = \ln\left(\frac{\pi}{1-\pi}\right)$

- ▷  $\phi = 1$

- ▷  $\psi(\theta) = \ln(1 - \pi) = \ln(1 + \exp(\theta))$

- ▷  $c(y, \phi) = 0$

- Mean and variance function:

- ▷  $\mu = \frac{\exp \theta}{1 + \exp \theta} = \pi$

- ▷  $v(\mu) = \frac{\exp \theta}{(1 + \exp \theta)^2} = \pi(1 - \pi)$

- Note that, under this model, the mean and variance are related:

$$\phi v(\mu) = \mu(1 - \mu)$$

- The link function here is the logit link:  $\theta = \ln\left(\frac{\mu}{1-\mu}\right)$

# Example: The Poisson Model

---

- Model:

$$Y \sim \text{Poisson}(\lambda)$$

- Density function:

$$f(y|\theta, \phi) = \frac{e^{-\lambda} \lambda^y}{y!}$$
$$= \exp\{y \ln \lambda - \lambda - \ln y!\}$$

- Exponential family:

- ▷  $\theta = \ln \lambda$

- ▷  $\phi = 1$

- ▷  $\psi(\theta) = \lambda = \exp \theta$

- ▷  $c(y, \phi) = -\ln y!$

- Mean and variance function:

- ▷  $\mu = \exp \theta = \lambda$

- ▷  $v(\mu) = \exp \theta = \lambda$

- Note that, under this model, the mean and variance are related:

$$\phi v(\mu) = \mu$$

- The link function is here the log link:  $\theta = \ln \mu$

# Standard Overdispersion Models

---

- $\phi \neq 1 \implies \text{Var}(Y_i) = \phi\mu_i$

- **Two-stage approach:**

- ▷ **Formulation:**

$$Y_i | \lambda_i \sim \text{Poi}(\lambda_i)$$

$$\mathbf{E}(\lambda_i) = \mu_i$$

$$\text{Var}(\lambda_i) = \sigma_i^2$$

- ▷ **Implications:**

$$\mathbf{E}(Y_i) = \mathbf{E}[\mathbf{E}(Y_i | \lambda_i)] = \mathbf{E}(\lambda_i) = \mu_i$$

$$\text{Var}(Y_i) = \mathbf{E}[\text{Var}(Y_i | \lambda_i)] + \text{Var}[\mathbf{E}(Y_i | \lambda_i)] = \mathbf{E}(\lambda_i) + \text{Var}(\lambda_i) = \mu_i + \sigma_i^2$$

- **Repeated-measures version:**

$$\boldsymbol{\lambda}_i = (\lambda_{i1}, \dots, \lambda_{in_i})'$$

$$E(\boldsymbol{\lambda}_i) = \boldsymbol{\mu}_i$$

$$\text{Var}(\boldsymbol{\lambda}_i) = \boldsymbol{\Sigma}_i$$

$$E(\mathbf{Y}_i) = \boldsymbol{\mu}_i$$

$$\text{Var}(\mathbf{Y}_i) = M_i + \boldsymbol{\Sigma}_i$$

# The Standard Poisson-normal Model

---

- **General formulation:**

$$f_i(y_{ij}|\mathbf{b}_i, \boldsymbol{\beta}, \phi) = \exp \{ \phi^{-1} [y_{ij}\theta_{ij} - \psi(\theta_{ij})] + c(y_{ij}, \phi) \}$$

$$\eta(\mu_{ij}) = \eta[E(Y_{ij}|\mathbf{b}_i)] = \mathbf{x}'_{ij}\boldsymbol{\beta} + \mathbf{z}'_{ij}\mathbf{b}_i$$

$$\mathbf{b}_i \sim N(\mathbf{0}, D)$$

- **Poisson formulation:**

$$Y_{ij} \sim \text{Poi}(\lambda_{ij})$$

$$\ln(\lambda_{ij}) = \mathbf{x}'_{ij}\boldsymbol{\beta} + \mathbf{z}'_{ij}\mathbf{b}_i$$

$$\mathbf{b}_i \sim N(\mathbf{0}, D)$$

- **Marginal moments:**

$$\mu_{ij} = \exp\left(\mathbf{x}'_{ij}\boldsymbol{\beta} + \frac{1}{2}\mathbf{z}'_{ij}D\mathbf{z}_{ij}\right)$$

$$\text{Var}(\mathbf{Y}_i) = M_i + M_i\left(e^{\mathbf{z}_i D \mathbf{z}'_i} - J_{n_i}\right) M_i$$

- **Univariate version of the marginal moments:**

$$\mu_i = \exp\left(x'_i\boldsymbol{\beta} + \frac{1}{2}d\right)$$

$$\text{Var}(Y_i) = \mu_i + \mu_i^2(e^d - 1)$$

# A Combined Model: The Poisson-(...)-normal Model

---

- Model formulation:

$$Y_{ij} \sim \text{Poi}(\lambda_{ij})$$

$$\lambda_{ij} = \theta_{ij} \exp(\mathbf{x}'_{ij}\boldsymbol{\beta} + \mathbf{z}'_{ij}\mathbf{b}_i)$$

$$\mathbf{b}_i \sim N(\mathbf{0}, D)$$

$$E(\boldsymbol{\theta}_i) = E[(\theta_{i1}, \dots, \theta_{in_i})'] = \Phi_i$$

$$\text{Var}(\boldsymbol{\theta}_i) = \Sigma_i$$



- **Marginal moments:**

$$\mu_{ij} = \phi_{ij} \exp \left( \mathbf{x}'_{ij} \boldsymbol{\beta} + \frac{1}{2} \mathbf{z}'_{ij} D \mathbf{z}_{ij} \right)$$

$$\text{Var}(\mathbf{Y}_i) = M_i + M_i (P_i - J_{n_i}) M_i$$

$$p_{i,jk} = \exp \left( \frac{1}{2} \mathbf{z}'_{ij} D \mathbf{z}_{ik} \right) \cdot \frac{\sigma_{i,jk} + \phi_{ij} \phi_{ik}}{\phi_{ij} \phi_{ik}} \cdot \exp \left( \frac{1}{2} \mathbf{z}'_{ik} D \mathbf{z}_{ij} \right)$$

- **Univariate version of the marginal moments:**

$$\mu_i = \phi_i \exp \left( x'_i \boldsymbol{\beta} + \frac{1}{2} d \right)$$

$$\text{Var}(Y_i) = \mu_i + \mu_i^2 \cdot \left( e^d - 1 + \frac{\sigma_i^2}{\phi_i^2} e^d \right)$$

# Fully Parametric Specification

---

- So far, random effects are specified only in a moment-based fashion:
  - ▷ Mean
  - ▷ Variance
- Fully parametric specification is possible
- Gamma specification is sensible
- **Univariate case: the negative-binomial model**
- **Hierarchical case: the Poisson-gamma-normal model**

# The Negative-binomial Model

---

$$\lambda \sim \text{Gamma}(\alpha, \beta)$$

$$f(\lambda) = \frac{1}{\beta^\alpha \Gamma(\alpha)} \lambda^{\alpha-1} e^{-\lambda/\beta}$$

$$\begin{aligned} \mathbf{P}(Y = y) &= \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^{+\infty} \lambda^{\alpha-1} e^{-\lambda/\beta} \frac{\lambda^y e^{-\lambda}}{y!} \\ &= \binom{\alpha + y - 1}{\alpha - 1} \left( \frac{\beta}{\beta + 1} \right)^y \left( \frac{1}{\beta} \right)^\alpha \end{aligned}$$

# The Poisson-gamma-normal Model

- The marginal probability mass function:

$$P(\mathbf{Y}_i = \mathbf{y}_i) = \sum_{\mathbf{t}} \left[ \prod_{j=1}^{n_i} \binom{y_{ij} + t_j}{y_{ij}} \cdot \binom{\alpha_j + y_{ij} + t_j - 1}{\alpha_j - 1} \cdot (-1)^{t_j} \cdot \beta_j^{y_{ij} + t_j} \right]$$

$$\times \exp \left( \sum_{j=1}^{n_i} (y_{ij} + t_j) \mathbf{x}'_{ij} \boldsymbol{\beta} \right) \times \exp \left\{ \frac{1}{2} \left[ \sum_{j=1}^{n_i} (y_{ij} + t_j) \mathbf{z}'_{ij} \right] D \left[ \sum_{j=1}^{n_i} (y_{ij} + t_j) \mathbf{z}_{ij} \right] \right\}$$

- The special case of the Poisson-normal model:

$$P(\mathbf{Y}_i = \mathbf{y}_i) = \frac{1}{\prod_{j=1}^{n_i} y_{ij}!} \sum_{\mathbf{t}} \frac{(-1)^{\sum_{j=1}^{n_i} t_j}}{\prod_{j=1}^{n_i} t_j!} \cdot \exp \left[ \sum_{j=1}^{n_i} (y_{ij} + t_j) \mathbf{x}'_{ij} \boldsymbol{\beta} \right]$$

$$\times \exp \left\{ \frac{1}{2} \left[ \sum_{j=1}^{n_i} (y_{ij} + t_j) \mathbf{z}'_{ij} \right] D \left[ \sum_{j=1}^{n_i} (y_{ij} + t_j) \mathbf{z}_{ij} \right] \right\}$$

- **Univariate case for Poisson-gamma-normal model:**

$$P(Y_i = y_i) = \sum_{t=0}^{+\infty} \binom{y_i + t}{y_i} \binom{\alpha + y_i + t - 1}{\alpha - 1} \beta^{y_i+t} (-1)^t \exp \left[ x_i' \boldsymbol{\beta}(y_i + t) + \frac{1}{2} d(y_i + t)^2 \right]$$

- **Univariate case for Poisson-normal model:**

$$P(Y_i = y_i) = \frac{1}{y_i!} \sum_{t=0}^{+\infty} \frac{(-1)^t}{t!} \exp \left[ x_i' \boldsymbol{\beta}(y_i + t) + \frac{1}{2} d(y_i + t)^2 \right]$$

# Estimation

---

- Integrate over the gamma random effects **only**:

$$P(Y_{ij} = y_{ij} | \mathbf{b}_i) = \binom{\alpha_j + y_{ij} - 1}{\alpha_j - 1} \cdot \left( \frac{\beta_j}{1 + \kappa_{ij}\beta_j} \right)^{y_{ij}} \cdot \left( \frac{1}{1 + \kappa_{ij}\beta_j} \right)^{\alpha_j} \kappa_{ij}^{y_{ij}}$$

$$\kappa_{ij} = \exp[\mathbf{x}'_{ij}\boldsymbol{\beta} + \mathbf{z}'_{ij}\mathbf{b}_i]$$

- Can be fitted using the SAS procedure NLMIXED
- Model for the epilepsy data:

$$\ln(\lambda_{ij}) = \begin{cases} (\beta_{00} + b_i) + \beta_{01}t_{ij} & \text{if placebo} \\ (\beta_{10} + b_i) + \beta_{11}t_{ij} & \text{if treated,} \end{cases}$$

$$b_i \sim N(0, d)$$

# SAS Code: Poisson-normal Model

---

```
proc nlmixed data=test qpoints=50;
title 'Poisson-normal Model';
parms int0=0.5 slope0=-0.1 int1=1 slope1=0.1 sigma=1;
if (trt = 0) then eta = int0 + b + slope0*time;
else if (trt = 1) then eta = int1 + b + slope1*time;
lambda = exp(eta);
model nseizw ~ poisson(lambda);
random b ~ normal(0,sigma**2) subject = id;
estimate 'difference in slope' slope1-slope0;
estimate 'ratio of slopes' slope1/slope0;
estimate 'variance RIs' sigma**2;
run;
```

# SAS Code: Poisson-gamma-normal Model

---

```
proc nlmixed data=test qpoints=50;
title 'Poisson-Combined Model';
parms int0=0.5 slope0=-0.1 int1=1 slope1=0.1 sigma=1 alpha=5;
if (trt = 0) then eta = int0 + b + slope0*time;
else if (trt = 1) then eta = int1 + b + slope1*time;
lambda = exp(eta);
beta=1;
loglik=lgamma(alpha+y)-lgamma(alpha)+y*log(beta)
      -(y+alpha)*log(1+beta*lambda+y*eta);
model y ~ general(loglik);
random b ~ normal(0,sigma**2) subject = id;
estimate 'difference in slope' slope1-slope0;
estimate 'ratio of slopes' slope1/slope0;
estimate 'variance RIs' sigma**2;
run;
```



# The Epilepsy Data

Effect	Parameter	Poisson	Negative-binomial
		Estimate (s.e.)	Estimate (s.e.)
Intercept placebo	$\beta_{00}$	1.2662 (0.0424)	1.2594 (0.1119)
Slope placebo	$\beta_{01}$	-0.0134 (0.0043)	-0.0126 (0.0111)
Intercept treatment	$\beta_{10}$	1.4531 (0.0383)	1.4750 (0.1093)
Slope treatment	$\beta_{11}$	-0.0328 (0.0038)	-0.0352 (0.0101)
Negative-binomial parameter	$\alpha_1$	—	0.5274 (0.0255)
Negative-binomial parameter	$\alpha_2 = 1/\alpha_1$	—	1.8961 (0.0918)
Variance of random intercepts	$d$	—	—

Effect	Parameter	Poisson-normal	Combined
		Estimate (s.e.)	Estimate (s.e.)
Intercept placebo	$\beta_0$	0.8179 (0.1677)	0.9112 (0.1755)
Slope placebo	$\beta_1$	-0.0143 (0.0044)	-0.0248 (0.0077)
Intercept treatment	$\beta_0$	0.6475 (0.1701)	0.6555 (0.1782)
Slope treatment	$\beta_2$	-0.0120 (0.0043)	-0.0118 (0.0074)
Negative-binomial parameter	$\alpha_1$	—	2.4640 (0.2113)
Negative-binomial parameter	$\alpha_2 = 1/\alpha_1$	—	0.4059 (0.0348)
Variance of random intercepts	$d$	1.1568 (0.1844)	1.1289 (0.1850)

# Implications for Correlation Function

Model	Arm	Smallest value		Largest value	
		$\rho$	time pair	$\rho$	time pair
Poisson-normal	placebo	0.8577	26 & 27	0.8960	1 & 2
Poisson-normal	treatment	0.8438	26 & 27	0.8794	1 & 2
Combined	placebo	0.3041	26 & 27	0.3134	1 & 2
Combined	treatment	0.2234	1 & 2	0.3410	26 & 27

# Implications for Hypothesis Testing

---

<i>p-values</i>		
Model	$H_0 : \beta_{11} - \beta_{01} = 0$	$H_0 : \beta_{11}/\beta_{01} = 1$
<b>Poisson</b>	<b>0.0008</b>	<b>0.0038</b>
<b>Poisson-normal</b>	0.7115	<b>0.0376</b>
<b>negative-binomial</b>	<b>0.0131</b>	0.2815
<b>combined</b>	0.2260	0.1591

# The General Case

---

$$f_i(y_{ij}|\mathbf{b}_i, \boldsymbol{\xi}) = \exp \{ \phi^{-1} [y_{ij} \lambda_{ij} - \psi(\lambda_{ij})] + c(y_{ij}, \phi) \}$$

$$E(Y_{ij}|\theta_{ij}, \mathbf{b}_i) = \mu_{ij}^c = \theta_{ij} \kappa_{ij}$$

$$\theta_{ij} \sim \mathcal{G}_{ij}(\xi_{ij}, \sigma_{ij}^2)$$

$$\kappa_{ij} = g(\mathbf{x}'_{ij} \boldsymbol{\xi} + \mathbf{z}'_{ij} \mathbf{b}_i)$$

$$\mathbf{b}_i \sim N(\mathbf{0}, D)$$

$$\eta_{ij} = \mathbf{x}'_{ij} \boldsymbol{\xi} + \mathbf{z}'_{ij} \mathbf{b}_i$$

# Classical Conjugacy

---

$$f(y|\theta) = \exp \{ \phi^{-1} [yh(\theta) - g(\theta)] + c(y, \phi) \}$$

$$f(\theta) = \exp \{ \gamma [\psi h(\theta) - g(\theta)] + c^*(\gamma, \psi) \}$$

$$f(y) = \exp \left[ c(y, \phi) + c^*(\gamma, \psi) - c^* \left( \phi^{-1} + \gamma, \frac{\phi^{-1}y + \gamma\psi}{\phi^{-1} + \gamma} \right) \right]$$

# Strong Conjugacy

---

- Write:

$$f(y|\kappa\theta) = \exp \{ \phi^{-1} [y h(\kappa\theta) - g(\kappa\theta)] + c(y, \phi) \}$$

- Apply the transformation theorem:

$$f(\theta|\gamma, \psi) = \kappa \cdot f(\kappa\theta|\tilde{\gamma}, \tilde{\psi})$$

- Request the form:

$$f(\kappa\theta) = \exp \{ \gamma^* [\psi^* h(\kappa\theta) - g(\kappa\theta)] + c^{**}(\gamma^*, \psi^*) \}$$

- Leading to the marginal model:

$$f(y|\kappa) = \exp \left\{ c(y, \phi) + c^{**}(\gamma^*, \psi^*) + c^{**} \left( \phi^{-1} + \gamma^*, \frac{\phi^{-1}y + \gamma^*\psi^*}{\phi^{-1} + \gamma^*} \right) \right\}$$

# The Weibull Case

---

- The Weibull-gamma-normal model:

$$f(\mathbf{y}_i | \boldsymbol{\theta}_i, \mathbf{b}_i) = \prod_{j=1}^{n_i} \lambda \rho \theta_{ij} y_{ij}^{\rho-1} e^{\mathbf{x}'_{ij} \boldsymbol{\beta} + \mathbf{z}'_{ij} \mathbf{b}_i} e^{-\lambda y_{ij}^\rho \theta_{ij} e^{\mathbf{x}'_{ij} \boldsymbol{\beta} + \mathbf{z}'_{ij} \mathbf{b}_i}}$$

$$f(\boldsymbol{\theta}_i) = \prod_{j=1}^{n_i} \frac{1}{\beta_j^{\alpha_j} \Gamma(\alpha_j)} \theta_{ij}^{\alpha_j-1} e^{-\theta_{ij}/\beta_j}$$

$$f(\mathbf{b}_i) = \frac{1}{(2\pi)^{q/2} |D|^{1/2}} e^{-\frac{1}{2} \mathbf{b}_i' D^{-1} \mathbf{b}_i}$$

- **Marginal density:**

$$f(\mathbf{y}_i) = \sum_{(m_1, \dots, m_{n_i})} \prod_{j=1}^{n_i} \frac{(-1)^{m_j} \Gamma(\alpha_j + m_j + 1) \beta_j^{m_j+1}}{m_j! \Gamma(\alpha_j)} \lambda^{m_j+1} \rho y_{ij}^{(m_j+1)\rho-1} \\ \times \exp \left\{ (m_j + 1) \left[ \mathbf{x}'_{ij} \boldsymbol{\beta} + \frac{1}{2} (m_j + 1) \cdot \mathbf{z}'_{ij} D \mathbf{z}_{ij} \right] \right\}$$

- **Marginal moments:**

$$E(Y_{ij}^k) = \frac{\alpha_j B(\alpha_j - k/\rho, k/\rho + 1)}{\lambda^{k/\rho} \beta_j^{k/\rho}} \exp \left( -\frac{k}{\rho} \mathbf{x}'_{ij} \boldsymbol{\beta} + \frac{k^2}{2\rho^2} \mathbf{z}'_{ij} D \mathbf{z}_{ij} \right)$$

$$E(Y_{ij}) = \frac{\alpha_j B(\alpha_j - 1/\rho, 1/\rho + 1)}{\lambda^{1/\rho} \beta_j^{1/\rho}} \exp \left( -\frac{1}{\rho} \mathbf{x}'_{ij} \boldsymbol{\beta} + \frac{1}{2\rho^2} \mathbf{z}'_{ij} D \mathbf{z}_{ij} \right)$$



$$\begin{aligned} \text{Var}(Y_{ij}) &= \frac{\alpha_j}{\lambda^{2/\rho} \beta_j^{2\rho}} \exp\left(-\frac{2}{\rho} \mathbf{x}'_{ij} \boldsymbol{\beta} + \frac{1}{\rho^2} \mathbf{z}'_{ij} D \mathbf{z}_{ij}\right) \\ &\times \left[ B\left(\alpha_j - 2/\rho, 2/\rho + 1\right) \exp\left(\frac{1}{\rho^2} \mathbf{z}'_{ij} D \mathbf{z}_{ij}\right) - \alpha_j B\left(\alpha_j - \frac{1}{\rho}, \frac{1}{\rho} + 1\right)^2 \right] \end{aligned}$$

$$\begin{aligned} \text{Cov}(Y_{ij}, Y_{ik}) &= \frac{\alpha_j \alpha_k}{\lambda^{2/\rho} \beta_j^{1/\rho} \beta_k^{1/\rho}} \exp\left[-\frac{1}{\rho} (\mathbf{x}'_{ij} \boldsymbol{\beta} + \mathbf{x}'_{ik} \boldsymbol{\beta})\right] \\ &\times B\left(\alpha_j - \frac{1}{\rho}, \frac{1}{\rho} + 1\right) B\left(\alpha_k - \frac{1}{\rho}, \frac{1}{\rho} + 1\right) \\ &\times \exp\left[\frac{1}{2\rho^2} (\mathbf{z}'_{ij} D \mathbf{z}_{ij} + \mathbf{z}'_{ik} D \mathbf{z}_{ik})\right] \left[ \exp\left(\frac{1}{\rho^2} \mathbf{z}'_{ij} D \mathbf{z}_{ik}\right) - 1 \right] \end{aligned}$$

- **Identifiability** can be achieved in, for example, two ways:

$$f(\boldsymbol{\theta}_i) = \prod_{j=1}^{n_i} \frac{1}{\left(\frac{1}{\alpha_j}\right)^{\alpha_j} \Gamma(\alpha_j)} \theta_{ij}^{\alpha_j-1} e^{-\alpha_j \theta_{ij}}$$

$$f(\boldsymbol{\theta}_i) = \prod_{j=1}^{n_i} \delta_j e^{-\delta_j \theta_{ij}}$$

- For the second case:

$$E(Y_{ij}^k) = \frac{\delta_j^{k/\rho} k}{\lambda^k} \Gamma\left(1 - \frac{k}{\rho}\right) \Gamma\left(\frac{k}{\rho}\right) \exp\left(-\frac{k}{\rho} \mathbf{x}'_{ij} \boldsymbol{\beta} + \frac{k^2}{2\rho^2} \mathbf{z}'_{ij} D \mathbf{z}_{ij}\right)$$

- **Weibull**  $\longrightarrow$  **Exponential**

$$E(Y_{ij}^k) = \frac{\delta_j^k k}{\lambda^k} \Gamma(1 - k) \Gamma(k) \exp\left(-k \mathbf{x}'_{ij} \boldsymbol{\beta} + \frac{k^2}{2} \mathbf{z}'_{ij} D \mathbf{z}_{ij}\right)$$

- **Univariate case, Weibull-exponential:**

$$f(y) = \frac{\varphi \rho y^{\rho-1} \delta}{(\delta + \varphi y^\rho)^2}$$

$$E(Y^k) = \frac{k}{\rho} \left( \frac{\delta}{\varphi} \right)^{k\rho} \cdot \Gamma(1 - k/\rho) \cdot \Gamma(k/\rho)$$

- **Univariate case, exponential-exponential:**

$$f(y) = \frac{\varphi \delta}{(\delta + \varphi y)^2}$$

$$E(Y^k) = k \left( \frac{\delta}{\varphi} \right)^k \cdot \Gamma(1 - k) \cdot \Gamma(k)$$

- This leads to a class of **“Cauchy-type”** distributions, including the log-logistic distribution, of which there are known to be problems.
- **Estimation using the marginal-conditional approach:**

$$f(y_{ij}|\mathbf{b}_i) = \frac{\lambda \kappa_{ij} e^{\mu_{ij}} \rho y_{ij}^{\rho-1} \alpha_j \beta_j}{(1 + \lambda \kappa_{ij} e^{\mu_{ij}} \beta_j y_{ij}^{\rho})^{\alpha_j+1}}$$

- **Further issues:**
  - ▷ censorship
  - ▷ related functions
  - ▷ non-parametric baseline hazard functions

# Analysis of the Asthma Data

---

- **Model:**

$$\kappa_{ij} = \xi_0 + b_i + \xi_1 T_i$$

- where

- ▷  $T_i$ : indicator for treatment

- ▷  $b_i \sim N(0, d)$

Effect	Parameter	Exponential	Exponential-gamma
		Estimate (s.e.)	Estimate (s.e.)
Intercept	$\xi_0$	-3.3709 (0.0772)	-3.9782 (15.354)
Treatment effect	$\xi_1$	-0.0726 (0.0475)	-0.0755 (0.0605)
Shape parameter	$\lambda$	0.8140 (0.0149)	1.0490 (16.106)
Std. dev random effect	$\sqrt{d}$	—	—
Gamma parameter	$\gamma$	—	3.3192 (0.3885)
–2log-likelihood		18,693	18,715

Effect	Parameter	Exponential-normal	Combined
		Estimate (s.e.)	Estimate (s.e.)
Intercept	$\xi_0$	-3.8095 (0.1028)	3.9923 (20.337)
Treatment effect	$\xi_1$	-0.0825 (0.0731)	-0.0887 (0.0842)
Shape parameter	$\lambda$	0.8882 (0.0180)	0.8130 (16.535)
Std. dev random effect	$\sqrt{d}$	0.4097 (0.0386)	0.4720 (0.0416)
Gamma parameter	$\gamma$	—	6.8414 (1.7146)
–2log-likelihood		18,611	18,629

# The Asthma Data: Wald Tests

---

<b>Model</b>	<b><i>Z</i> value</b>	<b><i>p</i> value</b>
<b>Exponential</b>	-1.5283	0.1264
<b>Exponential-gamma</b>	-1.1293	0.2588
<b>Exponential-normal</b>	-1.2480	0.2120
<b>Combined</b>	-1.0534	0.2921

# The Binary Case

---

- **A logit formulation:**

$$Y_{ij} \sim \text{Bernoulli}(\pi_{ij} = \theta_{ij}\kappa_{ij})$$

$$\kappa_{ij} = \frac{\exp(\mathbf{x}'_{ij}\boldsymbol{\beta} + \mathbf{z}'_{ij}\mathbf{b}_i)}{1 + \exp(\mathbf{x}'_{ij}\boldsymbol{\beta} + \mathbf{z}'_{ij}\mathbf{b}_i)}$$

- No closed forms
- Conjugacy  $\longrightarrow$  no **strong** conjugacy
- **Marginal-conditional version** (for NLMIXED implementation)

$$f(y_{ij}|\mathbf{b}_i) = \frac{1}{\alpha_j + \beta_j} \cdot (\kappa_{ij}\alpha_j)^{y_{ij}} \cdot [(1 - \kappa_{ij})\alpha_j + \beta_j]^{1-y_{ij}}$$



- A probit formulation:

$$\kappa_{ij} = \Phi_1(\mathbf{x}'_{ij}\boldsymbol{\beta} + \mathbf{z}'_{ij}\mathbf{b}_i)$$

$$\theta_{ij} \sim \text{Beta}(\alpha, \beta)$$

$$f_{n_i}(\mathbf{y}_i = \mathbf{1}) = \left(\frac{\alpha}{\alpha + \beta}\right)^{n_i} \cdot \Phi_{n_i}(X_i\boldsymbol{\beta}; L_{n_i}^{-1})$$

$$L_{n_i} = I_{n_i} - Z_i (D^{-1} + Z_i'Z_i)^{-1} Z_i'$$

$$E(Y_{ij}) = \frac{\alpha}{\alpha + \beta} \cdot \Phi_1(\mathbf{x}'_{ij}\boldsymbol{\beta}; L_1^{-1}) = \frac{\alpha}{\alpha + \beta} \cdot \Phi_1(|I + Dz_{ij}z'_{ij}|^{-1/2} \mathbf{x}'_{ij}\boldsymbol{\beta})$$

$$\text{Var}(Y_{ij}) = \frac{\alpha}{\alpha + \beta} \cdot \Phi_1(\mathbf{x}'_{ij}\boldsymbol{\beta}; L_1^{-1}) \cdot \left[1 - \frac{\alpha}{\alpha + \beta} \cdot \Phi_1(\mathbf{x}'_{ij}\boldsymbol{\beta}; L_1^{-1})\right]$$

$$\text{Cov}(Y_{ij}, Y_{ik}) = \left( \frac{\alpha}{\alpha + \beta} \right)^2 \cdot \left\{ \Phi_2 \left[ \begin{pmatrix} \mathbf{x}'_{ij} \\ \mathbf{x}'_{ik} \end{pmatrix} \boldsymbol{\beta}, L_{2jk}^{-1} \right] - \Phi_1(\mathbf{x}'_{ij}\boldsymbol{\beta}; L_{1j}^{-1})\Phi_1(\mathbf{x}'_{ik}\boldsymbol{\beta}; L_{1k}^{-1}) \right\}$$

$$L_{2jk} = I_2 - \begin{pmatrix} \mathbf{z}'_{ij} \\ \mathbf{z}'_{ik} \end{pmatrix} \left[ D^{-1} + \begin{pmatrix} \mathbf{z}'_{ij} \\ \mathbf{z}'_{ik} \end{pmatrix} \begin{pmatrix} \mathbf{z}_{ij} & \mathbf{z}_{ik} \end{pmatrix} \right]^{-1} \begin{pmatrix} \mathbf{z}_{ij} & \mathbf{z}_{ik} \end{pmatrix}$$

- The above gives the “success probabilities;”  $\longrightarrow$  all other joint probabilities can be derived in a similar fashion

# Toenail Data: Unaffected Nail Length

---

- **Model:**

$$E(Y_{ij}|T_i, t_j, \boldsymbol{\beta}) = \beta_0 + \beta_1 T_i + \beta_2 t_j + \beta_3 T_i t_j$$

- where:

- ▷  $Y_{ij}$ : unaffected nail length for subject  $i$  at occasion  $j$
- ▷  $t_j$ : time at which the  $j$ th measurement is made
- ▷  $T_i$ : treatment indicator for subject  $i$

Effect	Parameter	Logistic	Beta-binomial
		Estimate (s.e.)	Estimate (s.e.)
Intercept treatment A	$\xi_0$	-0.5571 (0.1090)	17.9714 (1482.6)
Slope treatment A	$\xi_1$	-0.1769 (0.0246)	5.2454 (12970.0)
Intercept treatment B	$\xi_2$	-0.5335 (0.1122)	18.6744 (2077.13)
Slope treatment B	$\xi_3$	-0.2549 (0.0309)	4.7775 (12912.0)
Std. dev random effect	$\sqrt{d}$	—	—
Ratio	$\alpha/\beta$	—	3.6739 (0.2051)
-2log-likelihood		1812	1980

Effect	Parameter	Logistic-normal	Combined
		Estimate (s.e.)	Estimate (s.e.)
Intercept treatment A	$\xi_0$	-1.6299 (0.4354)	-1.6042 (4.0263)
Slope treatment A	$\xi_1$	-0.4042 (0.0460)	-6.4783 (1.4386)
Intercept treatment B	$\xi_2$	-1.7486 (0.4478)	-16.2079 (3.5830)
Slope treatment B	$\xi_3$	-0.5634 (0.0602)	-8.0745 (1.5997)
Std. dev random effect	$\sqrt{d}$	4.0150 (0.3812)	60.8835 (14.2237)
Ratio	$\alpha/\beta$	—	0.2805 (0.0350)
-2log-likelihood		1248	1240

# Concluding Table

Element	notation	continuous	binary	count	time to event	
<b>Standard univariate exponential family</b>						
Model		<b>normal</b>	<b>Bernoulli</b>	<b>Poisson</b>	<b>Exponential</b>	<b>Weibull</b>
Model	$f(y)$	$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2\sigma^2}}$	$\pi^y (1-\pi)^{1-y}$	$\frac{e^{-\lambda}\lambda^y}{y!}$	$\varphi e^{-\varphi y}$	$\varphi\rho y^{\rho-1} e^{-\varphi y^\rho}$
Nat. param	$\eta$	$\mu$	$\ln[\pi/(1-\pi)]$	$\ln \lambda$	$-\varphi$	
Mean function	$\psi(\eta)$	$\eta^2/2$	$\ln[1 + \exp(\eta)]$	$\lambda = \exp(\eta)$	$-\ln(-\eta)$	
Norm. constant	$c(y, \phi)$	$\frac{\ln(2\pi\phi)}{2} - \frac{y^2}{2\phi}$	0	$-\ln y!$	0	
(Over)dispersion	$\phi$	$\sigma^2$	1	1	1	
Mean	$\mu$	$\mu$	$\pi$	$\lambda$	$\varphi^{-1}$	$\varphi^{-1/\rho}\Gamma(\rho^{-1} + 1)$
Variance	$\phi v(\mu)$	$\sigma^2$	$\pi(1-\pi)$	$\lambda$	$\varphi^{-2}$	$\varphi^{-2/\rho} [\Gamma(2\rho^{-1} + 1) - \Gamma(\rho^{-1} + 1)^2]$
<b>Exponential family with conjugate random effects</b>						
Model		<b>normal-normal</b>	<b>Beta-binomial</b>	<b>Negative binomial</b>	<b>Exponential-gamma</b>	<b>Weibull-gamma</b>
Hier. model	$f(y \theta)$	$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-\theta)^2}{2\sigma^2}}$	$\theta^y (1-\theta)^{1-y}$	$\frac{e^{-\theta}\theta^y}{y!}$	$\varphi\theta e^{-\varphi\theta y}$	$\varphi\theta\rho y^{\rho-1} e^{-\varphi\theta y^\rho}$
RE model	$f(\theta)$	$\frac{1}{\sqrt{d}\sqrt{2\pi}} e^{-\frac{(\theta-\mu)^2}{2d}}$	$\frac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{B(\alpha,\beta)}$	$\frac{\theta^{\alpha-1}e^{-\theta/\beta}}{\beta^\alpha\Gamma(\alpha)}$	$\frac{\theta^{\alpha-1}e^{-\theta/\beta}}{\beta^\alpha\Gamma(\alpha)}$	$\frac{\theta^{\alpha-1}e^{-\theta/\beta}}{\beta^\alpha\Gamma(\alpha)}$
Marg. model	$f(y)$	$\frac{1}{\sqrt{\sigma^2+d}\sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2(\sigma^2+d)}}$	$(\alpha+\beta) \frac{\Gamma(\alpha)}{\Gamma(\alpha+y)} \frac{\Gamma(\beta)}{\Gamma(\beta+1-y)}$	$\frac{\Gamma(\alpha+y)}{y!\Gamma(\alpha)} \left(\frac{\beta}{\beta+1}\right)^y \left(\frac{1}{\beta+1}\right)^\alpha$	$\frac{\varphi\alpha\beta}{(1+\varphi\beta y)^{\alpha+1}}$	$\frac{\varphi\rho y^{\rho-1}\alpha\beta}{(1+\varphi\beta y^\rho)^{\alpha+1}}$
	$h(\theta)$	$\theta$	$\ln[\theta/(1-\theta)]$	$\ln(\theta)$	$-\theta$	$-\theta$
	$g(\theta)$	$-\frac{1}{2}\theta^2$	$-\ln(1-\theta)$	$\theta$	$-\ln(\theta)/\varphi$	$-\ln(\theta)/\varphi$
	$\phi$	$\sigma^2$	1	1	$1/\varphi$	$1/\varphi$
	$\gamma$	$1/d$	$\alpha + \beta - 2$	$1/\beta$	$\varphi(\alpha - 1)$	$\varphi(\alpha - 1)$
	$\psi$	$\mu$	$\frac{\alpha-1}{\alpha+\beta-2}$	$\beta(\alpha - 1)$	$[\beta\varphi(\alpha - 1)]^{-1}$	$[\beta\varphi(\alpha - 1)]^{-1}$
	$c(y, \phi)$	$-\frac{1}{2}\phi y^2 - \frac{1}{2}\ln\left(\frac{2\pi}{\phi}\right)$	0	$-\ln(y!)$	$\ln(\varphi)$	$\ln(\varphi\rho y^{\rho-1})$
	$c^*(\gamma, \psi)$	$-\frac{1}{2}\gamma\psi^2 - \frac{1}{2}\ln\left(\frac{2\pi}{\gamma}\right)$	$-\ln B(\gamma\psi + 1, \gamma - \psi\gamma + 1)$	$(1 + \gamma\psi) \ln \gamma - \ln \Gamma(1 + \gamma\psi)$	$\frac{\gamma+\varphi}{\varphi} \ln(\gamma\psi) - \ln \Gamma\left(\frac{\gamma+\varphi}{\varphi}\right)$	$\frac{\gamma+\varphi}{\varphi} \ln(\gamma\psi) - \ln \Gamma\left(\frac{\gamma+\varphi}{\varphi}\right)$
Mean	$E(Y)$	$\mu$	$\frac{\alpha}{\alpha+\beta}$	$\alpha\beta$	$[\varphi(\alpha - 1)\beta]^{-1}$	$\frac{\Gamma(\alpha-\rho^{-1})\Gamma(\rho^{-1}+1)}{(\varphi\beta)^{1/\rho}\Gamma(\alpha)}$
Variance	$\text{Var}(Y)$	$\sigma^2 + d$	$\frac{\alpha\beta}{(\alpha+\beta)^2}$	$\alpha\beta(\beta + 1)$	$\alpha[\varphi^2(\alpha - 1)^2(\alpha - 2)\beta^2]^{-1}$	$\frac{1}{\rho(\varphi\beta)^{2/\rho}\Gamma(\alpha)} \left[ 2\Gamma(\alpha - 2\rho^{-1})\Gamma(2\rho^{-1}) - \frac{\Gamma(\alpha-\rho^{-1})^2\Gamma(\rho^{-1})^2}{\rho\Gamma(\alpha)} \right]$

# References

---

- Molenberghs, G. and Verbeke, G. (2011). On the Weibull-Gamma frailty model, its infinite moments, and its connection to generalized log-logistic, logistic, Cauchy, and extreme-value distributions. *Journal of Statistical Planning and Inference*, **141**, 861–868.
- Molenberghs, G., Verbeke, G., and Demétrio, C. (2007). An extended random-effects approach to modeling repeated, overdispersed count data. *Lifetime Data Analysis*, **13**, 513–531.
- Molenberghs, G., Verbeke, G., Demétrio, C.G.B., and Vieira, A. (2010). A family of generalized linear models for repeated measures with normal and conjugate random effects. *Statistical Science*, **25**, 325–347.
- Vangeneugden, T., Molenberghs, G., Verbeke, G., and Demétrio, C. (2011). Marginal correlation from an extended random-effects model for repeated and overdispersed counts. *Journal of Applied Statistics*, **38**, 215–232.