

# Likelihood-Based Inference for Generalized Linear Mixed Models: Inference with R Package glmm

Christina Knudson, Ph.D.

University of St. Thomas

Symposium for Data Science and Statistics 2020

Co-authors: Charles J. Geyer, Ph.D. and Galin Jones, Ph.D.  
(School of Statistics, University of Minnesota)

## R Package glmm Preview

The R package `glmm` enables all frequentist likelihood-based inference via Monte Carlo likelihood approximation.

- LBI: maximum likelihood, likelihood ratio tests, etc
- Poisson and binomial response, indep. normal random effects
- MCLA uses importance sampling
- Approximates log likelihood and two derivatives
- Parallel computing (bring your own cores)
- Weights (weighted likelihood)

# Extending the Linear Model

The usual assumptions for frequentist linear model inference:

- responses are normally distributed.
- responses are independent.
- responses have equal variance.

# Extending the Linear Model

The usual assumptions for frequentist linear model inference:

- responses are normally distributed.
- responses are independent.
- responses have equal variance.

If observations are non-normal and correlated, try a “generalized linear mixed model” (GLMM), which:

- incorporates random effects to include correlation.
- models the log odds or log mean.

## Salamander Example

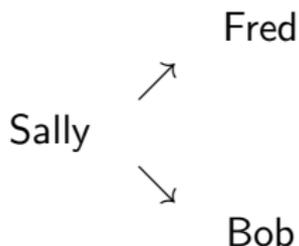


Two salamander populations: Rough Butt (R) and White Side (W)

Do salamanders prefer mating with their own population?

# Salamander Example

Each salamander is used multiple times, has its own random effect to represent its general tendency to mate.



Random effect assumptions:

- Normally distributed with mean 0
- 2 variance components:  $\sigma_M^2$  for males,  $\sigma_F^2$  for females
- Each salamander's random effect is independent of others'

## Salamander Example

What affects the probability that a pair of salamanders mate?

- Type of cross (RR, RW, WR, WW)
- Female's individualized tendency to mate
- Male's individualized tendency to mate

# Salamander Example

What affects the probability that a pair of salamanders mate?

- Type of cross (RR, RW, WR, WW)
- Female's individualized tendency to mate
- Male's individualized tendency to mate

Translation to statistical modeling:

- Response: whether or not the pair mated
- Fixed effects:  $\beta_{RR}, \beta_{RW}, \beta_{WR}, \beta_{WW}$  (log odds of mating)
- Random effects: one per salamander, independent
- Variance components:  $\sigma_F^2$  and  $\sigma_M^2$

# Extending the Linear Model

GLMM predicted response depends on:

- fixed effects
- random effects (individual-specific tendencies)

GLMM parameters:

- fixed effects
- variance components (variance of random effects)

# Inference for GLMMs

Likelihood: a function of the parameters given the observed data.

Likelihood-based inference includes:

- maximum likelihood
- standard errors and covariances of parameter estimates
- confidence intervals
- hypothesis tests (Wald, LRT, etc)

# Inference for GLMMs

Likelihood: a function of the parameters given the observed data.

Likelihood-based inference includes:

- maximum likelihood
- standard errors and covariances of parameter estimates
- confidence intervals
- hypothesis tests (Wald, LRT, etc)

Likelihood-based inference is hard for GLMMs

Likelihood cannot depend on random effects

Likelihood is integral (often high dimension)

# Statistical Theory Underlying glmm

R package `glmm` enables all likelihood-based inference:

- `glmm` approximates entire likelihood using Monte Carlo
- Monte Carlo MLEs converge to MLEs as Monte Carlo sample size increases
- Monte Carlo likelihood approximation converges to likelihood
- All likelihood-based inference converges

## Salamander Example

```
> library(glmm)
```

```
> data(salamander)
```

```
> head(salamander)
```

	Mate	Cross	Female	Male
1	1	R/R	10	10
2	1	R/R	11	14
3	1	R/R	12	11
4	1	R/R	13	13
5	1	R/R	14	12
6	1	R/W	15	28

## Salamander Example: Model Fitting

```
> m <- 10^6
> clust <- makeCluster(6)
> set.seed(1234)
>
> sal <- glmm(Mate~0+Cross,
+   random=list(~0+Female,~0+Male),
+   varcomps.names=c("F","M"), data=salamander,
+   family.glmm=bernoulli.glmm, m=m, cluster=clust)
```

## Salamander Example: Standard Errors

```
> cbind(mcse(sal), se(sal))
              "mcse"      "se"
CrossR/R 0.008261367 0.4061965
CrossR/W 0.008474469 0.3871796
CrossW/R 0.012265425 0.4597051
CrossW/W 0.006240152 0.3911193
F         0.018988405 0.4744714
M         0.025215655 0.5450108
```

Two sources of variability:

- SE: variability from data-set to data-set
- MCSE: variability from MC run to run, holding data constant

If MCSE large compared to SE, increase  $m$  to reduce MCSE.

(Increasing  $m$  will not decrease SE. You need more data to reduce SE.)

## Salamander Example: Summary

Part of the model summary:

Fixed Effects:

	Estimate	Std. Error	z value	Pr(> z )	
CrossR/R	1.011	0.406	2.489	0.0128	*
CrossR/W	0.317	0.387	0.818	0.4135	
CrossW/R	-1.920	0.460	-4.183	2.87e-05	***
CrossW/W	0.983	0.391	2.513	0.0120	*

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

## Salamander Example: Likelihood Ratio Test

Do the odds of mating differ by cross? Use LRT.

```
sal2 <- glmm(Mate~1, random=list(~0+Female,~0+Male),  
+   varcomps.names=c("F","M"), data=salamander,  
+   family.glmm=bernoulli.glmm, m=m, cluster=clust)  
> (teststat <- 2*(logLik(sal)-logLik(sal2)))  
[1] 56.07031  
> pchisq(teststat, lower.tail = FALSE, df = 3)  
[1] 4.058279e-12
```

3 df because one model has 4 betas, other has 1 beta.

P-value is quite small: model with cross is better than model w/o.

The log odds of mating are not the same between all 4 crosses.

# Salamander Example: Wald Tests

Hypothesis tests determine **which** odds differ significantly.

$$\text{(eg. } H_0 : \beta_{RR} = \beta_{WW} \quad H_A : \beta_{RR} \neq \beta_{WW} \text{)}$$

Use `vcov` function for (co)variances to calculate Wald test statistic:

$$\frac{\hat{\beta}_{RR} - \hat{\beta}_{WW} - 0}{\sqrt{\text{Var}(\hat{\beta}_{RR}) + \text{Var}(\hat{\beta}_{WW}) - 2\text{Cov}(\hat{\beta}_{RR}, \hat{\beta}_{WW})}} \sim N(0, 1).$$

## Salamander Example: Summary

More of the model summary:

Variance Components for Random Effects

(P-values are one-tailed):

	Estimate	Std. Error	z value	Pr(> z )/2	
F	1.31	0.47	2.768	0.00282	**
M	1.22	0.55	2.234	0.01276	*

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

The variances are significantly greater than 0: there's definitely variability from salamander to salamander.

## Salamander Example: Confidence Intervals

```
> confint(sal, level = .9)
              0.05      0.95
CrossR/R  0.3429750  1.6792426
CrossR/W -0.3202513  0.9534561
CrossW/R -2.6793000 -1.1670047
CrossW/W  0.3393784  1.6260465
F          0.5328952  2.0937671
M          0.3208338  2.1137599
```

# R package glmm Summary

- Enables all frequentist likelihood-based inference.
- Computes and maximizes a Monte Carlo likelihood approximation.
- Calculates the MCLA and two derivatives at the MCMLEs.
- Computes in parallel (if you have cores).
- Can weight observations (weighted likelihood).

Questions?

knud8583@stthomas.edu  
github.com/knudson1  
cknudson.com

## Related to MCLA and glmm

Geyer C. (1990). *Likelihood and Exponential Families*. PhD thesis, University of Washington.

Geyer C.J. (1994). "On the Convergence of Monte Carlo Maximum Likelihood Calculations." *Journal of the Royal Statistical Society, Series B*, 61, 261-274.

Geyer C.J., Thompson E. (1992). "Constrained Monte Carlo Maximum Likelihood for Dependent Data." *Journal of the Royal Statistical Society, Series B*, 54, 657-699.

Knudson C. (2015). *glmm: Generalized Linear Mixed Models via Monte Carlo Likelihood Approximation*. R package version 1.0.2, URL <http://CRAN.R-project.org/package=glmm>.

Knudson C. (2016). *Monte Carlo Likelihood Approximation for Generalized Linear Mixed Models*. Ph.D. Thesis, University of Minnesota.

Sung Y.J., Geyer C.J. (2007). "Monte Carlo Likelihood Inference for Missing Data Models." *Annals of Statistics*, 35, 990-1011.

# Inference for GLMMs with `glmm`

R package `glmm`

- 1 Based on data, selects importance sampling distribution  $\tilde{f}(u)$
- 2 Generates  $m$  random effects from  $\tilde{f}(u)$
- 3 Calculates and maximizes MCLA using `trust`
- 4 Returns
  - Monte Carlo MLEs
  - MCLA value, gradient and Hessian at MCMLEs
  - Lots of other info (`trust` output, etc)

Families currently allowed: Binomial and Poisson

Random effect structure currently allowed: independent normals

## Salamander Example: Model Fitting

```
> m <- 10^6
> clust <- makeCluster(6)
> set.seed(1234)
> ptm <- proc.time()
> sal <- glmm(Mate~0+Cross, random=list(~0+Female,~0+Male),
  varcomps.names=c("F","M"), data=salamander,
  family.glmm=bernoulli.glmm, m=m, cluster=clust)
> proc.time()-ptm
  user  system elapsed
69.751  11.037 2783.557
```

(About 45 minutes)

## Salamander Example: Model Fitting

Fitting the nested model:

```
> ptm <- proc.time()
> sal2 <- glmm(Mate~1, random=list(~0+Female,~0+Male),
  varcomps.names=c("F","M"), data=salamander,
  family.glmm=bernoulli.glmm, m=m, cluster=clust)
> proc.time()-ptm
  user    system elapsed
66.843   10.124 1459.746
```

(About 24 minutes)

# Salamander Example

Back to probabilities:

$$P(\text{mating}) = \frac{\exp(\hat{\beta}_{RW})}{1 + \exp(\hat{\beta}_{RW})}$$

Cross	RR	WW	RW	WR
Probability of mating	0.73	0.73	0.58	0.13

Which probabilities are significantly different?

# Salamander Example

Hypothesis tests determine which probabilities differ significantly.

$$H_0 : \beta_{RR} = \beta_{WW}$$

$$H_A : \beta_{RR} \neq \beta_{WW}$$

First, use `vcov` function for (co)variances needed to calculate

$$\text{Var}(\hat{\beta}_{RR} - \hat{\beta}_{WW}) = \text{Var}(\hat{\beta}_{RR}) + \text{Var}(\hat{\beta}_{WW}) - 2\text{Cov}(\hat{\beta}_{RR}, \hat{\beta}_{WW}).$$

Then a Wald test statistic is

$$\frac{\hat{\beta}_{RR} - \hat{\beta}_{WW} - 0}{\sqrt{\text{Var}(\hat{\beta}_{RR} - \hat{\beta}_{WW})}} \sim N(0, 1).$$

# Salamander Example

```
> pvals
CrossR/R vs CrossR/W CrossR/R vs CrossW/R
      1.467231e-01      1.792076e-07
CrossR/W vs CrossW/R CrossR/W vs CrossW/W
      2.673051e-04      1.615877e-01
CrossR/R vs CrossW/W CrossW/R vs CrossW/W
      9.597193e-01      8.331821e-08
```

## Salamander Example

How do point estimates compare?

	$\hat{\beta}_{RR}$	$\hat{\beta}_{RW}$	$\hat{\beta}_{WR}$	$\hat{\beta}_{WW}$	$\hat{\nu}_F$	$\hat{\nu}_M$
glmm ( $m = 10^6$ )	1.0	.3	-1.9	1.0	1.3	1.2
MCEM	1.0	.3	-2.0	1.0	1.4	1.3
lme4 (PQL)	1.0	.3	-1.9	1.0	1.2	1.0

Monte Carlo EM finds Monte Carlo MLEs.

lme4 maximizes a penalized quasi-likelihood.

# Comparing R Packages

## Comparing `glmm` and `lme4`

- `lme4` much faster (penalized-quasi likelihood v. Monte Carlo)
- `lme4` performs maximum likelihood for simple models (one random effect per observation)
- `glmm` performs/enables all likelihood-based inference
- `glmm` inference converges as  $m \uparrow$
- `lme4` variance components are too small
- Impossible to know how close `lme4`'s PQL matches likelihood
- `glmm` random effect structure limited to independent normals

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.1-6.

Breslow, N. and Clayton, D. (1993). *Approximate inference in generalized linear mixed models*. Journal of the American Statistical Association, 88:9-25.

Breslow, N. and Lin, X. (1995). *Bias correction in generalized linear mixed models with a single component of dispersion*. Biometrika, 82:81-91.

Lin, X. and Breslow, N. (1996). *Bias correction in generalized linear mixed models with multiple components of dispersion*. Journal of the American Statistical Association, 91:1007-1016.