

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variable | Example | Type of Regression | R function / R function for mixed models | More Information | Bayesian Equivalent (w/ `brms`) |
| Continuous | Quality of Life, linear scales | linear | `lm()` |  | `brm(family = gaussian())` |
| - `lmer()` - `glmmTMB()` |  |
| Binary | Success yes/no | binary logistic | `glm(family=binomial)` | [UCLA](https://stats.idre.ucla.edu/r/dae/logit-regression/) | `brm(family = binomial())` |
| - `glmer(family=binomial)` - `glmmTMB(family=binomial)` |
| Trials (or proportions of \_counts\_) | 20 successes out of 30 trials | logistic | `glm(cbind(successes, failures), family=binomial)` | [Hadley’s notes](http://had.co.nz/notes/modelling/logistic-regression.html) | `brm(successes | trials(total), family = binomial())` |
| - `glmer(cbind(successes, failures), family=binomial)` - `glmmTMB(cbind(successes, failures), family=binomial)` |
| Count data | Number of usage, counts of events | Poisson | `glm(family=poisson)` | [UCLA](https://stats.idre.ucla.edu/r/dae/poisson-regression/) | `brm(family = poisson())` |
| - `glmer(family=poisson)` - `glmmTMB(family=poisson)` |
| Count data, with excess zeros or overdispersion | Number of usage, counts of events (with higher variance than mean of response) | negative binomial | `glm.nb()` | [UCLA](https://stats.idre.ucla.edu/r/dae/negative-binomial-regression/) | `brm(family = negbinomial())` |
| - `glmer.nb()` - `glmmTMB(family=nbinom)` |
| Count data with very many zeros (inflation) | see count data, but response is modelled as mixture of Bernoulli & Poisson distribution (two sources of zeros) | zero-inflated | `zeroinfl()` | [UCLA](https://stats.idre.ucla.edu/r/dae/zip/) | `brm(family = zero\_inflated\_poisson())` |
| `glmmTMB(ziformula, family=poisson)` |
| Count data, with very many zeros (inflation) and overdispersion | Number of usage, counts of events (with higher variance than mean of response) | zero-inflated negative binomial | `zeroinfl(dist="negbin")` | [UCLA](https://stats.idre.ucla.edu/r/dae/zinb/) | `brm(family = zero\_inflated\_negbinomial())` |
| `glmmTMB(ziformula, family=nbinom)` |
| Count data, zero-truncated | see count data, but only for positive counts (hurdle component models zero-counts) | hurdle (Poisson) | `hurdle()` | [UCLA](https://stats.idre.ucla.edu/r/dae/zero-truncated-poisson/) | `brm(family = hurdle\_poisson())` |
| `glmmTMB(family=truncated\_poisson)` |
| Count data, zero-truncated and overdispersion | see “Count data, zero-truncated”, but with higher variance than mean of response | hurdle (neg. binomial) | `vglm(family=posnegbinomial)` | [UCLA](https://stats.idre.ucla.edu/r/dae/zero-truncated-negative-binomial/) | `brm(family = hurdle\_negbinomial())` |
| `glmmTMB(family=truncated\_nbinom)` |
| Proportion / Ratio (without zero and one) | Percentages, proportion of \_continuous\_ data | Beta \_(see note below)\_ | `betareg()` | [ouR data generation]( https://www.rdatagen.net/post/binary-beta-beta-binomial/) | `brm(family = Beta())` |
| `glmmTMB(family=beta\_family)` |  |
| Proportion / Ratio (including zero and one) | Percentages, proportions of \_continuous\_ data | Beta-Binomial, zero-inflated Beta, ordered Beta \_(see note below)\_ | `BBreg()`, `betabin()`, `vglm(family=betabinomial)`, `ordbetareg()` | [ouR data generation]( https://www.rdatagen.net/post/binary-beta-beta-binomial/) | `brm(family = zero\_one\_inflated\_beta())` |
| `glmmTMB(ziformula, family=beta\_family)`, `glmmTMB(ziformula, family= betabinomial)`, `glmmTMB(ziformula, family= ordbeta)`, `ordbetareg()` |
| Ordinal | Likert scale, worse/ok/better | ordinal, proportional odds, cumulative | `polr()`, `clm()`, `bracl()` | [UCLA](https://stats.idre.ucla.edu/r/dae/ordinal-logistic-regression/) | `brm(family = cumulative())` |
| `clmm()`, `mixor()`, `MCMCglmm(family = "ordinal")` |
| Multinomial | No natural order of categories, like red/green/blue | multinomial | `multinom()`, `brmultinom()` | [UCLA](https://stats.idre.ucla.edu/r/dae/multinomial-logistic-regression/) | `brm(family = multinomial())` |
| `MCMCglmm(family = "multinomial")` |
| Continuous, right-skewed | Financial data, reaction times | Gamma | `glm(family=Gamma)` | [Sean Anderson]( https://seananderson.ca/2014/04/08/gamma-glms/) | `brm(family = Gamma())`,  but see also [Reaction time distributions in `brms`](https://lindeloev.github.io/shiny-rt/) |
| `glmer()`, `glmmTMB()` |  |
| (Semi-)Continuous, (right) skewed, probably with spike at zero (zero-inlfated) | Financial data, probably exponential dispersion of variance | Tweedie | `glm(family=tweedie)`, `cpglm()` | [Revolutions](https://blog.revolutionanalytics.com/2014/10/a-note-on-tweedie.html) |  |
| `cpglmm()`, `glmmTMB(family=tweedie)` |
| (Semi-)Continuous, (right) skewed, probably with spike at zero (zero-inlfated) | Normal distribution, but negative values are censored and stacked on zero | Tobit | `tobit()`, `censReg()` |  | `brm(y | cens(), family = gaussian())` |
| `semLme()` |
| Continuous, but truncated or outliers |  | truncated | `censReg()`, `tobit()`, `vglm(family=tobit)` | [UCLA-1](https://stats.idre.ucla.edu/r/dae/tobit-models/), [UCLA-2](https://stats.idre.ucla.edu/r/dae/truncated-regression/) | `brm(y | trunc(), family = gaussian())` |
| Continuous, but exponential growth |  | log-transformed, non-linear | `glm(family=Gaussian("log")`, `nls()` | [Some useful equations](https://www.statforbiology.com/nonlinearregression/usefulequations), [linear vs. non-linear regression](https://stats.stackexchange.com/a/61806/54740) |  |
| `glmmTMB(family=Gaussian("log"))`, `nlmer()`, `nlme()` |
| Proportion / Ratio with more than 2 categories | Biomass partitioning in plants (ratio of leaf, stem and root mass) | Dirichlet | `DirichReg()` |  | `brm(family = dirichlet())` |
| Time-to-Event | Survival-analysis, time until event/death occurs | Cox (proportional hazards) | `coxph` | [UCLA]( https://stats.idre.ucla.edu/r/dae/mixed-effects-cox-regression/) | `brm(family = cox())` |
| `coxme()` |