# Fitting a Gamma-Gompertz survival model to capture-recapture data collected on free-ranging animal populations

**Gilbert Marzolin**[1] **and Olivier Gimenez**[1]

**1** Centre National de la Recherche Scientifique

## Introduction

`GammaGompertzCR` is an R package (R Core Team 2016) that allows estimating survival in free-ranging animal populations using a Gompertz capture-recapture model with a Gamma frailty to deal with individual heterogeneity (Marzolin, Charmantier, and Gimenez 2011; Missov 2013). To use the package, users should be familiar with Bayesian MCMC techniques and in particular how to interpret convergence diagnostics. We refer to Robert and Casella (2010) for an introduction to Bayesian MCMC techniques with R and to King et al. (2009) for an introduction for ecologists. In this paper, we introduce the theory underlying the model we implement in `GammaGompertzCR`, and illustrate the approach using a real example.

## Theory

### The Gamma-Gompertz model

We consider a random variable $T \geq 0$ called time to event. When the event is death of some organism, $T$ is usually associated with a survivor function $S$ such as, for $t \geq 0$, $S(t) = P(T > t)$. Denoting $f$ the pdf of $T$, we have $S(t) = \int_t^{+\infty} f(t)dt$ or $f(t) = -dS(t)/dt$. The hazard function or mortality rate $h$ yields, for $t \geq 0$, the rate of death $h(t)$ given the animal survived up to time $t$, that is $h(t) = f(t)/S(t)$ or $h(t) = -d \log S(t)/dt$. In case of discrete time steps - for instance age $j$ in years - annual survival $s$ by age is obtained as $s(j, j+1) = P(T > j + 1 | T > j) = S(j+1)/S(j)$.

To deal with individual heterogeneity in survival, we can use individual random-effect models (Marzolin, Charmantier, and Gimenez 2011) or multiply the baseline mortality rate $h$ by a unit-specific random variable $u$ named frailty. When $h$ is Gompertz with $h(t) = ae^{bt}$ and $u$ is distributed as a $\Gamma(k, \lambda)$ (with mean $k/\lambda$), we have a multiplicative Gamma-Gompertz frailty model. Typically a $\Gamma$ with mean one is adopted, hence $k = \lambda$ and variance is $1/k$. In a Gamma-Gompertz model, the population survival function is obtained through marginalization. When two parameters are used, $S(t) = (1 + (a/b\lambda)(e^{bt} - 1))^{-k}$ (Missov 2013), hence the hazard equals $h(t) = a(k/\lambda)e^{bt}/(1 + a(e^{bt} - 1)/b\lambda)$. In our case, we use $\lambda = k$. For the detection probability, we used a yearly random effect normally distributed with mean $\psi$ and standard deviation $\sigma_\eta$ on the logit scale.

### Parameter estimation

To get maximum likelihood parameter estimates, we use data cloning in the Bayesian framework (Lele, Dennis, and Lutscher 2007) implemented in the R package `dclone` (Solymos 2010). Data cloning uses multiple copies of the data to produce prior-invariant inferences and converge towards a normal distribution centered at the maximum likelihood estimates. In addition, this method allows detecting non-identifiable parameters (Lele, Nadeem, and Schmuland 2010).

## Illustration

We now illustrate the estimation of Gamma-Gompertz model parameters using data cloning. We analysed capture-recapture data collected on 75 breeding females over 9 years in a Dipper (*Cinclus cinclus*) population (Marzolin, Charmantier, and Gimenez 2011). These data are just a subset of the complete dataset (>1000 individuals, 35 years) that is provided with the package. We used 1, 10, 50 and 100 clones with 3 parallel MCMC chains for a total of 5000 iterates with a burn-in period of 1000. The model fitting took less than 5 minutes.

Parameter estimates are provided in Table 1.

| Parameter | Mean | SD | Credible interval |
|:---:|:---:|:---:|:---:|
| $a$ | 0.12 | 0.02 | [0.09; 0.15] |
| $b$ | 0.48 | 0.03 | [0.43; 0.54] |
| $k$ | 4.90 | 0.03 | [4.84; 4.96] |
| $\psi$ | 1.40 | 0.03 | [1.34; 1.46] |
| $\sigma_\eta$ | 0.67 | 0.02 | [0.63; 0.71] |

Table 1: Parameter estimates of the Gamma-Gompertz model using the Dipper dataset. We refer to Theory section for more details about these parameters. We provide the posterior mean and standard deviation as well as the 95% credible interval.

We also provide a plot of the relationship between estimated survival and age in Figure 1.

More details on how to use the package, including conducting convergence diagnostics, performing a sensitivity analysis and checking parameter identifiability, are provided at the package development repository (see next section).

## Availability

The latest version of the package is available at 'https://github.com/oliviergimenez/GammaGompertzCR'.

## References

King, R., O. Gimenez, B. Morgan, and S. Brooks. 2009. *Bayesian Analysis for Population Ecology.* Chapman & Hall/CRC.
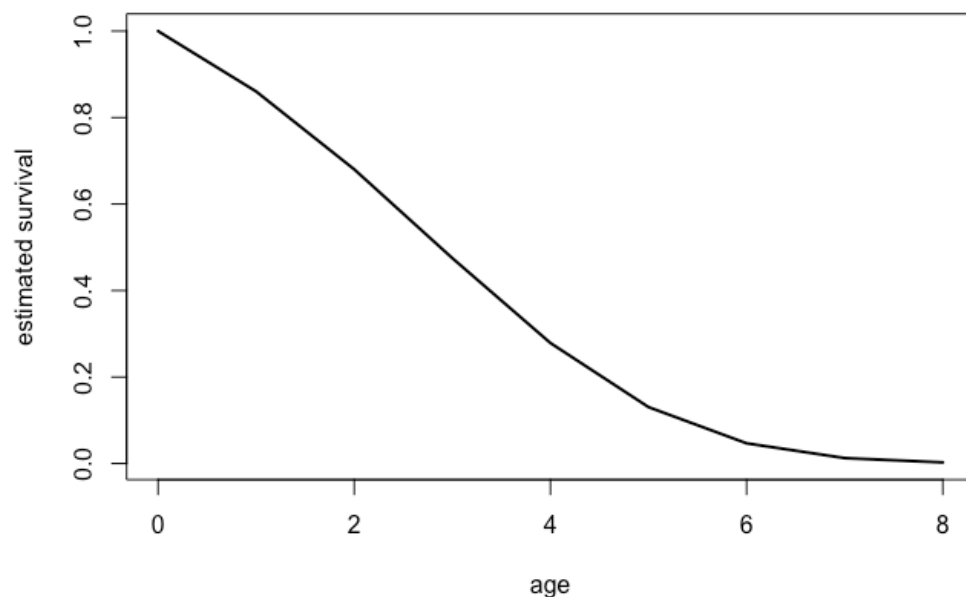
**Figure 1:** The relationship between Dipper survival and age as estimated by the Gamma-Gompertz model.

Lele, SR, D Dennis, and F Lutscher. 2007. "Data Cloning: Easy Maximum Likelihood Estimation for Complex Ecological Models Using Bayesian Markov Chain Monte Carlo Methods." *Ecology Letters* 10:551–63.

Lele, SR, K Nadeem, and B Schmuland. 2010. "Estimability and Likelihood Inference for Generalized Linear Mixed Models Using Data Cloning." *Journal of the American Statistical Association* 105:1617–25.

Marzolin, G, A Charmantier, and O Gimenez. 2011. "Frailty in State-Space Models: Application to Actuarial Senescence in the Dipper." *Ecology* 92:562–67.

Missov, TI. 2013. "Gamma-Gompertz Life Expectancy at Birth." *Demographic Research* 28:259–70.

R Core Team. 2016. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/.

Robert, C., and G. Casella. 2010. *Introducing Monte Carlo Methods with R.* Use R! Springer.

Solymos, P. 2010. "Dclone: Data Cloning in R." *The R Journal* 2:29–37.