

Q & A - Lecture 4

Nicolò Margaritella

March 22, 2020

Parameter interpretation in slide 68

Consider the analysis of the relationship between the number of satellite crabs (response variable) and the covariates “Width of the Carapace” and “Carapace Colour”. We set the following generalised linear model with a log-link connecting the response mean with the linear combination of the covariates:

$$\begin{aligned}\log(\mu_i) &= \beta_0 + \beta_1 \widetilde{\text{Width}}_i + \beta_2 \text{Li.Med}_i + \beta_3 \text{Med}_i + \beta_4 \text{dark.Med}_i, \\ \mu_i &= \exp\{\beta_0\} \exp\{\beta_1 \widetilde{\text{Width}}_i\} \exp\{\beta_2 \text{Li.Med}_i\} \exp\{\beta_3 \text{Med}_i\} \exp\{\beta_4 \text{dark.Med}_i\},\end{aligned}\tag{1}$$

where $\widetilde{\text{Width}}_i$ represents the centred covariate. The second line of Equation (1) highlights the multiplicative effect of the covariates on the mean μ_i . After running JAGS, we want to interpret our results.

UPDATED: Plug-in the posterior means obtained in slide 68 like shown here:

$$\begin{aligned}\mu_i &= \exp\{0.816\} \exp\{0.544 \times \widetilde{\text{Width}}_i\} \exp\{0.455 \times \text{Li.Med}_i\} \\ &\quad \exp\{0.258 \times \text{Med}_i\} \exp\{0.01 \times \text{dark.Med}_i\}\end{aligned}\tag{2}$$

is NOT RECOMMENDED. This is because we are approximating $E(\exp(\mu_i))$ with $\exp(E(\mu_i))$ and we don’t know how good this approximation is going to be!

The best approach is to use the posterior samples of each parameters obtained with JAGS and compute $E(\exp(\beta_j))$ in just one line of R-code:

```
fit.crabs<- as.data.frame( combine.mcmc( results.crabs.mult.B) )
mean( exp( fit.crabs$beta0 ) )# = 2.30
...
mean( exp( fit.crabs$beta.dk.med ) )# = 1.02
```

Now we have the correct results and we can interpret them as follows:

- $E[\exp\{\beta_0\}] = 2.30$ is the expected number of satellite crabs for a female with a dark carapace of average width (i.e. $\widetilde{\text{Width}}_i = 0$ and $\text{Li.Med}=\text{Med}=\text{dark.Med}=0$)
- A female crab with a carapace width one unit larger than the mean ($\widetilde{\text{Width}}$) has an expected number of satellite crabs increased of $E[\exp\{\beta_1\}] = 1.73$ times, *while the other variables are held constant*. Equivalently, for a one unit increase in $\widetilde{\text{Width}}_i$ there is a 73% increase in the expected number of satellites *while the other variables are held constant*.

- A female crab with a light medium coloured carapace attracts on average $E[\exp\{\beta_2\}] = 1.60$ times more satellite crabs than a female with dark coloured carapace, *while the other variables are held constant*. Equivalently, there is a 60% increase in the expected number of satellite crabs attracted by a female crab with light medium carapace than one with dark carapace *while the other variables are held constant*. We can interpret the other parameters in the same way.

Posterior distribution of μ_i

To avoid approximations and generate credible intervals, we can obtain the posterior distribution of μ_i conditional on **any** possible value of the covariates, from the posterior distribution of the relative parameters β_j with $j = \{0, 1, 2, 3, 4\}$. For example, the posterior mean and 95% credible interval for μ | Width = 2, Li.Med = 1 is obtained as follows:

```
post_mu=exp( fit.crabs$beta0+
fit.crabs$beta.width*2 + fit.crabs$beta.lt.med*1)
mean(post_mu)
# 10.74
quantile(post_mu,c(0.025,0.975))
#      2.5%      97.5%
# 7.286524 15.043891
```

One question was if we could obtain the same result by monitoring the parameter μ_i within JAGS (by adding “mu” in the *variable.names* arguments of the *coda.samples* function in slide 67). Two comments about this:

- Monitoring μ_i in JAGS means storing $i \times S$ additional MCMC samples, where S is the number of MCMC samples saved after burn-in. These additional data increase the size of *results.crabs.mult.B* from 1 MB to 41 MB (not very efficient for $i \gg 0$)
- By monitoring μ_i in JAGS we obtain the posterior distributions relative to *only those combinations of covariates observed in our sample*. In the example above, the combination (Width = 2, Li.Med = 1) does not appear in any of the crabs observed in our dataset but we can still make inference on that!

These comments are valid also for the case of the posterior predictive distributions in slide 72.