Interactions and chains

Richard Torkar

April 2021

INTERACTIONS are key to understand how predictors interact with each other. One issue, which is seldom discussed, is that for you to use interactions you'd need a much larger sample size to discover an effect the same size as when you look at population-level effects (i.e., β parameters).

Nevertheless, interactions (a.k.a. moderation) is common in analyses so we should understand them. But what is there to understand? Well, first of all, one should understand the concept of survivorship bias, which is about *what* you condition *on*. The reason you must know this bias well is that you sooner or later might face it. If you condition on the wrong thing, survival in this case, everything will be wrong in the end, i.e., the famous shit-in-shit-out principle.

Here are a few things you must know about interactions. First, they are not always easy. Second, how to define it in a model specification (pp. 248) and how to code it in the design using `rethinking`— this implies both categorical (Sect. 8.1) and continuous interactions (Sect. 8.3). Third, and final, how to plot interactions; preferably using *triptych* plots (pp. 258). If we want to plot priors only we can do that as usual, i.e., `prior <- extract.prior(m)` and then use `post=prior` as an argument to the `link()` function.

MARKOV CHAIN MONTE CARLO was invented by Stanisław Ulam and Ed Teller during the Manhattan project. Soon other extensions came and there are in particular two you should know about: Metropolis and Hamiltonian Monte Carlo (HMC).

You must understand how the Metropolis algorithm is constructed (four steps, see pp. 265). When you understand it you will easier understand HMC (Sect. 9.3). After that it's quite easy. Instead of using `quap()` to sample, you use `ulam()`. The other tools are the same and you specify the models in the same way as always.

Go through the example in Sect. 9.4 and pay particular attention to using pairs plots, trace plots, and trank plots. They indicate if the chains have mixed. Additionally, the following diagnostics you must know,

- Effective sample size. A good rule of thumb is 10% of the total sample size, i.e., 4 chains running for 1000 iterations leaves us with 4000 samples, of which half are thrown away as warmup; thus 200 is a minimum in this case.

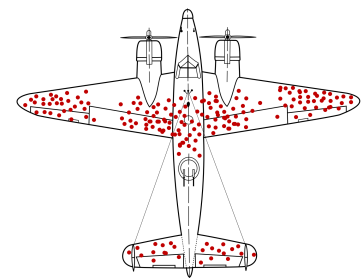According to Andrew Gelman one would need 16 times the sample size... `shorturl.at/dDKQS`

Figure 1: A plane that came back from a mission looking like this indicates that the bullets it caught didn't matter. (`https://bit.ly/3wwW9in`)

Figure 2: Dr. Arianna Wright Rosenbluth (1927–2020). Dr. Rosenbluth implemented the Metropolis algorithm for the first time in 1953. She received her Ph.D. in 1949, making her the fifth woman to earn a physics Ph.D. from Harvard since it began awarding them in 1873.

- Divergent transitions. What are they, what do they indicate and what can you do about them?

- $\widehat{R} < 1.01$. What it is and what happens when it's above 1.01? What can you do to fix this?