Introduction to Bayesian MCMC Techniques: A Practical Guide for the Lay-Actuary

Spencer Miller & Kenny Smart

2025-06-12

Abstract

To be completed later.

Introduction

Some text

Target Audience

- An actuary who has finished exams a few years ago
- An actuary who constantly finds their estimates to be out of line with results
- An actuary interested in learning about Stan but finds the math too onerous or intimidating
- An actuary with limited data

Current Theories on Prediction

- Point estimates do not provide for an adequate measure of process / parameter / model risk
- Simple MCMC models layer in process risk but still fail to account for parameter risk / model risk
- Using a copula for parameters ignores the likelihood of occurring given the data

Yes, Even Actuaries are Human

When making predictions, people often find themselves using heuristics. As outlined by Amos Tversky and Daniel Kahneman in their ground-breaking paper, there are several factors at play when making predictions:

- *Insensitivity to sample size*: An actuary may predict some future value based on limited data.
- *Insensitivity to prior probability of outcomes*: Example?
- *Misconception of chance*: Say something about process risk or that an actuary may infer some sort of pattern where one does not exist.
- *Insensitivity to predictive accuracy*: Taking a risk manager's assessment of a new claims handling process at face value.

- *The illusion of validity*: Using an exposure base as a proxy for loss?
- *Misconceptions of Regression*: Letting one bad AY influence the expectation of the next AY.
- *Anchoring*: Using the prior prediction.

Basics of Bayesian MCMC

Need an introductory sentence

- A **Markov process** is a random process for which the future (the next step) depends only on the present state; it has no memory of how the present state was reached.
- **Monte Carlo** simulation uses random sampling and statistical modeling to estimate mathematical functions and mimic the operations of complex systems.
- **Bayesian statistics** is an approach to data analysis based on Bayes' theorem, where available knowledge about parameters in a statistical model is updated with the information in observed data.
- This process is done using Stan , an open-source software, via common coding languages (e.g., python and R). In our example, we leverage the RStan package, though there are numerous other packages that can be used.

A Simple Walkthrough

Kenny to provide example. Ideally this is a step-by-step with graphics so that's super clear what is going on. We can gloss over the more complex stuff.

An Example in Action

For a simple example, we'll use the procedure to estimate the annual number of commercial fire claims using data compiled by the F'ed'eration Francaise des Soci'et'e d'Assurance. The data includes 9,613 fires occurring between 1982 and 1996. For simplicity, we'll assume a 0% frequency trend, the exposure base (commercial buildings) are level over the time period, and we are tasked with estimating the number of fires in 1993.

First, let's get the data we'll be using in our example:

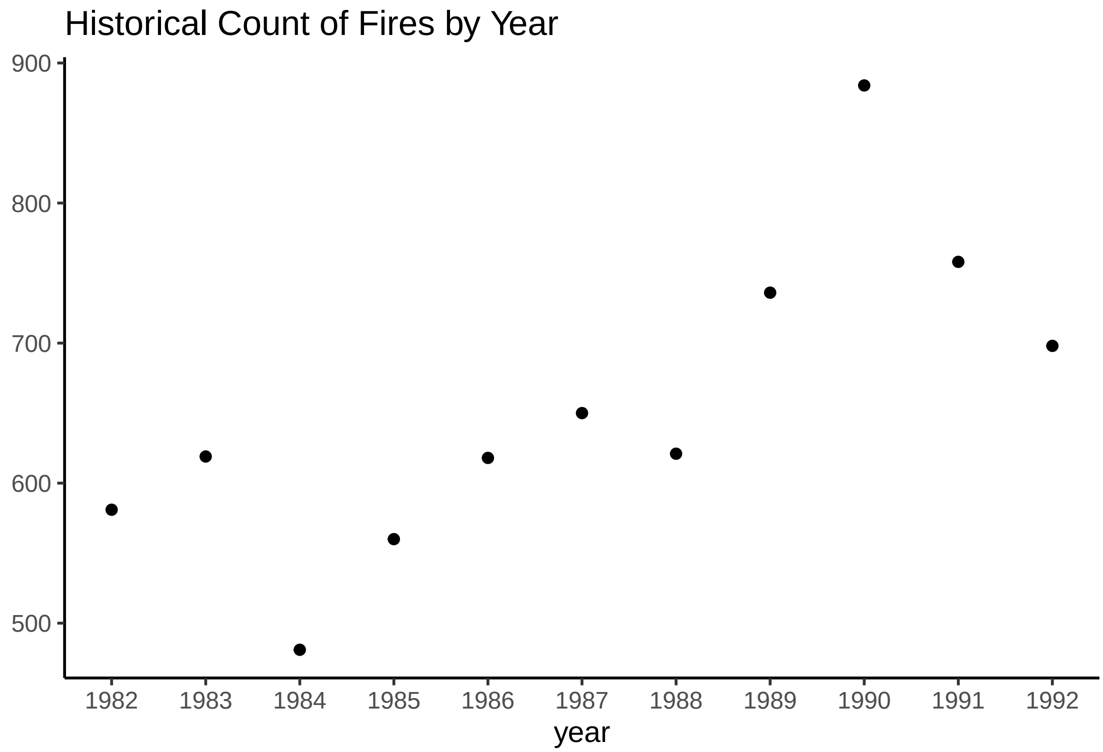
```r
data("frecomfire", package = 'CASdatasets')

data.table::setDT(x = frecomfire)

dplyr::glimpse(frecomfire)
```

```
## Rows: 9,613
## Columns: 4
## $ Year          <dbl> 1982, 1982, 1982, 1982, 1982, 1982, 1982, 1982, 1982, 19~
## $ OccurDate     <date> 1982-01-01, 1982-01-01, 1982-01-01, 1982-01-01, 1982-01~
## $ ClaimCost     <dbl> 2756559, 5408432, 8092851, 7517541, 2606663, 6697463, 84~
## $ ClaimCost2007 <dbl> 821.4924, 1611.7868, 2411.7805, 2240.3303, 776.8212, 199~
```

Next, we'll get the count data by year between 1982 and 1992.

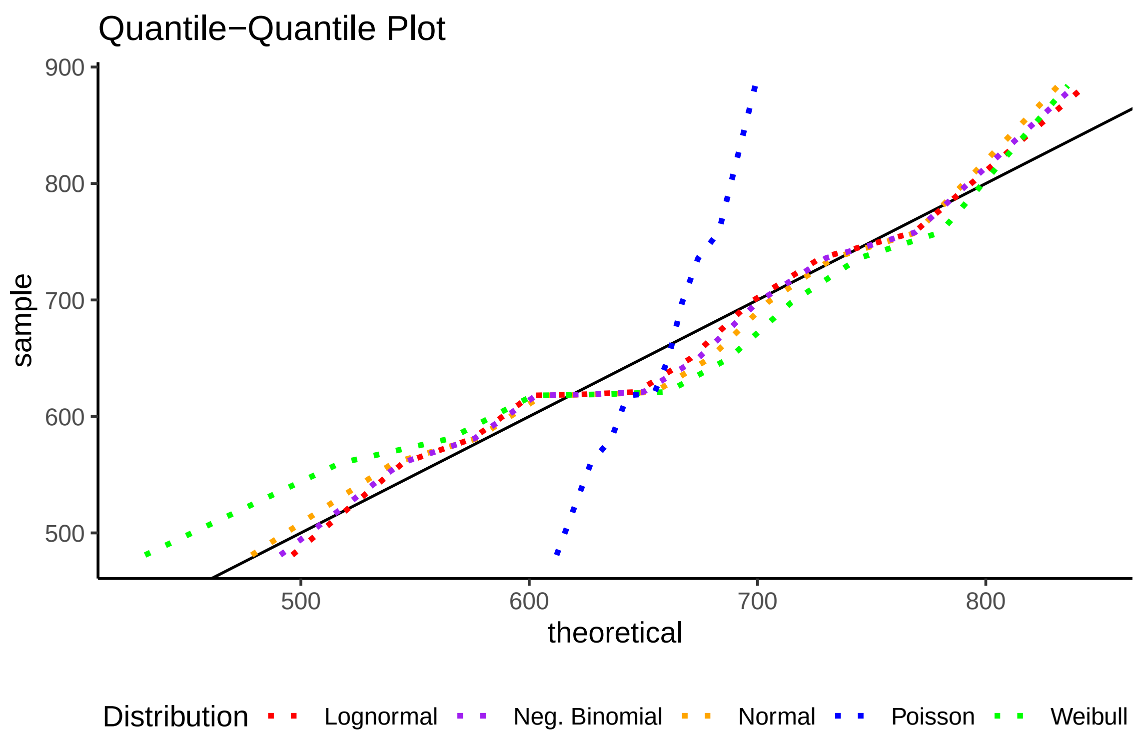


Add some commentary about the chart.

In the simplest case, an actuary may decide that a straight average of historical counts is the best approximation of future counts. In this case, the average number of fires per year is 655 ("Method A").

However, as actuaries, we like to incorporate more information than just raw data. Let's say you know there was a change to building construction in 1989 that accounts for the spike seen in the chart. You would give more weight to those periods and less to the pre-1989 values. Of course, 1990 might be an outlier, so you decide to use a four year average (excluding 1990) of 731 ("Method B").

Your next logical step might be to add some process risk around your estimate by fitting a distribution to the data. Using the maximum-likelihood-estimates of a few common distributions, you settle on using a Lognormal distribution parameterized using a mean of 731 and standard deviation of 15%.



After running 10,000 simulations with your selected distribution, you're left with a central estimate of 731 ("Method C").

But as we discussed earlier, there are a few problems with just relying on our data:

- The sample size is quite small (n = 11).
- That knowledge of a change to building construction may be misplaced.
- Regression to the mean.
- It might look like there is a "new normal" but you could be seeing a trend where there isn't one.

So how do we handle these possible baises? That's where Bayesian MCMC comes in. Our first model will assume a wide prior. We'll assume that our Lognormal distribution has the following prior distributions:

- mu ~ Normal(6.583, 0.658)
- $sigma$ ~ Exponential(6.704)

After running 10,000 simulations with your selected distribution, you're left with a central estimate of 659 ("Method D"). Since you used a wide prior, your result will be fairly close to Method A even though you had a much higher prior. That is because the data is overwhelmingly lower, thus a higher estimate is less likely.

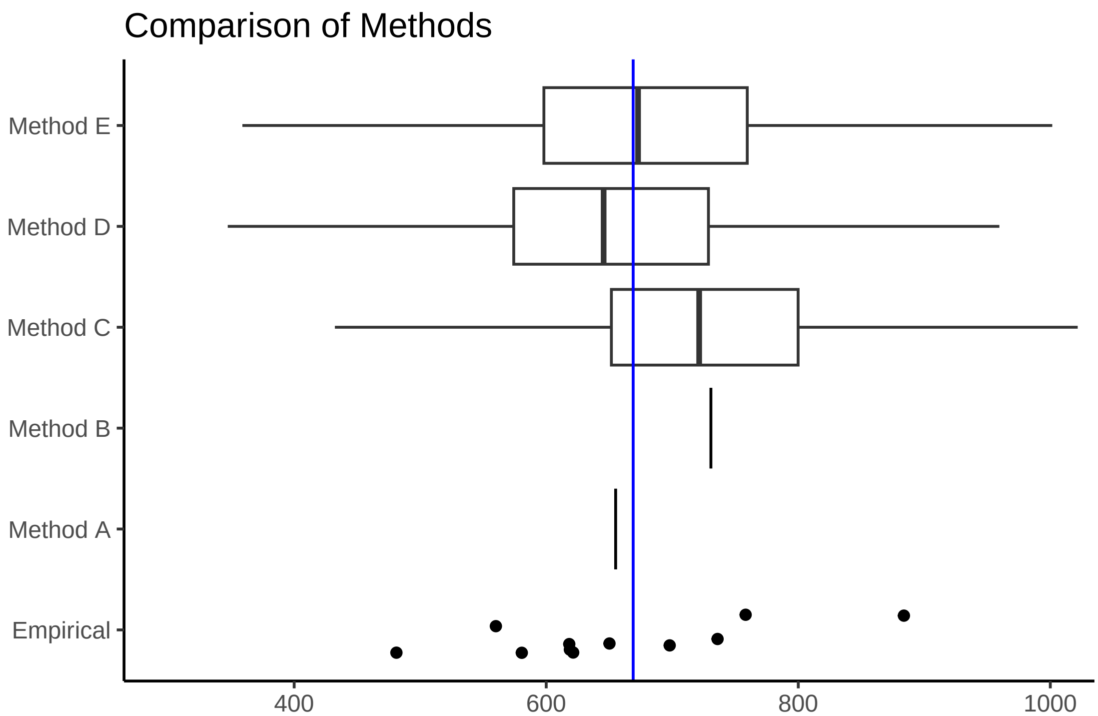
Next, let's test "tighter" our prior. After all, we're pretty confident of the change. We'll assume that our Lognormal distribution has the following prior distributions:

- mu ~ Normal(6.583, 0.066)

- *sigma* ~ Exponential(6.704)

After running 10,000 simulations with your selected distribution, you're left with a central estimate of 688 ("Method E"). While this is higher than your "wide" prior estimate in Method D, it still falls short of the estimate in Method B.

Let's compare each result to the true number of fires that occured in 1993 (669).



Add commentary.

Literature Review

Bayesian MCMC is currently being used by actuaries in various applications:

- Reserving (Meyers 2015)
- Renewal Functions (Aminzadeh & Deng 2022)
- Risk Margins (Meyers 2019)
- Trends (Schmid 2013)

We also believe there is a number of other possible applications, ranging from selection of development factors to estimating reserve variability. We leave it to the reader to explore these.

Acknowledgements

Thank everyone that peer review

Biographies of the Authors

Spencer Miller

To be completed later

Kenny Smart

To be completed later

Citations

References

A. Tversky, and D. Kahneman, Judgment under uncertainty: Heuristics and biases, Science 185, 1124–1131 (1974).

Meyers, G. 2015. "Stochastic Loss Reserving Using Bayesian MCMC Models." CAS Monograph #1.

Aminzadeh, M.S., and Min Deng. 2022. "Bayesian Estimation of Renewal Function Based on Pareto-Distributed Inter-Arrival Times via an MCMC Algorithm." Variance 15 (2).

Meyers, G. 2019. "A Cost-of-Capital Risk Margin Formula for Nonlife Insurance Liabilities." Variance 12 (2).

Schmid, F. 2013. "Bayesian Trend Selection." Casualty Actuarial Society E-Forum, Spring 2013.

Software

Barrett, Tyson, Matt Dowle, Arun Srinivasan, Jan Gorecki, Michael Chirico, Toby Hocking, Benjamin Schwendinger, and Ivan Krylov. 2025. "Data.table: Extension of 'Data.frame'." https://CRAN.Rproject.org/package=data.table.

Dutang, Christophe, and Arthur Charpentier. 2025. "CASdatasets: Insurance Datasets." https://doi.org/10.57745/P0KHAG.

Stan Development Team. 2025. "{RStan}: The {r} Interface to {Stan}." https://mc-stan.org/.

Wickham, Hadley. 2016. "Ggplot2: Elegant Graphics for Data Analysis." https://ggplot2.tidyverse.org.