

Introduction to Bayesian linear regression with brms

Stefano Coretta

18/01/2020

[mention installation]

Random variables

- We have a question about the world, so we collect data (sample from a population).
 - $y = (y_1, y_2, y_3, y_4, \dots, y_n)$
- We want to know how the data (the sample) was generated.
- In probability theory, data is generated by a random variable Y .

Random variables

- Y is uncertain.
 - We can describe Y as a probability distribution, expressed by a set of parameters $\Theta = (\theta_1, \dots, \theta_n)$.
- Probability distributions:
 - $Normal(\mu, \sigma)$,
 - $Binomial(n, p)$,
 - ...

$$vot \sim Normal(\mu, \sigma)$$

$$voiced \sim Bernoulli(p)$$

$$DoubleDative \sim Poisson(\lambda)$$

Frequentist vs Bayesian view

- Parameters: $\mu, \sigma, p, \lambda, \dots$
- Frequentist view:
 - The parameters are **fixed** (they are unknown but certain).
 - They take on a specific value.
- Bayesian view:
 - The parameters are **random variables** (they are unknown and uncertain).
 - We describe each parameter as a probability distribution, expressed by a set of **hyperparameters**.

Continuous random variable

$$vot_i \sim Normal(\mu, \sigma)$$

$$\mu \sim Normal(\mu_1, \sigma_1)$$

$$\sigma \sim HalfCauchy(x_0, \gamma)$$

Priors

- We can incorporate previous knowledge about the hyperparameters as **priors** (prior distributions).
- Priors are chosen based on expert knowledge, previous studies, pilot data...
 - Priors must **not** be chosen based on the data to be analysed.

Priors

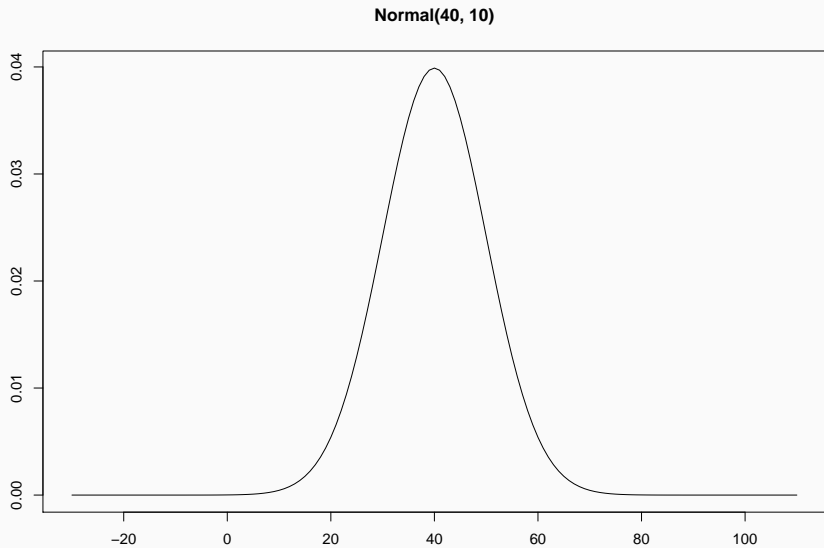
- Informative and uninformative priors.
- Informative priors.
 - Weakly informative.
 - Strongly informative.
- Uninformative priors.
 - Uniform distribution.

Normal prior

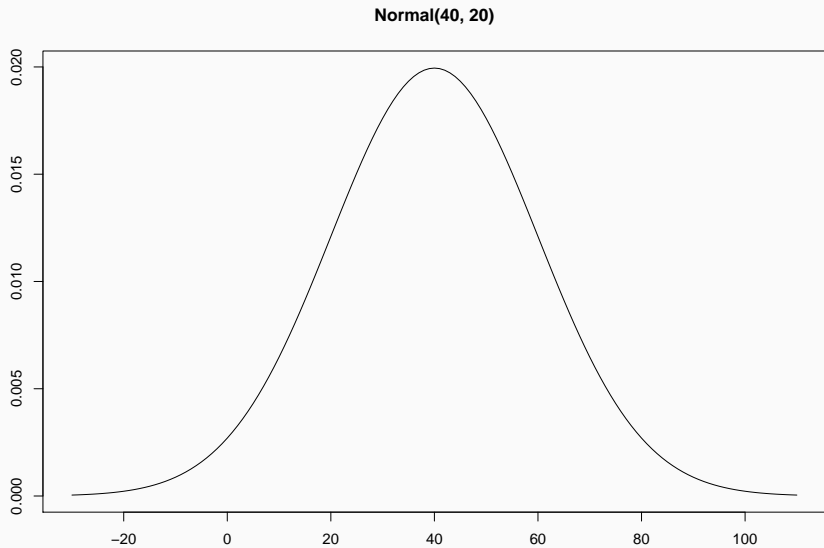
[empirical rule]

- Previous literature on VOT in Italian (Esposito, 2002; Stevens & Hajek, 2010) report VOT values for voiceless stops in the range of 20–60 ms.
 - We can express this knowledge with the prior $Normal(40, 10)$.
 - This is a somewhat strongly informative prior.

Italian VOT



Italian VOT

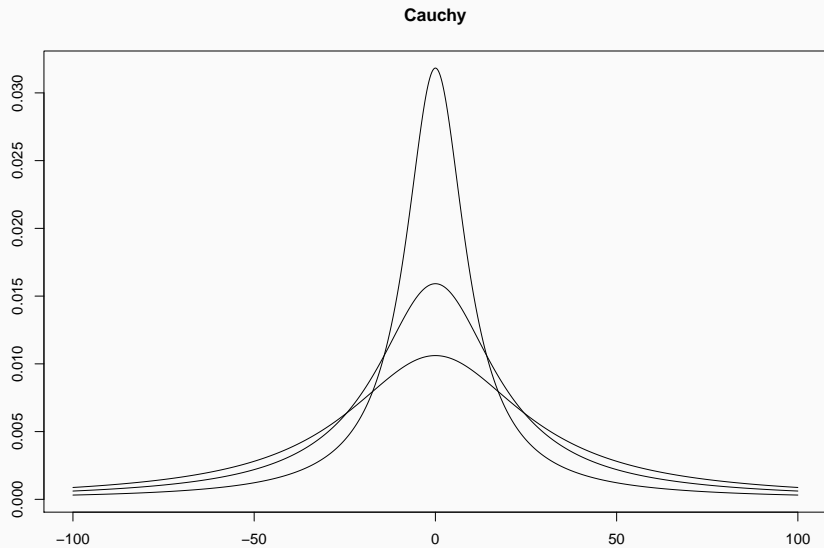


$$vot_i \sim Normal(\mu, \sigma)$$

$$\mu \sim Normal(40, 10)$$

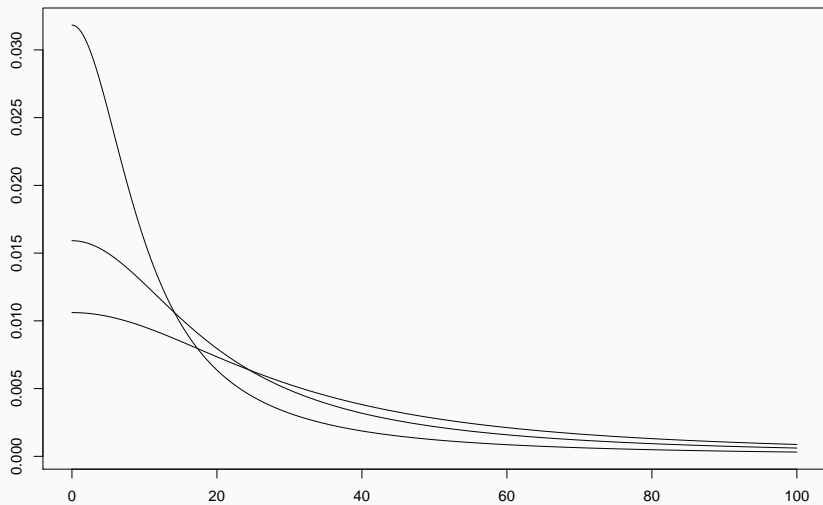
$$\sigma \sim HalfCauchy(x_0, \gamma)$$

Cauchy prior



Cauchy prior

HalfCauchy



$$vot_i \sim Normal(\mu, \sigma)$$

$$\mu \sim Normal(40, 10)$$

$$\sigma \sim HalfCauchy(0, 10)$$

- We have a model which incorporates our knowledge about VOT (through the priors for μ and σ).
- Now we want to obtain the **posterior distributions** of μ and σ .
 - The posterior distribution is the prior distribution *conditioned* on the data.
- **brms** R package: Bayesian Regression Models using Stan (Bürkner, 2018).

- Stan (Stan Development Team, 2017).
 - Statistical programming language written in C++ for fitting Bayesian models (calculate posterior distributions).
 - Calculation can be complex and/or impossible, so we take many samples from the data and from the possible parameter values to find the posterior distributions of the hyperparameters.
 - Markov Chain Monte Carlo (MCMC) sampling using the No-U-Turn sampler (NUTS).
- brms is an interface between R and Stan.
- `brm()` function from brms.
 - lme4 syntax ($y \sim x + (1|w)$).
 - Creates a Stan model, which is compiled and run.

```
vot1 <- brm(  
  <model_formula>,  
  <family>,  
  <prior>,  
  <data>,  
  chains = 4,  
  iter = 2000  
)
```

```
vot1 <- brm(  
  vot ~ 1,  
  family = gaussian(),  
  <prior>,  
  data = ita_egg,  
  chains = 4,  
  iter = 2000  
)
```

Get prior

```
get_prior(  
  vot ~ 1,  
  family = gaussian(),  
  data = ita_egg  
)
```

```
##               prior      class coef group resp dpa  
## 1 student_t(3, 19, 14) Intercept  
## 2  student_t(3, 0, 14)      sigma
```

References

- Bürkner, Paul-Christian. 2018. Advanced Bayesian multilevel modeling with the R package brms. *The R Journal* 10(1). 395–411. doi:10.32614/RJ-2018-017.
- Esposito, Anna. 2002. On vowel height and consonantal voicing effects: Data from Italian. *Phonetica* 59(4). 197–231. doi:10.1159/000068347.
- Stan Development Team. 2017. Stan: A C++ library for probability and sampling, version 2.14.0. <http://mc-stan.org/>.

Stevens, Mary & John Hajek. 2010. Post-aspiration in standard Italian: some first cross-regional acoustic evidence. Paper presented at Interspeech, 26-30 September 2010, Makuhari, Chiba, Japan.