Statistical Analysis

Ramya Ghantasala

15 March, 2022

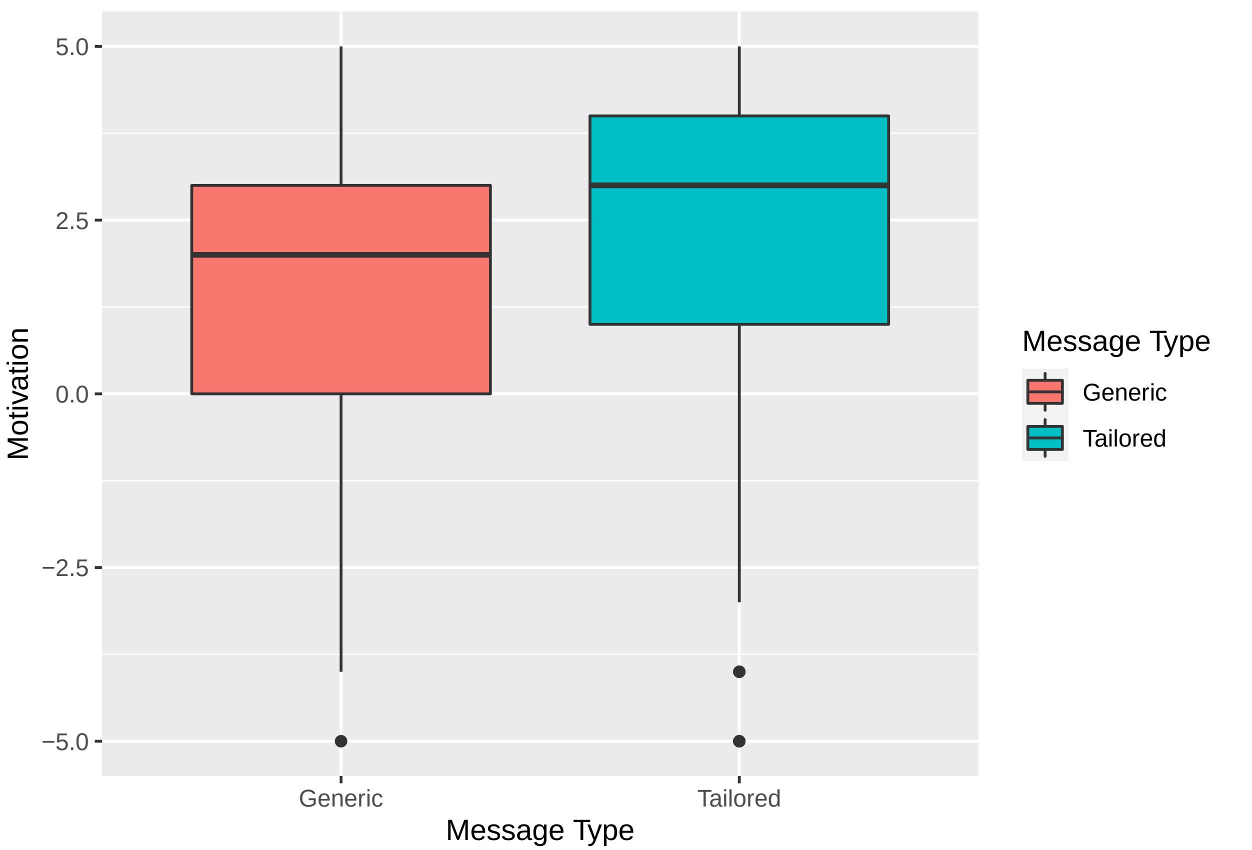
This file reproduces Figure 4.4, "Comparison of the motivation of the two message types.", Figure 4.5, "Comparison of the two models fitted on motivation.", values in Table 4.3, "Summary of model 2".

Package Installation

```r
library(rethinking)
library(reshape2)
library(ggplot2)
```

Comparing motivational levels of the two message types.

The below figure compares the motivational level of generic and tailored messages.



This figure corresponds to Figure 4.4, "Comparison of the motivation of the two message types." from the thesis report.

Hierarchical Modelling

Person as random intercept

The data obtained from the experiment is fit for motivation to a base linear model with random intercepts for each participant (m1).

```
#Fixed effect of motivation + Person as random intercept
set.seed(4)
m1 <- ulam(
  alist(
    motivation ~ dstudent(v, mu, sigma),
    mu <- a + a_person[pid] * sigma_p,
    #adaptive prior
    a_person[pid] ~ dnorm(0, 1),
    #hyper prior
    sigma_p ~ dexp(1),
    #fixed priors
    v ~ gamma(2, 0.1),
    a ~ dnorm(0, 5),
    sigma ~ dexp(1)
  ), data = combinedMotivation, iter = 10000, chains = 4, cores = 4,  log_lik = TRUE,
  control=list(adapt_delta=.99, max_treedepth = 15)
)
```

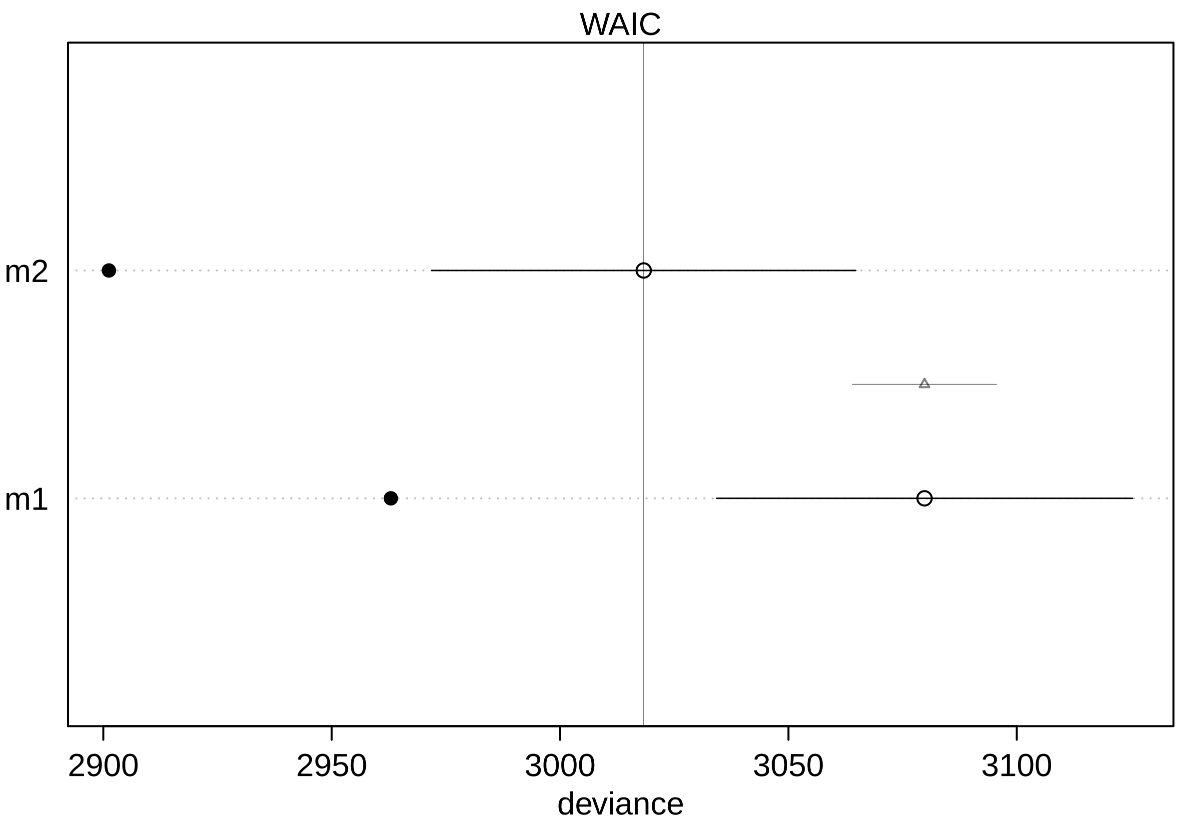
Person as random intercept + Message type as fixed effect

The second model (m2) extends m1 along with a fixed effect for message type.

```
#Message type as fixed effect
set.seed(4)
m2 <- ulam(
  alist(
    motivation ~ dstudent(v, mu, sigma),
    #model
    mu <- a + (a_person[pid] * sigma_p) + (b * messageType),
    #adaptive prior
    a_person[pid] ~ dnorm(0, 1),
    #hyper prior
    sigma_p ~ dexp(1),
    #fixed priors
    v ~ gamma(2, 0.1),
    a ~ dnorm(0, 5),
    b ~ dnorm(0, 10),
    sigma ~ dexp(1)
  ), data = combinedMotivation, iter = 10000, chains = 4, cores = 4, log_lik = TRUE,
  control=list(adapt_delta=.99, max_treedepth = 15)
)
```

Comparison of the two models

The fit of the two models, m1 and m2 are compared. Model two has lesser deviance than model one, and hence, model two will perform better for out-of-sample predictions.



The above figure corresponds to Figure 4.5, "Comparison of the two models fitted on motivation."

Posterior Probability

We are interested in the value of b here, as it is the coefficient of message type.

```
set.seed(4)
post <- extract.samples(m2, 10000)
```

The samples extracted from the model m2 are surmised below:

```
precis(post, depth=1, prob=.95)
```

```
##              mean         sd      2.5%    97.5%       histogram
## sigma_p 1.119632 0.12730724 0.8921293 1.400583
## v       3.517768 0.41557221 2.7825106 4.410004
## a       1.530863 0.17320517 1.1899976 1.868219
## b       1.015278 0.13370934 0.7575785 1.280364
## sigma   1.431181 0.06481896 1.3044911 1.560906
```

```
cat("Calculated posterior probability value is ",  (length(post[which(post$b>0)])/length(post$b)))
```

```
## Calculated posterior probability value is  1
```

We are interested in the parameter `b`, which is the coefficient for message type. The values for `b` correspond to the vales given in Table 4.3, "Summary of model 2". The posterior probability calculated by extracting samples from model 2 is 1, supporting the hypothesis that tailored messages are more motivating than generic messages. The posterior probability is interpreted according to Chechile's guidelines (Chechile 2020) as "Virtually certain for H1".

References

Chechile, Richard A. 2020. *Bayesian Statistics for Experimental Scientists*.