Bayesian Analysis of Steps

Nele Albers

23 January, 2024

Contents

This file is to reproduce our Bayesian analyses comparing the steps taken before and on the last day of the study, reported in section 2.3 ("Data Collection and Model Training").

Authored by Martin Dierikx, Nele Albers, Bouke L. Scheltinga, and Willem-Paul Brinkman.

Required and output files

The following files are required: preprocessed_step_data.csv

And these files are created: None.

Load packages

First, we load the packages that we need.

```
library(BayesianFirstAid)  # For Bayesian t-test
library(formatR)  # To wrap lines
```

Load data

Now we load the pre-processed data.

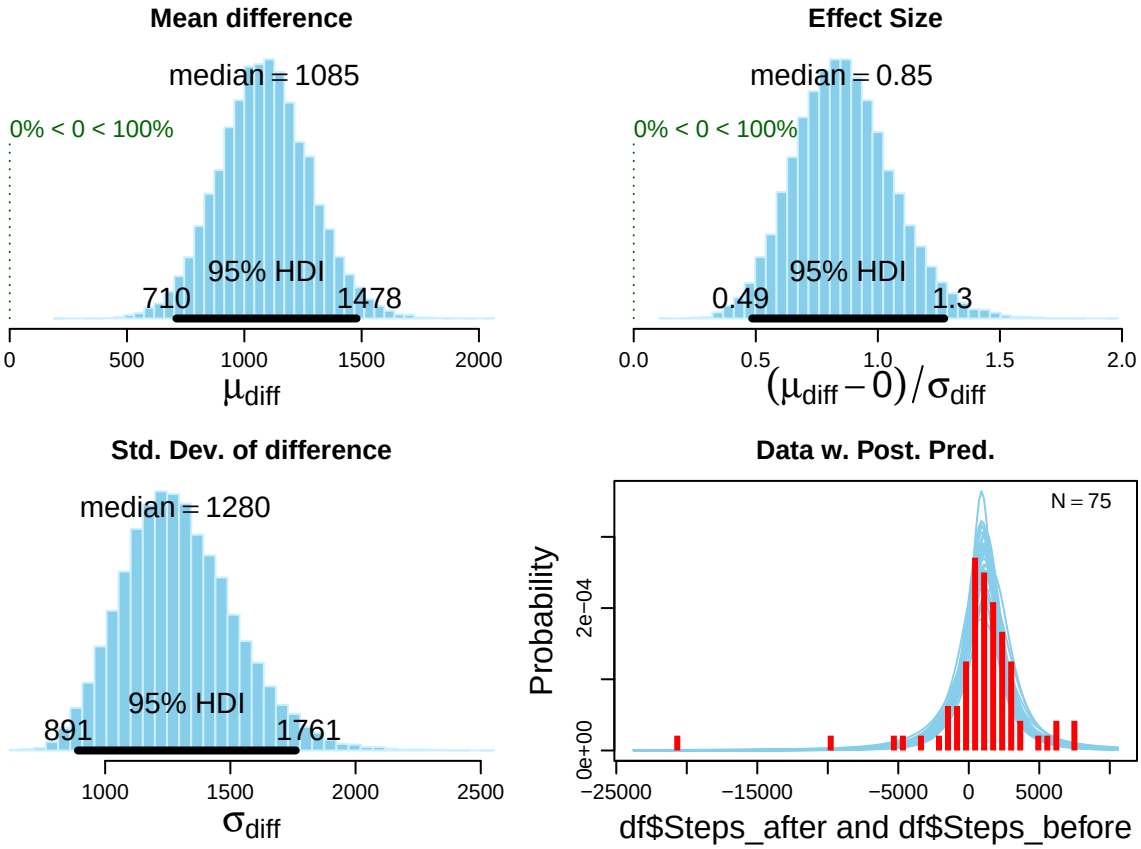
```
df <- read.csv(file = 'preprocessed_step_data.csv')
```

Bayesian analysis of change in steps

Now we conduct a paired Bayesian t-test comparing the steps that the 75 people who finished the post-questionnaire took before the study and on the last day of the study.

```r
set.seed(22)  # For reproducibility

fit1 <- bayes.t.test(df$Steps_after, df$Steps_before,
    paired = TRUE)
plot(fit1)
```




```r
summary(fit1)
```

```
##   Data
## df$Steps_after, n = 75
## df$Steps_before, n = 75
##
##   Model parameters and generated quantities
## mu_diff: the mean pairwise difference between df$Steps_after and df$Steps_before
## sigma_diff: the scale of the pairwise difference, a consistent
##   estimate of SD when nu is large.
## nu: the degrees-of-freedom for the t distribution fitted to the pairwise difference
## eff_size: the effect size calculated as (mu_diff - 0) / sigma_diff
## diff_pred: predicted distribution for a new datapoint generated
##   as the pairwise difference between df$Steps_after and df$Steps_before
##
##   Measures
##                 mean        sd     HDIlo    HDIup %<comp %>comp
## mu_diff    1087.310   194.585   710.497 1478.353  0.000  1.000
## sigma_diff 1297.506   226.225   891.420 1761.081  0.000  1.000
## nu            1.866     0.544     1.001    2.864  0.000  1.000
## eff_size      0.861     0.202     0.485    1.272  0.000  1.000
```

```
## diff_pred   921.500 19230.700 -5332.093 7559.016  0.252  0.748
##
## 'HDIlo' and 'HDIup' are the limits of a 95% HDI credible interval.
## '%<comp' and '%>comp' are the probabilities of the respective parameter being
## smaller or larger than 0.
##
##   Quantiles
##                q2.5%     q25%   median     q75%   q97.5%
## mu_diff      713.152  957.330 1085.208 1213.771 1481.160
## sigma_diff   905.126 1139.312 1280.117 1438.628 1786.550
## nu             1.113    1.479    1.774    2.139    3.166
## eff_size       0.493    0.721    0.852    0.989    1.282
## diff_pred  -5275.216   -8.396 1067.655 2161.023 7613.881
```

```
# Mean, SD, CI
diff_mean <- fit1$stats[1, 1]
diff_SD <- fit1$stats[2, 1]
diff_ci_low <- fit1$stats[1, 5]
diff_ci_high <- fit1$stats[1, 6]

# Posterior probability that difference is
# greater than 0
diff_post_p <- fit1$stats[1, 7]

print(paste("Posterior probability that difference is greater than 0:",
    round(diff_post_p, 3)))
```

```
## [1] "Posterior probability that difference is greater than 0: 1"
```

```
print(paste("Mean =", round(diff_mean, 2), "SD =",
    round(diff_SD, 2), "95% HDI = [", round(diff_ci_low,
        2), ",", round(diff_ci_high, 2), "]"))
```

```
## [1] "Mean = 1087.31 SD = 1297.51 95% HDI = [ 710.5 , 1478.35 ]"
```

This posterior probability can be evaluated based on the guidelines from (Chechile (2020)) and their extension to posterior probabilities below 0.5 by (Andraszewicz et al. (2015)).

```
if (diff_post_p < 0.0005){
  evaluation_diff = "Nearing certainty against"
}else if (diff_post_p < 0.005){
  evaluation_diff = "Very strong bet against"
}else if (diff_post_p < 0.01){
  evaluation_diff = "Strong bet against - irresponsible to avoid"
}else if (diff_post_p < 0.1){
  evaluation_diff = "A promising but risky bet against"
}else if (diff_post_p < 0.25){
  evaluation_diff = "Only a casual bet against"
}else if (diff_post_p < 0.5){
  evaluation_diff = "Not worth betting against"
}else if (diff_post_p < 0.75){
  evaluation_diff = "Not worth betting on"
}else if (diff_post_p < 0.9){
  evaluation_diff = "Only a casual bet"
}else if (diff_post_p < 0.95){
  evaluation_diff = "A promising but risky bet"
```

```
}else if (diff_post_p < 0.99){
  evaluation_diff = "Good bet - too good to disregard"
}else if (diff_post_p < 0.995){
  evaluation_diff = "Strong bet - irresponsible to avoid"
}else if (diff_post_p < 0.9995){
  evaluation_diff = "Very strong bet"
}else if (diff_post_p < 0.99995){
  evaluation_diff= "Nearing certainty"
}else{
  evaluation_diff = "Virtually certain"
}

evaluation_diff
```

```
## [1] "Virtually certain"
```

References

Andraszewicz, Sandra, Benjamin Scheibehenne, Jörg Rieskamp, Raoul Grasman, Josine Verhagen, and Eric-Jan Wagenmakers. 2015. "An Introduction to Bayesian Hypothesis Testing for Management Research." *Journal of Management* 41 (2): 521–43.

Chechile, Richard A. 2020. *Bayesian Statistics for Experimental Scientists: A General Introduction Using Distribution-Free Methods*. MIT Press.