03 Linear Modeling Exercise: Power Posing

Shravan Vasishth

September 02, 2019

Power posing: Background

The claim is that adopting a high power pose for two minutes will increase your testosterone, improving your performance in, e.g., job interviews. We are going to evaluate this claim based on Cuddy's data. The data were released by Cuddy via Nathan Fosse:

https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/FMEGS6

Incidentally, the first author of the Cuddy paper has disavowed the paper: see here.

Carney, D. R., Cuddy, A. J., & Yap, A. J. (2010). Power posing: Brief nonverbal displays affect neuroendocrine levels and risk tolerance. Psychological science, 21(10), 1363-1368.

Preparation for the exercise

Load the data, basic checks

```
## cleaned data
datc<-read.csv("data/ccy-clean-data.csv",
               header=TRUE)

## sanity check: one subject, one row
dim(datc)
```

```
## [1] 47 41
```

```
length(unique(datc$id))
```

```
## [1] 47
```

```
#drop ineligible and something else as in stata code:
datc<-subset(datc,inelig!="Ineligible (drop)" & anyoutv1!="Selected")
```

A fairly typical characteristic of data is the sheer messiness of it.

```
head(datc)
```

```
##   id                inelig ccydrop cortm1v2 cortm2v2 cdiffv2 testm1v2
## 2 29 Analytic sample (keep)       0    0.003    0.025   0.022    38.72
## 3 30 Analytic sample (keep)       0    0.086    0.167   0.081    32.77
## 4 31 Analytic sample (keep)       0    0.047    0.059   0.012    32.32
## 5 32 Analytic sample (keep)       0    0.106    0.112   0.006    17.99
## 7 34 Analytic sample (keep)       1       NA    0.171      NA    73.58
## 8 35 Analytic sample (keep)       0    0.153    0.183   0.030    80.69
##   testm2v2 tdiffv2    testoutv1    cortoutv1     anyoutv1    testoutv2
## 2    62.37   23.65 Not selected Not selected Not selected Not selected
## 3    29.23   -3.54 Not selected Not selected Not selected Not selected
## 4    27.51   -4.81 Not selected Not selected Not selected Not selected
```

```
## 5    28.66   10.67 Not selected Not selected Not selected Not selected
## 7    44.67  -28.91 Not selected Not selected Not selected Not selected
## 8   105.48   24.79 Not selected Not selected Not selected Not selected
##      cortoutv2     anyoutv2 pose1rate pose2rate poseratem saldiff
## 2 Not selected Not selected         6         6       6.0  24.306
## 3 Not selected Not selected         7         6       6.5  20.833
## 4 Not selected Not selected         6         5       5.5  29.167
## 5 Not selected Not selected         6         7       6.5  18.750
## 7     Selected     Selected         7         7       7.0  23.611
## 8 Not selected Not selected         7         7       7.0  34.028
##   sal2manip hptreat female age cort1a1 cort1a2 cort2a1 cort2a2 cortm1
## 2    19.806    High   Male  19   0.004   0.001   0.027   0.023 0.0025
## 3    16.333     Low Female  20   0.085   0.086   0.174   0.161 0.0855
## 4    24.667    High Female  20   0.049   0.045   0.056   0.062 0.0470
## 5    14.250     Low Female  18   0.107   0.105   0.111   0.113 0.1060
## 7    19.111     Low Female  21   0.486   0.482   0.175   0.166 0.4840
## 8    29.528    High Female  20   0.159   0.147   0.179   0.188 0.1530
##   cortm2   cdiff test1a1 test1a2 test2a1 test2a2 testm1  testm2   tdiff
## 2 0.0250  0.0225   39.87   37.58   64.22   60.53 38.725  62.375  23.650
## 3 0.1675  0.0820   33.22   32.32   29.43   29.04 32.770  29.235  -3.535
## 4 0.0590  0.0120   32.52   32.12   27.98   27.04 32.320  27.510  -4.810
## 5 0.1120  0.0060   19.74   16.25   28.17   29.14 17.995  28.655  10.660
## 7 0.1705 -0.3135   78.85   68.31   46.14   43.20 73.580  44.670 -28.910
## 8 0.1835  0.0305   83.51   77.88  105.92  105.05 80.695 105.485  24.790
##    feelpower   incharge powm diceroll
## 2          3          2  2.5      Yes
## 3 Not at all          2  1.5       No
## 4          2 Not at all  1.5      Yes
## 5          3  Very much  3.5      Yes
## 7          2          2    2       No
## 8          3          3    3      Yes
```

Critical variables:

- id (numerical, should be factor): subject id
- testm1 (numerical): pre-treatment testosterone
- testm2 (numerical): post-treatment testosterone
- cortm1 (numerical): pre-treatment cortisone
- cortm2 (numerical): post-treatment cortisone
- female (factor): Female, Male
- hptreat (factor): High, Low (power pose)

So, isolate relevant columns:

```
dat<-datc[,c(1,21,22,28,29,35,36)]
head(dat)
```

```
##   id hptreat female cortm1 cortm2 testm1  testm2
## 2 29    High   Male 0.0025 0.0250 38.725  62.375
## 3 30     Low Female 0.0855 0.1675 32.770  29.235
## 4 31    High Female 0.0470 0.0590 32.320  27.510
## 5 32     Low Female 0.1060 0.1120 17.995  28.655
## 7 34     Low Female 0.4840 0.1705 73.580  44.670
## 8 35    High Female 0.1530 0.1835 80.695 105.485
```

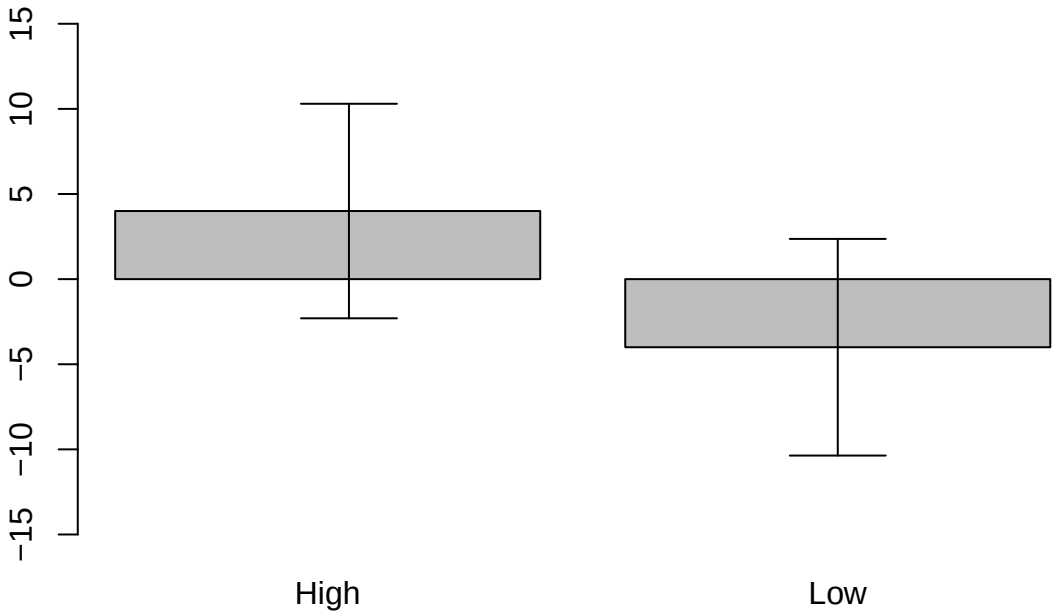


Figure 1: Exercise 4: Recreation of Figure 3 in the Cuddy power posing study.

Exercise 1

Plot the distributions of male vs female testosterone levels before and after treatment.

Exercise 2

Calculate mean post-treatment testosterone by gender and by hptreatment (high or low power pose).

Exercise 3

Calculate mean post-treatment testosterone ignoring gender, by hptreatment (high or low power pose).

Exercise 4

By how much did testosterone increase after treatment (testm2-testm1 tells you the increase in testosterone), taking gender and hptreat (treatment: high vs low power pose) into account?

Exercise 4

By how much did testosterone increase after treatment (testm2-testm1 tells you the increase in testosterone), ignoring gender but taking hptreat (treatment: high vs low power pose) into account?

You should be able to reproduce Figure 1 here; this is the rough and ready version I reproduced of fig 3 from the paper. The effects are a bit smaller in this data-set than the published result, probably because of Fosse (the statistician) cleaning up of the publicly released data.

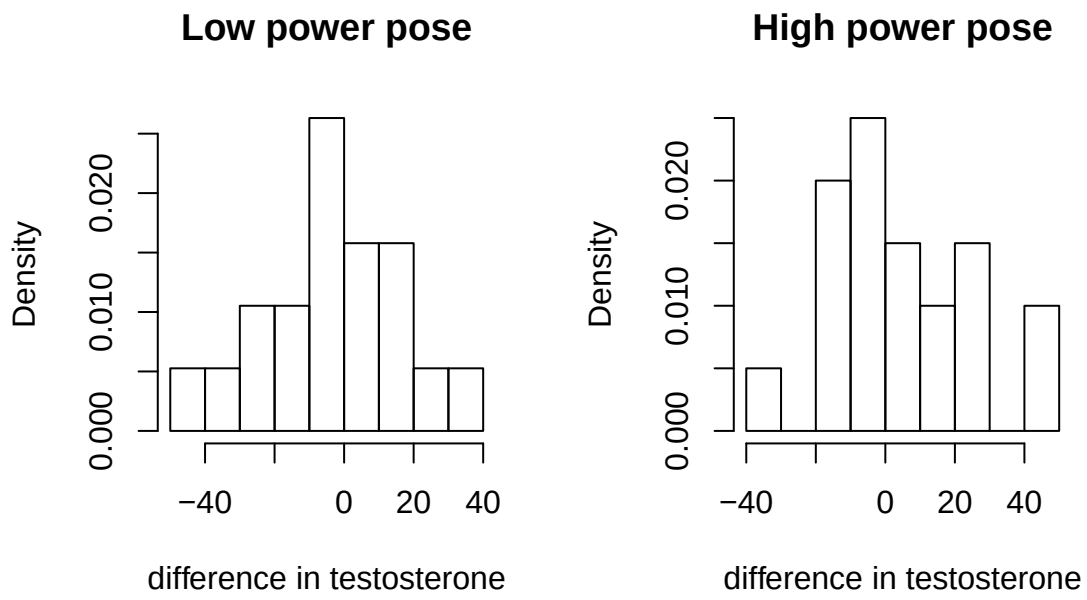


Figure 2: Do high and low power pose lead to differences in testosterone?

Exercise 5

Carry out a frequentist two-sample t-test on the differences in testosterone pre- vs. post-test. What do you conclude from this test?

It will help to visualize the distributions. See Figure ~2.

Exercise 6

Now, refit the model using various predictors:

- the initial testosterone value
- the initial and final cortisone levels
- the gender of the subject.

The question we ask here is, is post-treatment testosterone higher for subjects exposed to high vs low power, controlling for the above variables?

First, center all predictors. We show you below how to do this:

```
## center all predictors
datc$ctestm1<-scale(datc$testm1,scale=FALSE)
datc$chptreat<-ifelse(datc$hptreat=="High",1,-1)
datc$cortm1<-scale(datc$cortm1,scale=FALSE)
datc$cortm2<-scale(datc$cortm2,scale=FALSE)
datc$female<-ifelse(datc$female=="Female",1,-1)
```

Exercise 6.1

Fit the Bayesian version of model m0 below, using brms Here is the frequentist linear model:

```
##
## Call:
## lm(formula = testm2 ~ ctestm1 + chptreat + female, data = datc)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -31.83 -11.54   0.14   7.73  40.63
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   52.415      3.175   16.51  < 2e-16
## ctestm1        0.595      0.128    4.65  4.6e-05
## chptreat       3.044      2.956    1.03    0.310
## female        -7.464      3.689   -2.02    0.051
##
## Residual standard error: 18.3 on 35 degrees of freedom
## Multiple R-squared:  0.593,  Adjusted R-squared:  0.558
## F-statistic:   17 on 3 and 35 DF,  p-value: 5.51e-07
```

Because we are starting with a state of no knowledge (and we have no expert opinions), we will use vague Cauchy(0,1) priors on all parameters. You could also use Cauchy(0,2.5).

Then visualize the result using shinystan (I assume the model in brms is called m0brms):

```
library(shiny)
library(shinystan)
launch_shinystan(m0brms)
```

Exercise 6.2

How does gender affect the conclusions? With the frequentist linear model we get the following:

```
##
## Call:
## lm(formula = testm2 ~ ctestm1 + chptreat + cortm1 + cortm2 +
##     female + chptreat:female, data = datc)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -27.09  -9.51  -2.59   9.34  33.54
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)       54.133      2.906   18.63   <2e-16
## ctestm1            0.386      0.147    2.63   0.0129
## chptreat           4.853      2.944    1.65   0.1090
## cortm1            -1.478     27.693   -0.05   0.9578
## cortm2           147.923     43.610    3.39   0.0019
## female           -12.540      3.890   -3.22   0.0029
## chptreat:female    2.806      3.255    0.86   0.3951
##
## Residual standard error: 16.1 on 32 degrees of freedom
## Multiple R-squared:  0.712,  Adjusted R-squared:  0.657
## F-statistic: 13.2 on 6 and 32 DF,  p-value: 1.87e-07
```

```
##
## Call:
## lm(formula = testm2 ~ ctestm1 + chptreat + cortm1 + cortm2 +
##     female, data = datc)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -26.63  -9.16  -3.29   8.47  33.00
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   53.581      2.824   18.98  < 2e-16
## ctestm1        0.455      0.123    3.70  0.00078
## chptreat       5.822      2.710    2.15  0.03913
## cortm1        -4.989     27.285   -0.18  0.85603
## cortm2       142.046     42.906    3.31  0.00226
## female       -11.173      3.538   -3.16  0.00339
##
## Residual standard error: 16 on 33 degrees of freedom
## Multiple R-squared:  0.705,  Adjusted R-squared:  0.66
## F-statistic: 15.8 on 5 and 33 DF,  p-value: 6.22e-08
```

Fit the above model m1 using brms, and compare the Bayesian posteriors of the effect of treatment to the model m0 above.

Exercise 6.3

An alternative analysis using difference in testosterone: We could also have as dependent measure the **change** in testosterone in low vs high power subjects. This corresponds to the Fig 3 plot in the paper.

```
datc$change<-datc$testm2-datc$testm1
m2a<-lm(change~chptreat,datc)
summary(m2a) ## same t-value as in two-sample t-test
```

```
##
## Call:
## lm(formula = change ~ chptreat, data = datc)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -41.33 -14.90  -0.67  15.11  39.08
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.0507     3.2311    0.02     0.99
## chptreat      4.4173     3.2311    1.37     0.18
##
## Residual standard error: 20.2 on 37 degrees of freedom
## Multiple R-squared:  0.0481, Adjusted R-squared:  0.0224
## F-statistic: 1.87 on 1 and 37 DF,  p-value: 0.18
```

Fit the above model m2a in brms.

Exercise 6.4

Now take gender into account and check whether gender has an effect. Here is the lm model:

```
##
## Call:
## lm(formula = change ~ chptreat + female, data = datc)
##
## Residuals:
##    Min     1Q Median     3Q    Max
##  -42.9  -14.9   -0.8   14.7   39.9
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.432      3.470    0.12     0.90
## chptreat       4.378      3.273    1.34     0.19
## female        -1.140      3.470   -0.33     0.74
##
## Residual standard error: 20.4 on 36 degrees of freedom
## Multiple R-squared:  0.0509, Adjusted R-squared:  -0.0018
## F-statistic: 0.966 on 2 and 36 DF,  p-value: 0.39
```

Fit m3a in brms.

Exercise 6.5

Now look at the effect of treatment and gender and their interaction on the change in testosterone.

```
##
## Call:
## lm(formula = change ~ chptreat * female, data = datc)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -36.17 -12.36  -0.26   9.95  42.98
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.192      3.435    0.06    0.956
## chptreat           5.948      3.435    1.73    0.092
## female            -0.900      3.435   -0.26    0.795
## chptreat:female   -4.673      3.435   -1.36    0.182
##
## Residual standard error: 20.2 on 35 degrees of freedom
## Multiple R-squared:  0.0986, Adjusted R-squared:  0.0213
## F-statistic: 1.28 on 3 and 35 DF,  p-value: 0.298
```

This is an exciting result!!!! ... Well, almost. Sadly, the p-value for treatment *just* misses significance. Fit model m4a using brms.

Exercise 6.6

Interpretation: What can we conclude from this data? Does power posing change testosterone levels?

Exercise 6.7

Fit model m4a using RStan

The model code shown below (needs to be saved in a text file called linearmodel.stan under a directory called StanModels):

```
data { int N;
  real change[N];
  real female[N];
  real treatment[N];
  }
parameters {
  real alpha;
  real beta_female;
  real beta_treatment;
  real beta_interaction;
  real<lower = 0> sigma;
}
model {
  // priors probably need more thought:
  beta_female ~ cauchy(0,2.5);
  beta_treatment ~ cauchy(0,2.5);
  beta_interaction ~ cauchy(0,2.5);
  sigma ~ cauchy(0,2.5);

  for (n in 1:N) {
    change[n] ~ normal(alpha + beta_female * female[n] + beta_treatment * treatment[n] +
    beta_interaction * female[n] * treatment[n], sigma);
  } }
generated quantities {
  real change_ppc[N];
  for (n in 1:N) {
    change_ppc[n] = normal_rng(alpha + beta_female * female[n] + beta_treatment * treatment[n] +
    beta_interaction * female[n] * treatment[n], sigma);
  }
}
```

Here is the code for fitting this model in rstan. You simply need to run it.

```
dat<-list(change=datc$change,
          female=datc$female,
          treatment=datc$chptreat,
          N=length(datc$change))

library(rstan)
```

```
## Loading required package: StanHeaders
```

```
## rstan (Version 2.19.2, GitRev: 2e1f913d3ca3)
```

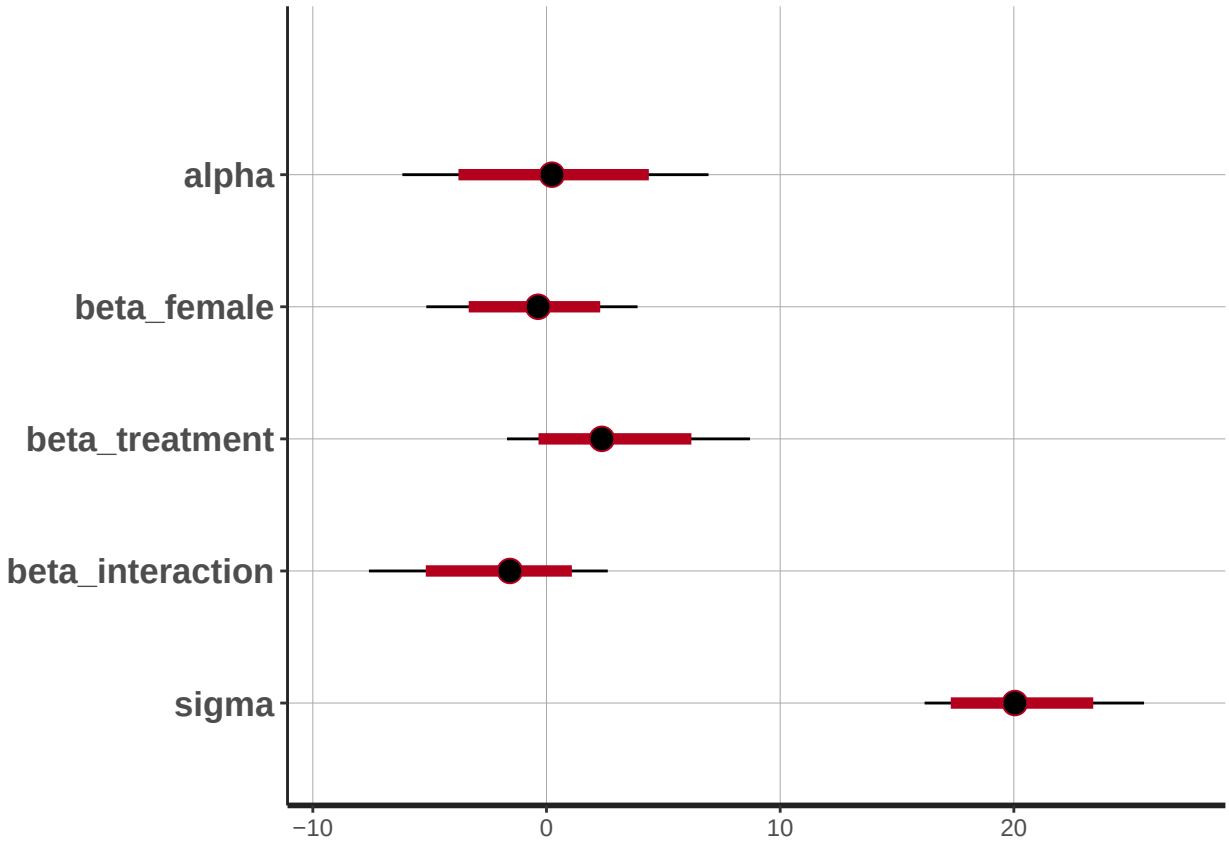
```
## For execution on a local, multicore CPU with excess RAM we recommend calling
## options(mc.cores = parallel::detectCores()).
## To avoid recompilation of unchanged Stan programs, we recommend calling
## rstan_options(auto_write = TRUE)
```

```r
options(mc.cores=parallel::detectCores())

## check if model compiles:
output <- stanc("StanModels/linearmodel.stan")

fit <- stan(file='StanModels/linearmodel.stan', data=dat,
            iter=2000, chains=4, seed=4938483,
            control = list(adapt_delta = 0.8))

paramnames<-c("alpha","beta_female","beta_treatment","beta_interaction","sigma")
print(fit,pars=paramnames)
```

```
## Inference for Stan model: linearmodel.
## 4 chains, each with iter=2000; warmup=1000; thin=1;
## post-warmup draws per chain=1000, total post-warmup draws=4000.
##
##                   mean se_mean   sd  2.5%   25%   50%   75% 97.5% n_eff
## alpha              0.30    0.05 3.26 -6.17 -1.83  0.23  2.42  6.93  5020
## beta_female       -0.46    0.03 2.30 -5.14 -1.84 -0.36  0.95  3.89  4417
## beta_treatment     2.68    0.05 2.66 -1.69  0.80  2.37  4.30  8.71  3450
## beta_interaction -1.82    0.04 2.56 -7.59 -3.30 -1.56 -0.09  2.64  3948
## sigma            20.26    0.04 2.41 16.18 18.52 20.05 21.73 25.57  4366
##                  Rhat
## alpha               1
## beta_female         1
## beta_treatment      1
## beta_interaction    1
## sigma               1
##
## Samples were drawn using NUTS(diag_e) at Mon Sep  2 11:06:03 2019.
## For each parameter, n_eff is a crude measure of effective sample size,
## and Rhat is the potential scale reduction factor on split chains (at
## convergence, Rhat=1).
```

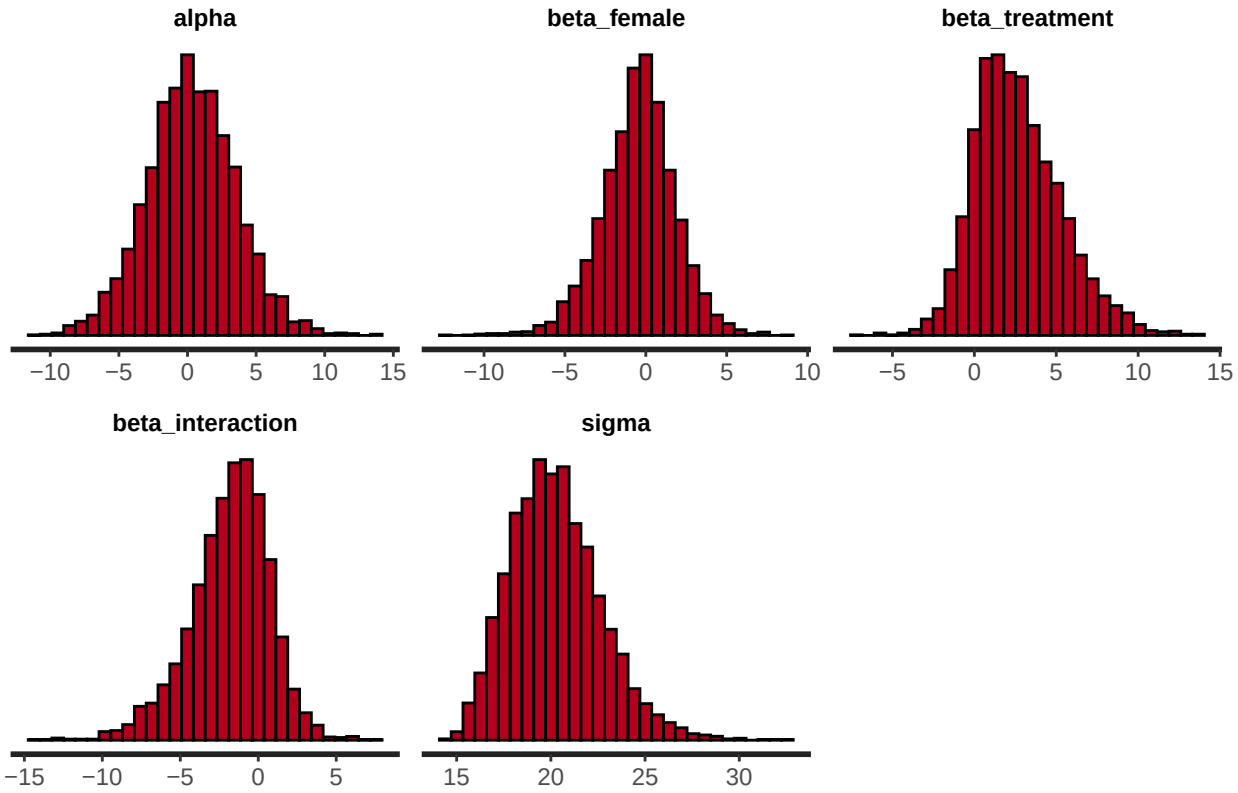
```r
fit_matrix<-as.matrix(fit)
prob<-mean(fit_matrix[,3]>0)

params<-extract(fit,pars=paramnames)

stan_plot(fit,pars=paramnames)
```

```
## ci_level: 0.8 (80% intervals)
## outer_level: 0.95 (95% intervals)
```




```
stan_hist(fit,pars=paramnames)
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```


alpha
beta_female
beta_treatment
beta_interaction
sigma