6. Applicazioni pratiche

Massimiliano Pastore
Università di Padova

Contenuti

Stime bayesiane in generale

When you have eliminated the impossible
all that remain, no matter how improbable,
must be the truth.
(Sherlock Holmes)

- L'inferenza bayesiana può essere considerata una riallocazione di credibilità in uno spazio di possibili candidati.
- In pratica vuol dire che riconsideriamo i *belief* dei valori dei parametri che sono consistenti con i dati ed eliminiamo i valori non consistenti.

t-test

Ripensare il t-test

- A differenza di NHST, il metodo bayesiano consente anche di accettare l'ipotesi nulla, non solo rigettarla.
- Il metodo bayesiano tiene conto di tutte le informazioni relative alle distribuzioni delle medie e delle deviazioni standard dei due gruppi.
- Inoltre consente di gestire outliers e distribuzioni asimmetriche piuttosto che normali.
- In particolare ci occuperemo di stima dei parametri, così da ottenere un risultato molto più ricco ed informativo rispetto al semplice t-test.

Esempio 1

- Uno psicologo vuole valutare se una nuova tecnica di meditazione sia efficace per la riduzione dello stress.
- A tal fine recluta un campione di 60 soggetti stressati, li divide in due gruppi (controllo e trattamento).
- Al termine dell'esperimento calcola i punteggi di stress dei soggetti nei due gruppi e li mette a confronto.

Esempio 1: t test

Le ipotesi:

- H_0 :

Esempio 1: t test

Le ipotesi:

- $H_0 : \mu_t = \mu_c$
- $H_1 :$

Esempio 1: t test

Le ipotesi:

- $H_0 : \mu_t = \mu_c$
- $H_1 : \mu_t < \mu_c$

I risultati:

Esempio 1: t test

Le ipotesi:

- $H_0 : \mu_t = \mu_c$
- $H_1 : \mu_t < \mu_c$

I risultati:

- Gruppo di controllo: $\overline{x}_c = 80.8$, $s_c = 10.05$
- Gruppo di trattamento: $\overline{x}_t = 75.6$, $s_t = 14.76$

Esempio 1: t test

Le ipotesi:

- $H_0 : \mu_t = \mu_c$
- $H_1 : \mu_t < \mu_c$

I risultati:

- Gruppo di controllo: $\overline{x}_c = 80.8$, $s_c = 10.05$
- Gruppo di trattamento: $\overline{x}_t = 75.6$, $s_t = 14.76$

$$t_{(58)} = \frac{(75.6 - 80.8)}{\sqrt{14.76^2/30 + 10.05^2/30}} = -1.59$$

Esempio 1: t test

Le ipotesi:

- $H_0 : \mu_t = \mu_c$
- $H_1 : \mu_t < \mu_c$

I risultati:

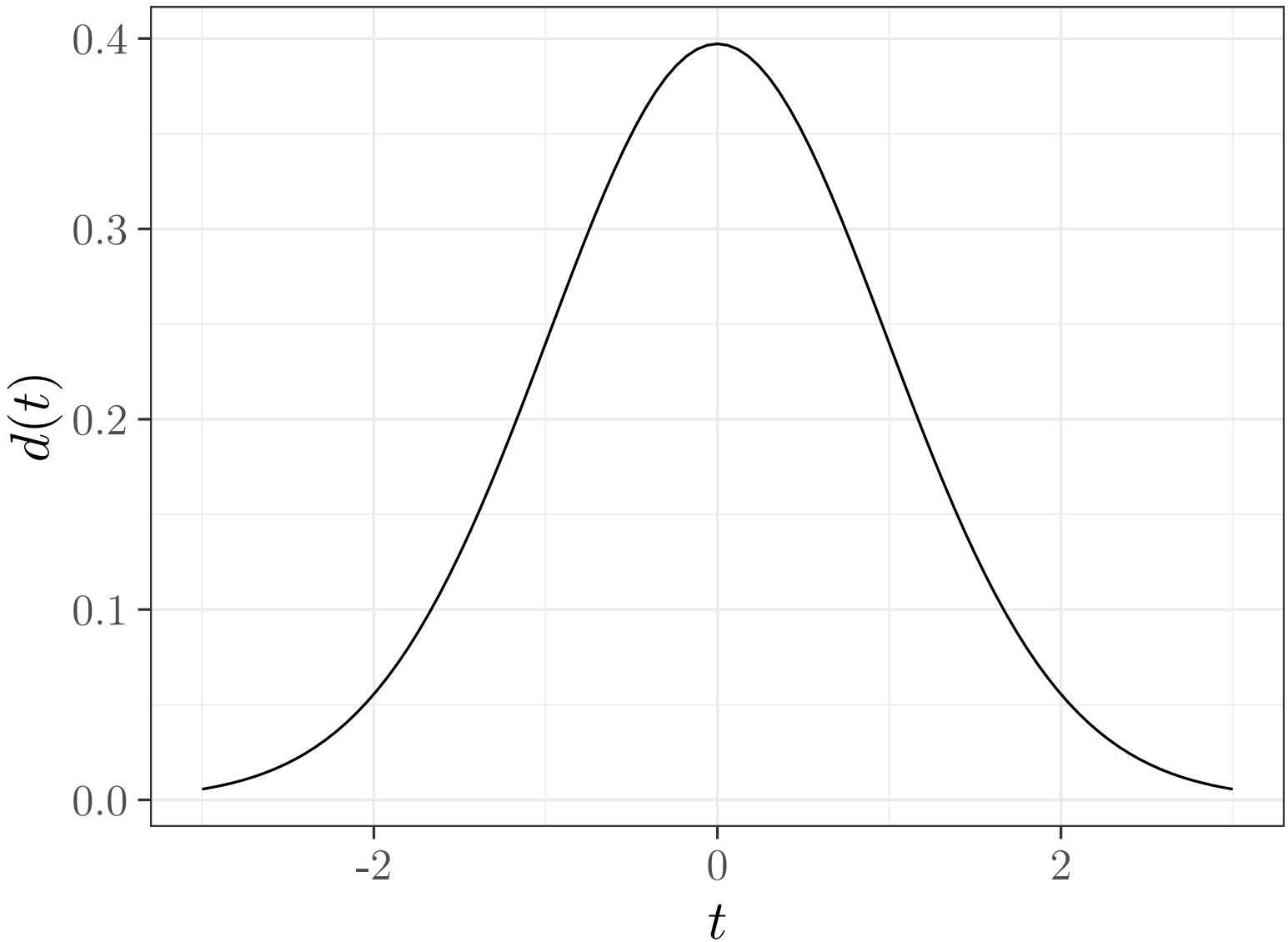
- Gruppo di controllo: $\overline{x}_c = 80.8$, $s_c = 10.05$
- Gruppo di trattamento: $\overline{x}_t = 75.6$, $s_t = 14.76$

$$t_{(58)} = \frac{(75.6 - 80.8)}{\sqrt{14.76^2/30 + 10.05^2/30}} = -1.59$$

$$p = 0.06$$

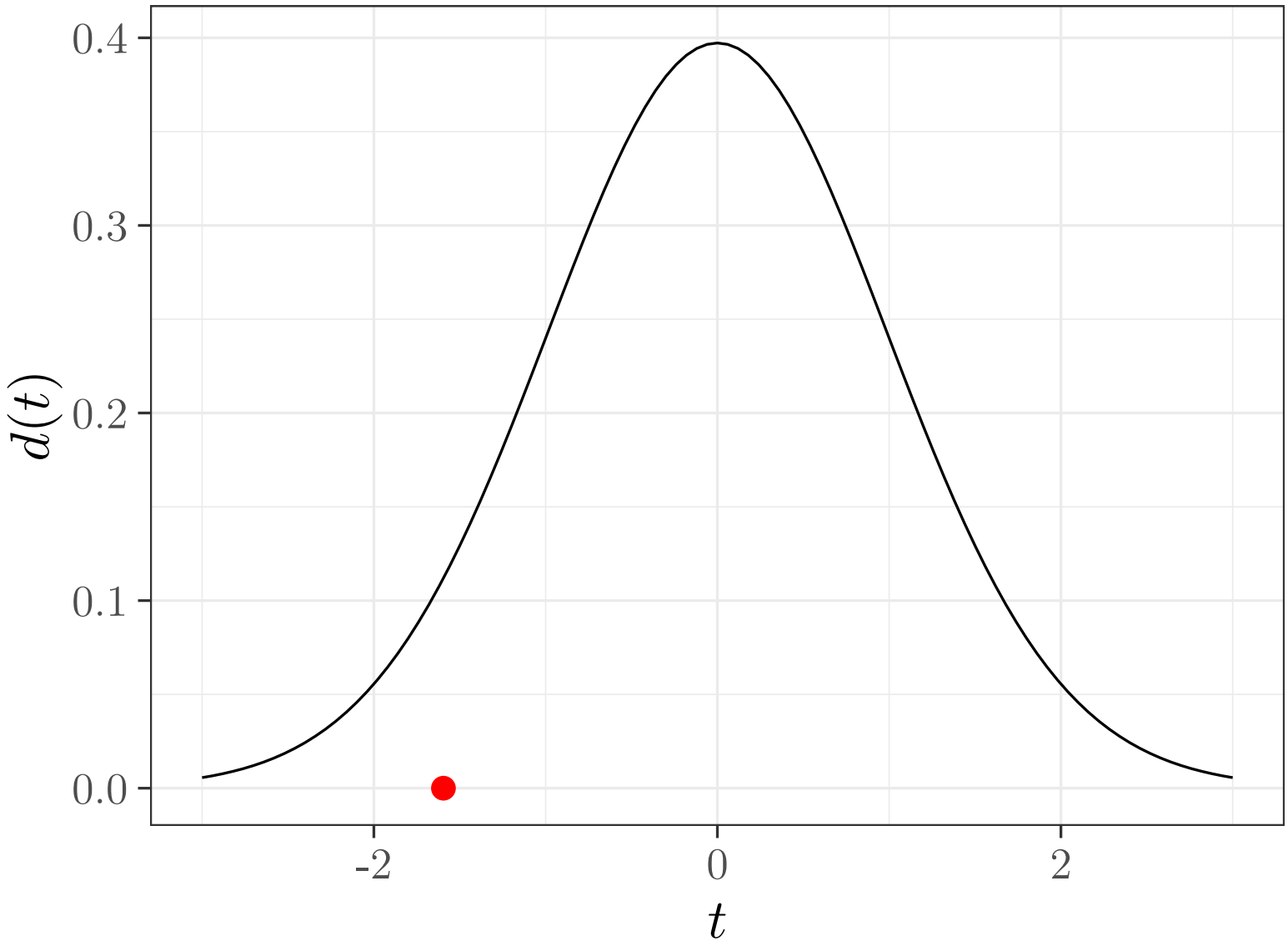
```
curve( dt( x, 58 ), -3, 3 )
```

```
curve( dt( x, 58 ), -3, 3 )
```



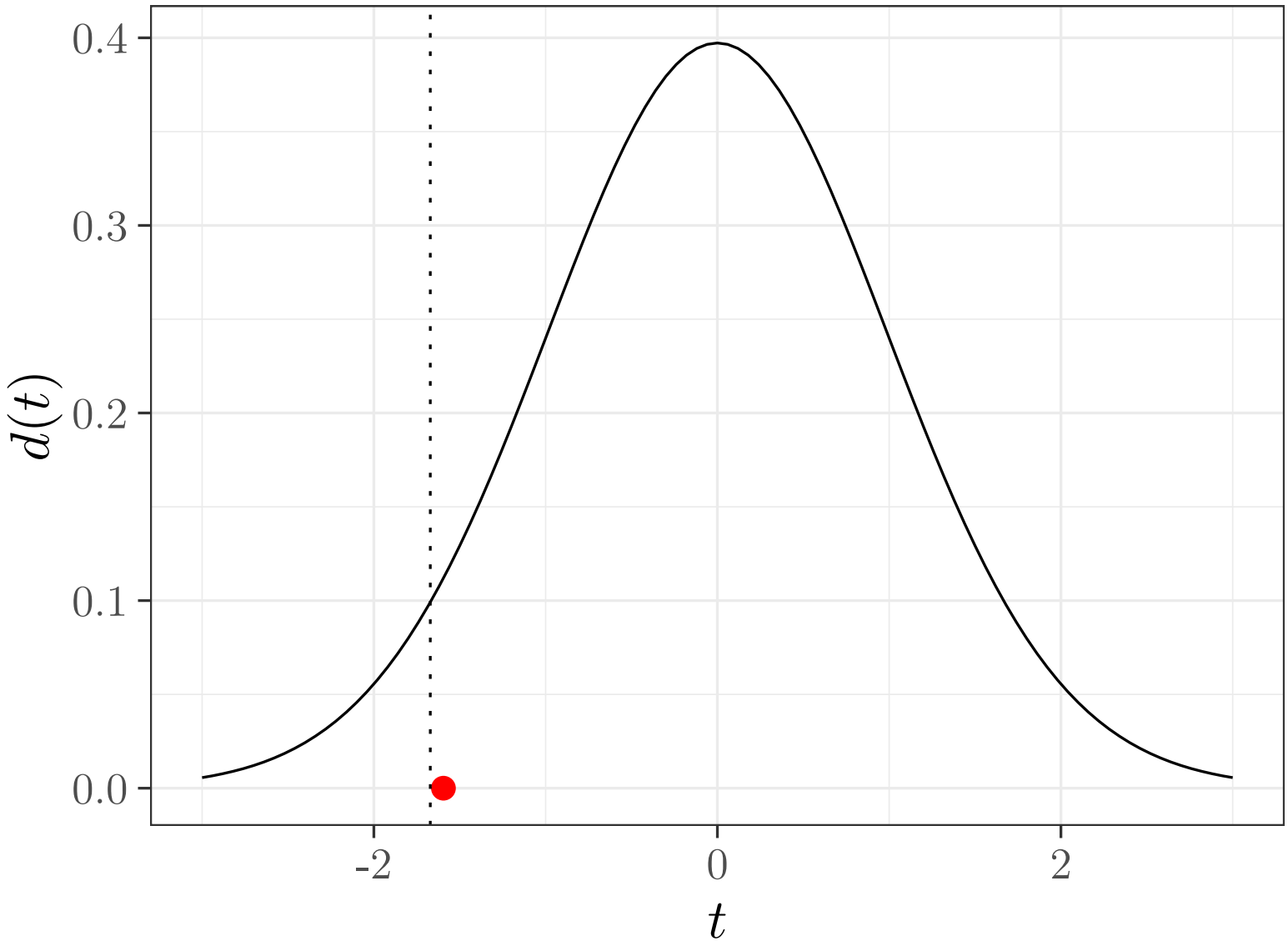
```
points( -1.59, 0, pch = 3, col = 'red')
```

```
points( -1.59, 0, pch = 3, col = 'red')
```



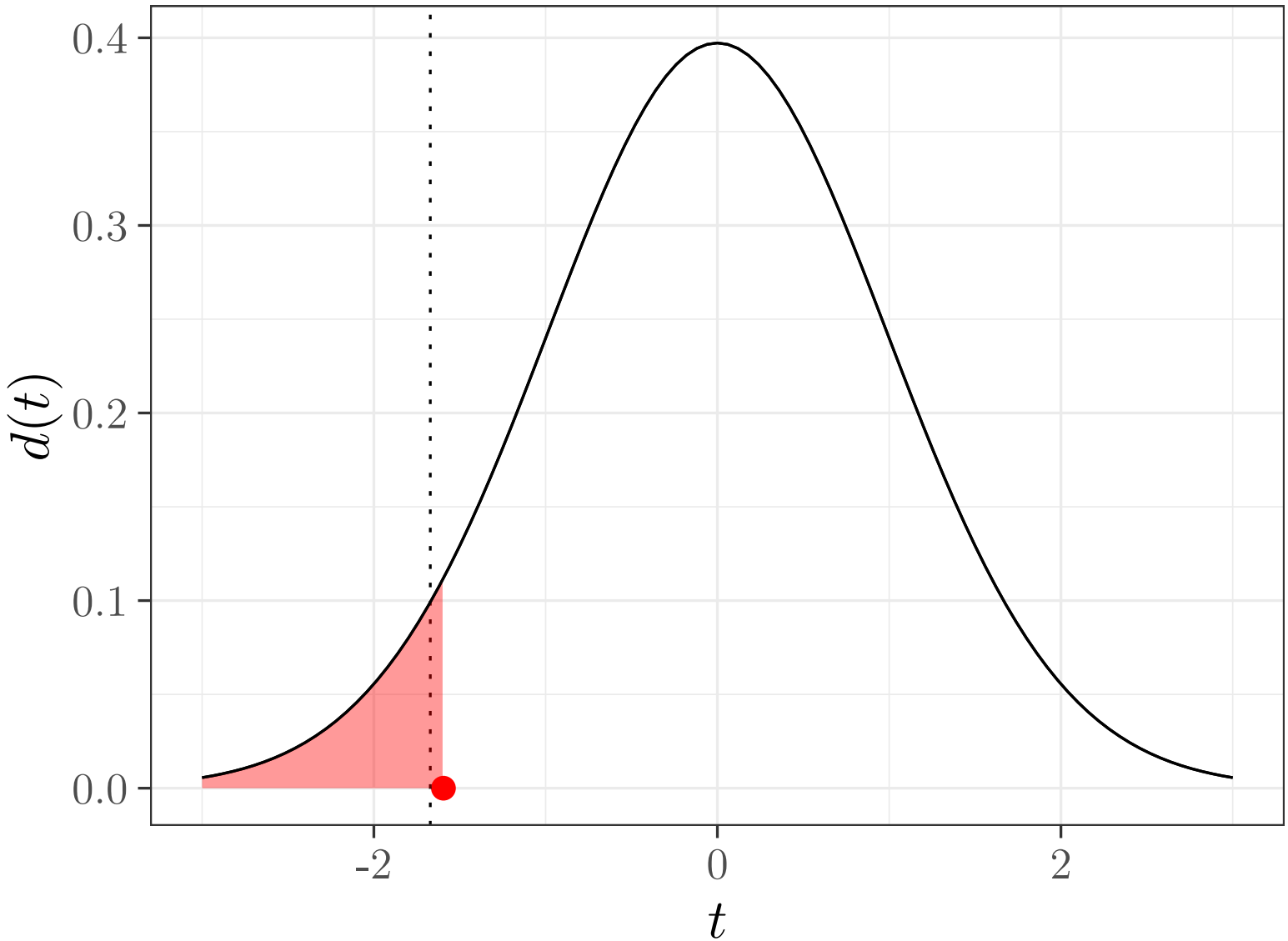
```
abline( v = qt( .05, 58 ), lty = 3 )
```

```
abline( v = qt( .05, 58 ), lty = 3 )
```



```
pt( -1.59, 58 )
```

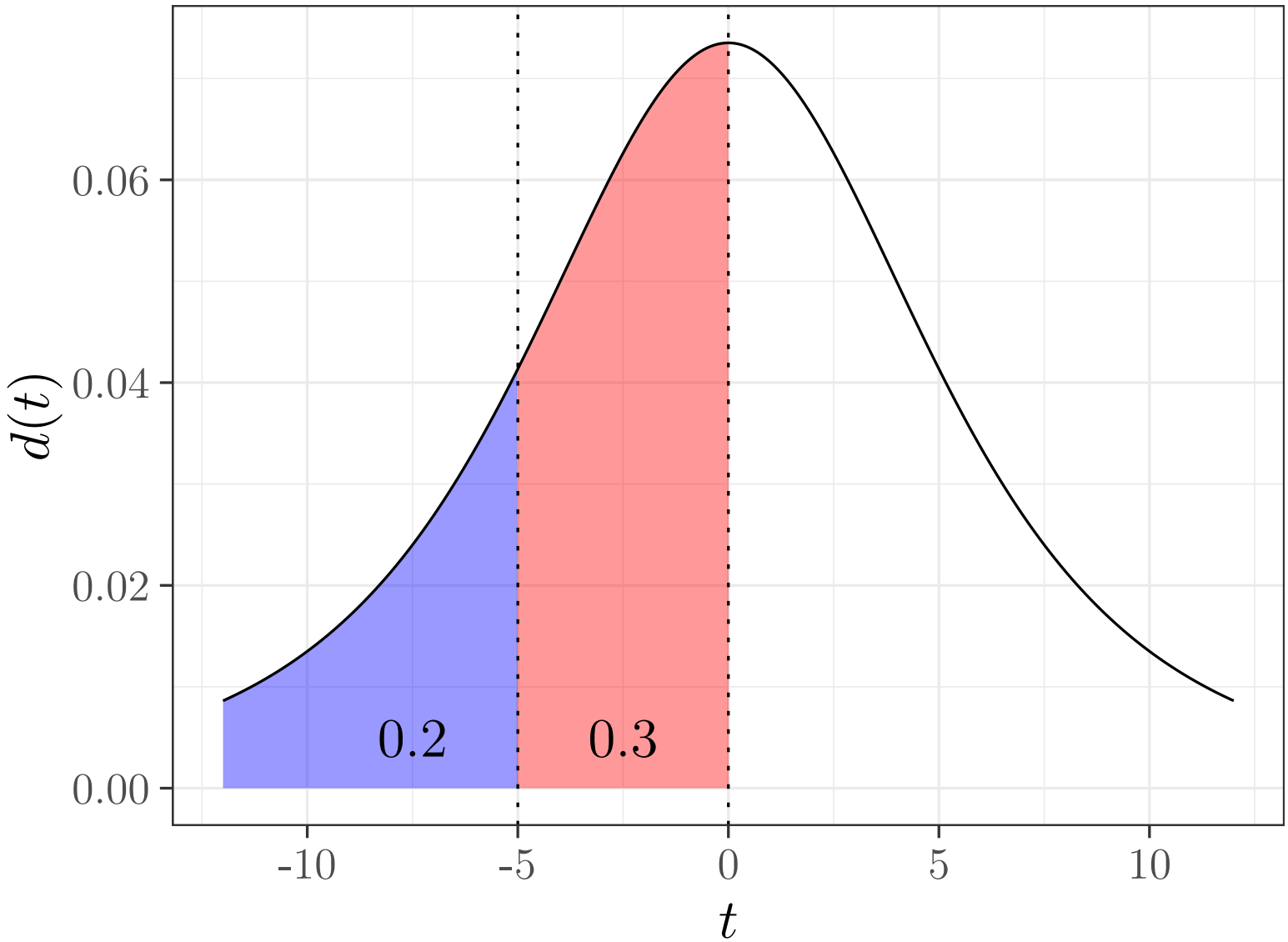
```
pt( -1.59, 58 )
```



Esempio 1: t-test bayesiano

- Supponiamo che il trattamento, per essere efficace, deve produrre una differenza di almeno 5 punti.
- In altre parole, una differenza compresa nell'intervallo $[-5, 0]$ è una differenza sostanzialmente nulla.
- Quindi, scegliamo come prior per la differenza una $t(3, 0, 5)$.
- Sulla base di questa prior, abbiamo che $p(-5 < \delta < 0) = 0.3$ e $p(\delta < -5) = 0.2$.

Esempio 1: prior



Esempio 1: t con Stan

```
data {
  int<lower=0> N;
  vector[N] y;
  vector[N] x;
}
parameters {
  real beta0;
  real beta1;
  real<lower=0> sigma;
}
```

Esempio 1: t con Stan

```
data {
  int<lower=0> N;
  vector[N] y;
  vector[N] x;
}
parameters {
  real beta0;
  real beta1;
  real<lower=0> sigma;
}
transformed parameters {
  vector[N] mu;
  mu = beta0 + beta1 * x;
}
model {
  target += student_t_lpdf( beta0 | 3, 0, 10 );
  target += student_t_lpdf( beta1 | 3, 0, 5 );
  target += student_t_lpdf( sigma | 3, 0, 10 );
  target += normal_lpdf( y | mu, sigma);
}
```

Esempio 1: t con Stan

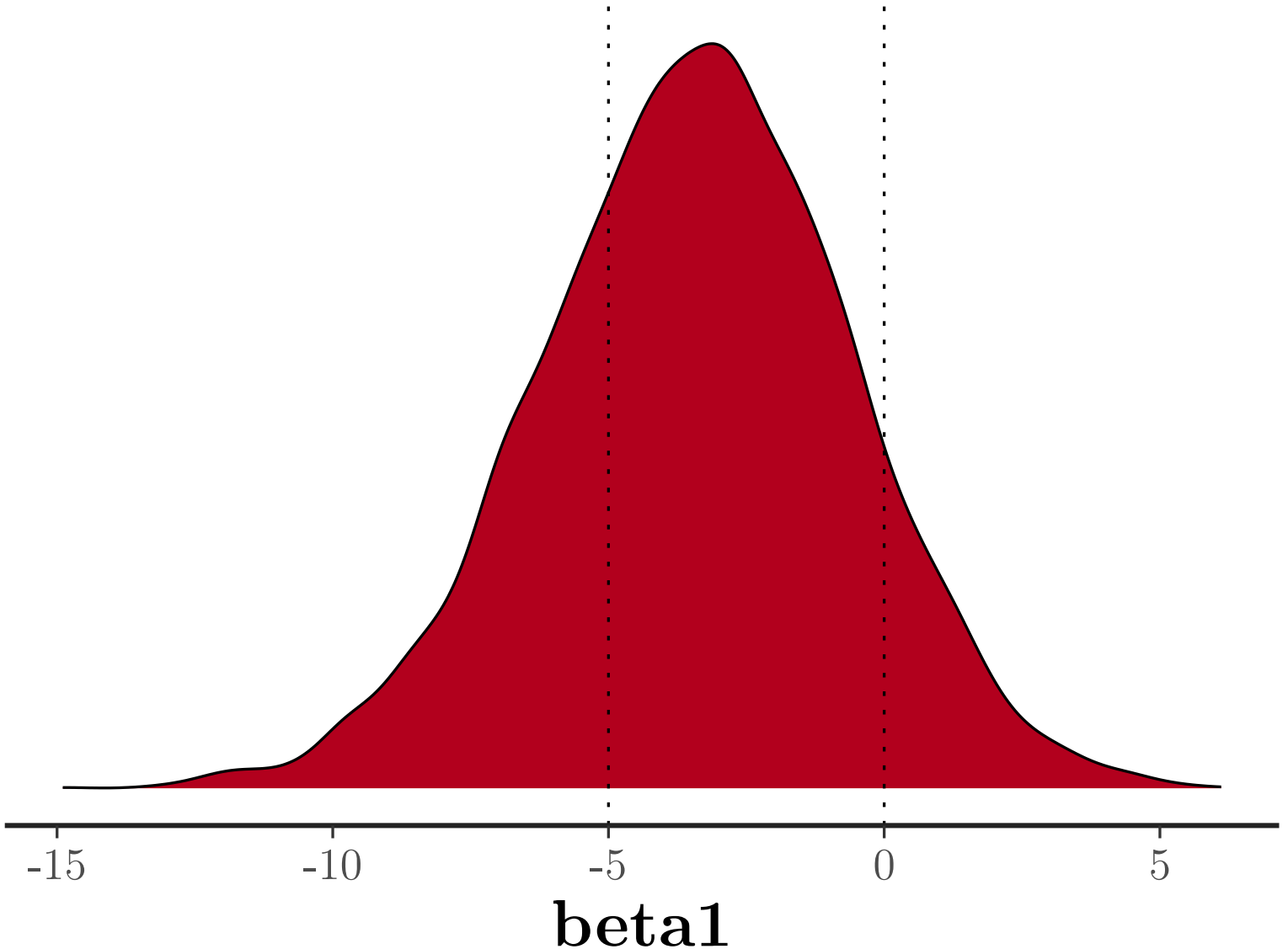
```
> library(rstan)
> fit1 <- stan( ... , data = dataList, seed = 3, iter = 3000 )
> print( fit, pars = c('beta0','beta1' ),
+         probs = c( .025, .5, .975 ) )
```

Esempio 1: t con Stan

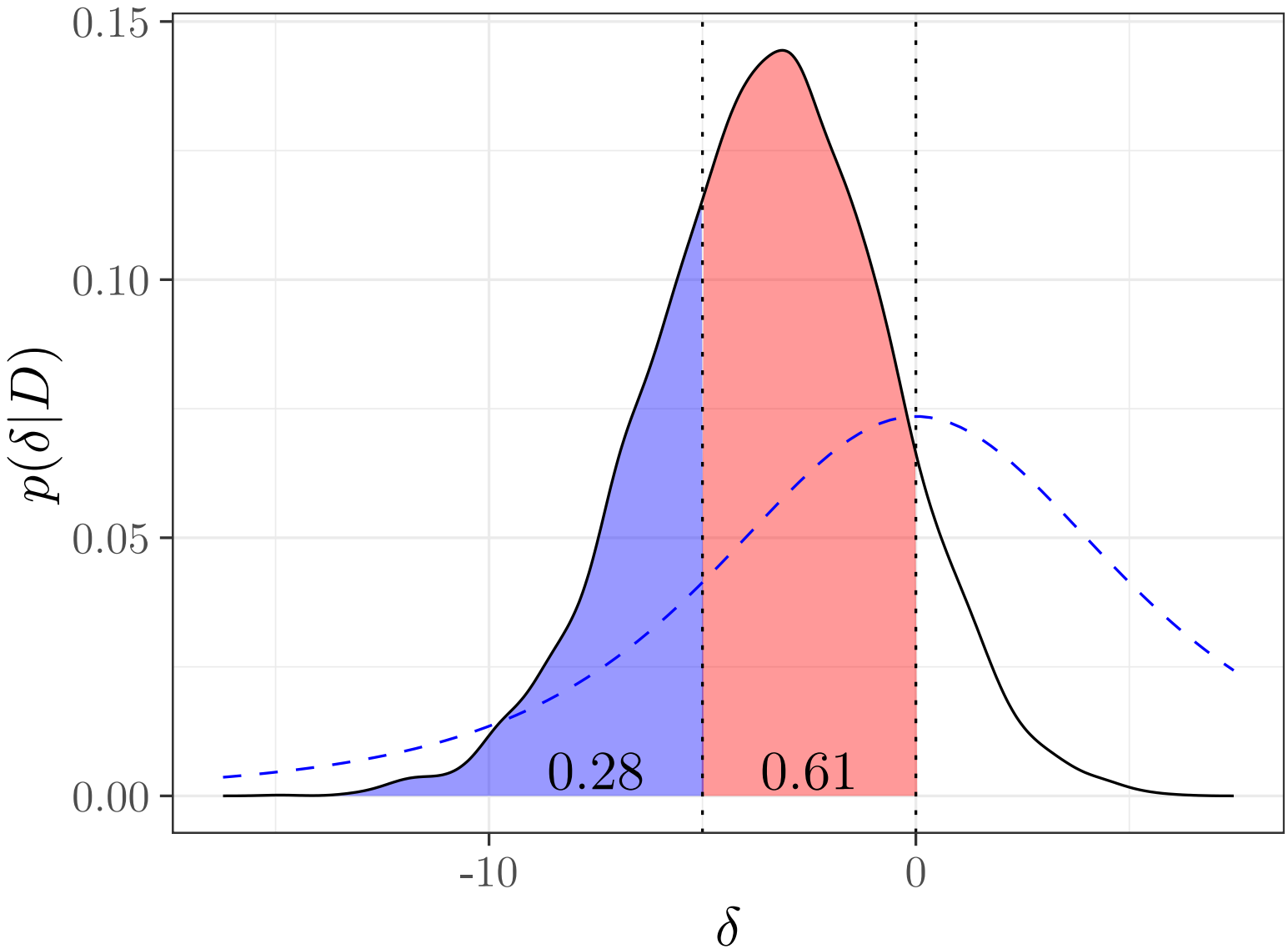
```
> library(rstan)
> fit1 <- stan( ... , data = dataList, seed = 3, iter = 3000 )
> print( fit, pars = c('beta0','beta1' ),
+         probs = c( .025, .5, .975 ) )
```

```
Inference for Stan model: anon_model.
4 chains, each with iter=3000; warmup=1500; thin=1;
post-warmup draws per chain=1500, total post-warmup draws=6000.

       mean se_mean   sd  2.5%   50% 97.5% n_eff Rhat
beta0 79.79    0.04 2.14 75.57 79.80  84.0  3468    1
beta1 -3.45    0.05 2.81 -9.17 -3.39   1.9  3556    1

Samples were drawn using NUTS(diag_e) at Sat Mar 29 09:26:22 202
For each parameter, n_eff is a crude measure of effective sample
and Rhat is the potential scale reduction factor on split chains
convergence, Rhat=1).
```

```
plot( fit1, pars = "beta1", plotfun = "dens" ) +
  geom_vline( xintercept = ROPE, lty = 3 )
```



```
plot( fit1, pars = "beta1", plotfun = "dens" ) +
  geom_vline( xintercept = ROPE, lty = 3 )
```



Uno studio clinico

Esempio 2: Ortoressia

Abstract

Purpose Orthorexia Nervosa (ON) exhibits specific features that may overlap with Obsessive-Compulsive Disorder (OCD), Perfectionism Striving (PS), and Perfectionism Concern (PC). While previous literature has shown predictiveness in different characteristics of ON, this study aimed to determine if PS, PC and OCD symptoms could predict ON dimensions in at-risk populations using Bayesian models.

Method The study enrolled 622 individuals from three different at-risk populations: people who were following treatment for an Eating Disorder (Patients), people who were following a diet (Dieters) and University students with a degree in medicine or nursing (Students).

Results Using Bayesian probabilistic methods and considering group factors, the model was enhanced highlighting that ON characterised Patients, Dieters, and Students. The predictiveness of OC features, PS, and PC in different ON characteristics was confirmed in each group, and different patterns were observed in the three groups. Regarding problems related to ON, predictors were higher in Patients, followed by Dieters and Students. Posterior Predictive Distribution (PPD) showed that almost 50% of Patients incur ON-related problems. In ON knowledge, Patients and Dieters are very similar. When feelings related to ON were considered, Patients and Dieters showed different interactions.

Conclusion Clinicians should consider that one in two patients suffering from EDs might develop ON-related problems. People on a diet could show similar patterns of symptoms to patients in ON knowledge and feelings. Finally, our results confirm that perfectionism represents a risk factor for ON in each group considered.

Keywords Orthorexia nervosa, Perfectionism, Obsessive-compulsive symptoms, Risk factors

Novara, C., Maggio, E., Pastore, M., Piasentin, S., Pardini, S., Mattioli, S. (2025). What is more likely in orthorexia nervosa: perfectionism or OC symptoms? A Bayesian method in clinical and non-clinical samples. *BMC Psychology*, 13, 230. doi.org/10.1186/s40359-025-02517-2.

I dati

- Il set di dati include 318 soggetti $\times$ 7 variabili:
  - Tre variabili endogene:
    EHQ_PROBLEMI, EHQ_CONVINZIONI,
    EHQ_EMOZIONI
  - Quattro variabili esogene:
    GROUP, MPS_POS, MPS_NEG, OCI_TOT
- In particolare ci sono tre gruppi di soggetti:
  clinical (144), diet (85), normal (89).

Le ipotesi

EHQ – item a 4 punti:

- Scala Problemi (12 items) Per i clinici attendiamo punteggi superiori ai 32 punti e 4/5 punti sopra i soggetti degli altri due gruppi.
- Scala Convinzioni (5 items) Per i clinici ed i diet attendiamo punteggi superiori ai 10 punti e circa 2 punti sopra i normali.
- Scala Emozioni (4 items). Per i clinici ed i diet attendiamo punteggi superiori a 8 punti e circa 2 punti sopra i normali.

Le ipotesi

EHQ – item a 4 punti:

- Scala Problemi (12 items) Per i clinici attendiamo punteggi superiori ai 32 punti e 4/5 punti sopra i soggetti degli altri due gruppi.
- Scala Convinzioni (5 items) Per i clinici ed i diet attendiamo punteggi superiori ai 10 punti e circa 2 punti sopra i normali.
- Scala Emozioni (4 items). Per i clinici ed i diet attendiamo punteggi superiori a 8 punti e circa 2 punti sopra i normali.

Perfezionismo: 1) Il gruppo clinico dovrebbe avere maggior perfezionismo negativo (`MPS_concern`) rispetto agli altri.
2) Non dovrebbe esserci differenza tra i gruppi nel perfezionismo positivo (`MPS_striving`).
3) Il perfezionismo negativo ha maggiori effetti sulla scala Problemi, il perfezionismo positivo su Convinzioni e Emozioni

I modelli

Utilizzando queste variabili definiamo un insieme di 6 modelli multivariati:

- M00, modello nullo:

$$\boldsymbol{Y} \sim 1$$

- M01, modello delle differenze tra i gruppi:

$$\boldsymbol{Y} \sim \text{GROUP}$$

- M02, effetti additivi GRUPPO + MPS:

$$\boldsymbol{Y} \sim \text{GROUP} + \text{MPS\_striving} + \text{MPS\_concern}$$

- M03, effetti additivi GRUPPO + MPS + OCI:

$$\boldsymbol{Y} \sim \text{GROUP} + \text{MPS\_striving} + \text{MPS\_concern} + \text{OCI\_TOT}$$

- M04, modello con interazione MPS_NEG × OCI:

$$\boldsymbol{Y} \sim \text{GROUP} + \text{MPS\_striving} + \text{MPS\_concern} \times \text{OCI\_TOT}$$

- M05, modello con interazioni rispetto al GRUPPO:

$$\boldsymbol{Y} \sim \text{GROUP} \times \text{MPS\_striving} + \text{GROUP} \times \text{MPS\_concern} + \text{GROUP} \times \text{OCI\_TO}$$

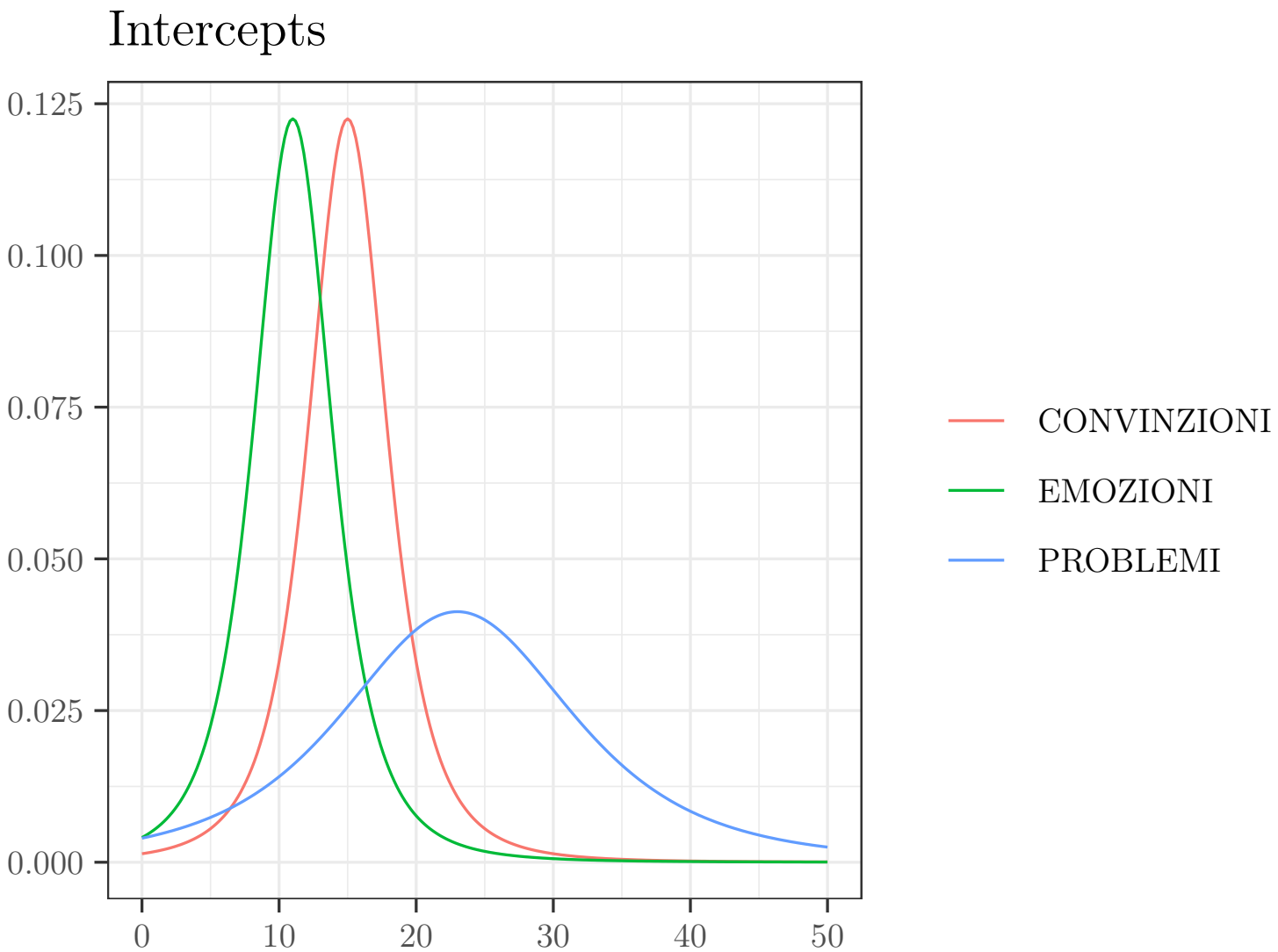
in cui $\boldsymbol{Y}$ rappresenta la matrice delle variabili endogene EHQ: PROBLEMI, CONVINZIONI, EMOZIONI. Ciascuno di questi modelli viene specificato con i residui correlati.

Prior specification

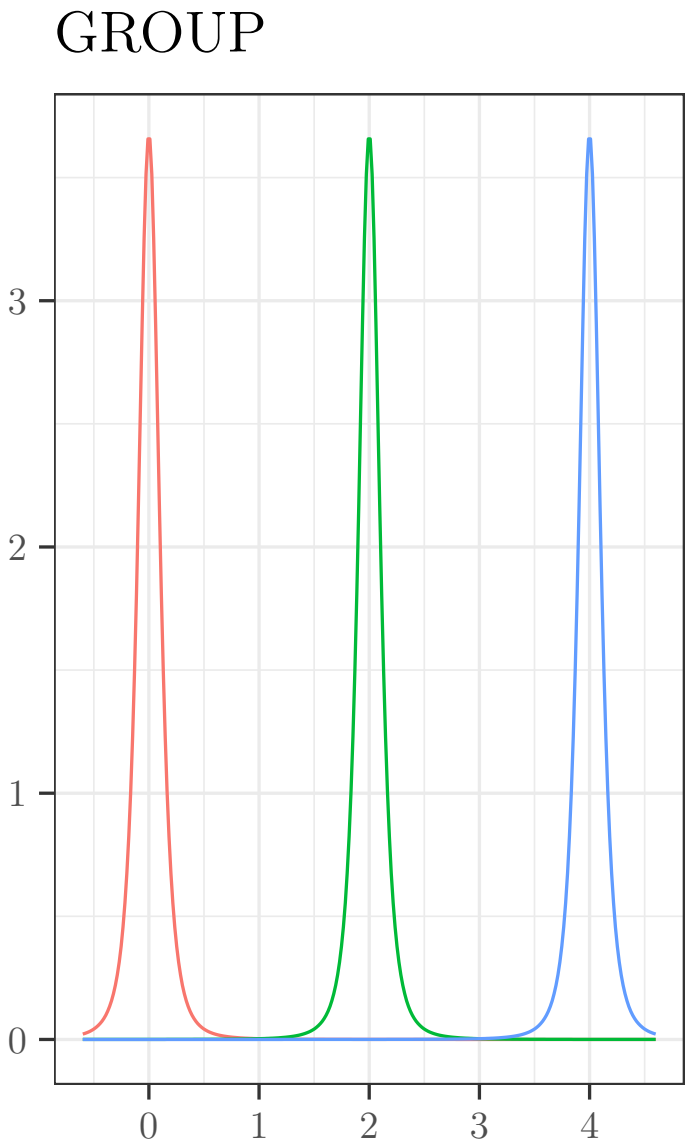
- Il numero di parametri in questi modelli varia da un minimo di 9 - le intercette e le deviazioni standard e le correlazioni dei residui nel modello nullo, M00 - ad un massimo di 42 nel modello che include le interazioni con il GRUPPO, M05.

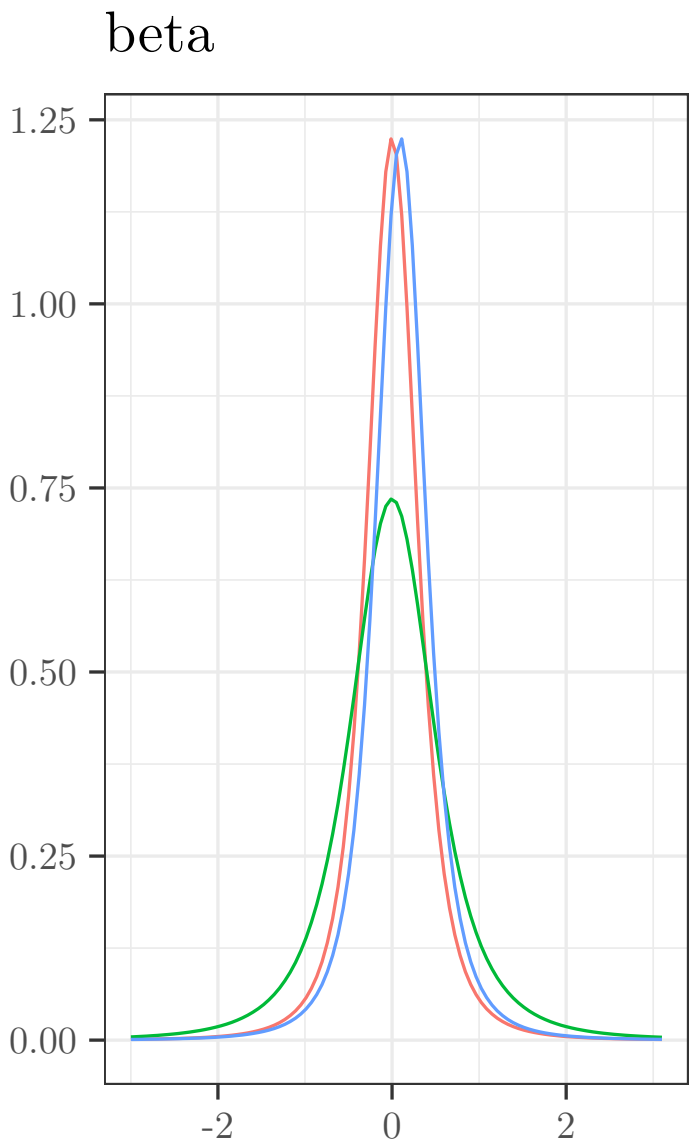
- Per ogni parametro dobbiamo definire una prior; in particolare utilizzeremo delle Student's t per le intercette, i coefficienti di regressione e le deviazioni standard, una LKJ (Lewandowski et al, 2009) per le correlazioni tra residui.

Prior specification

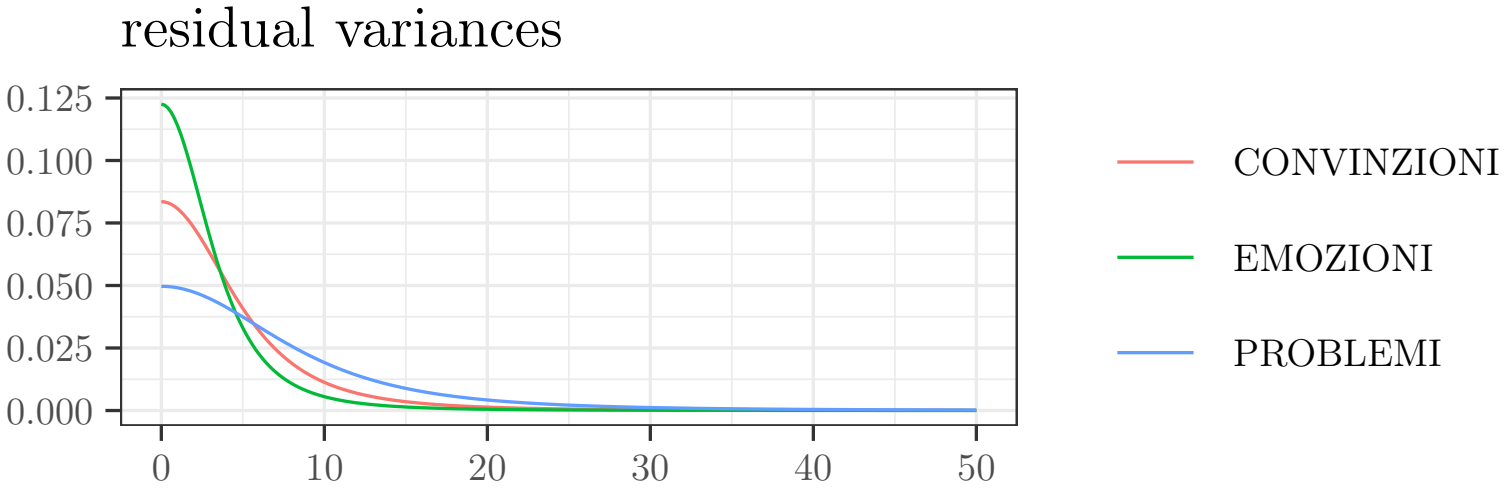


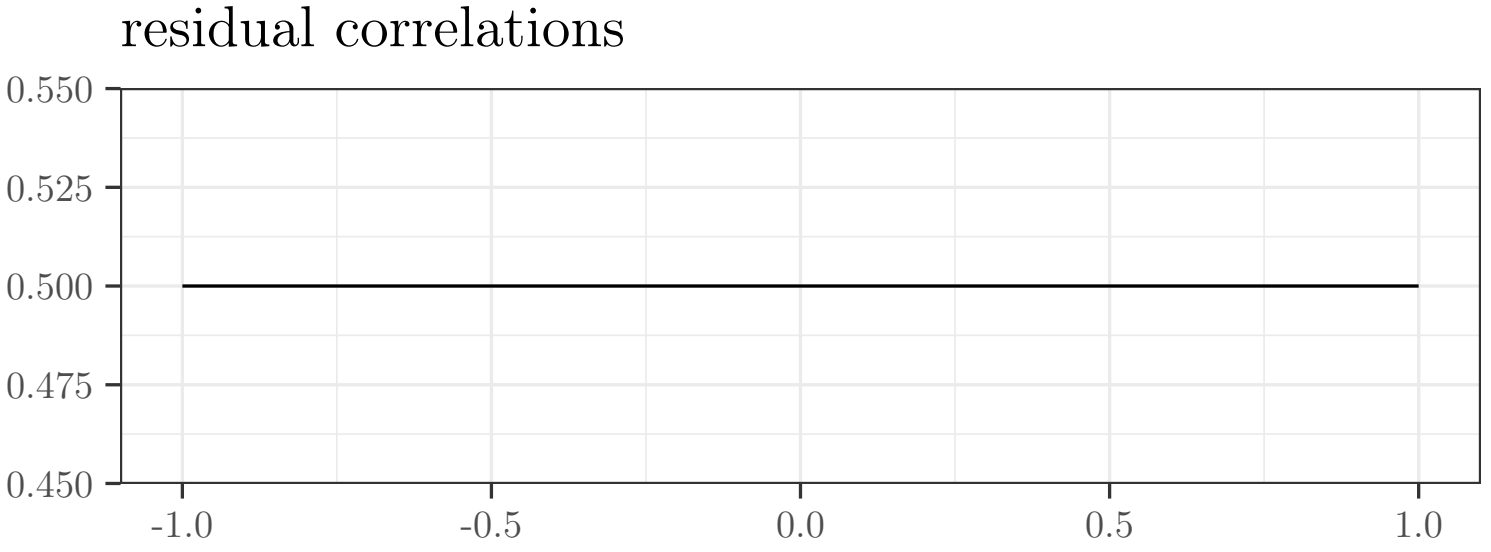
Prior specification





Prior specification

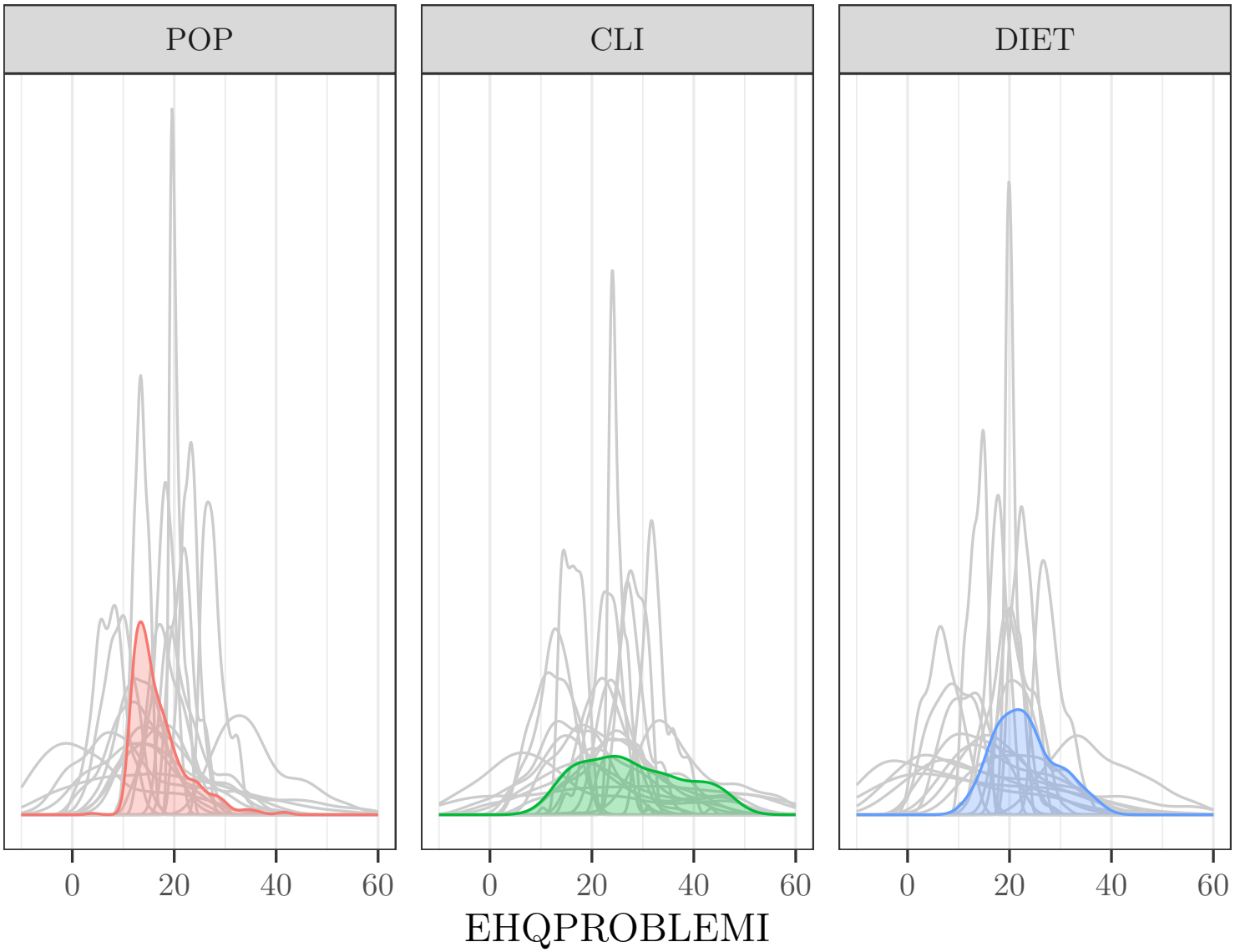




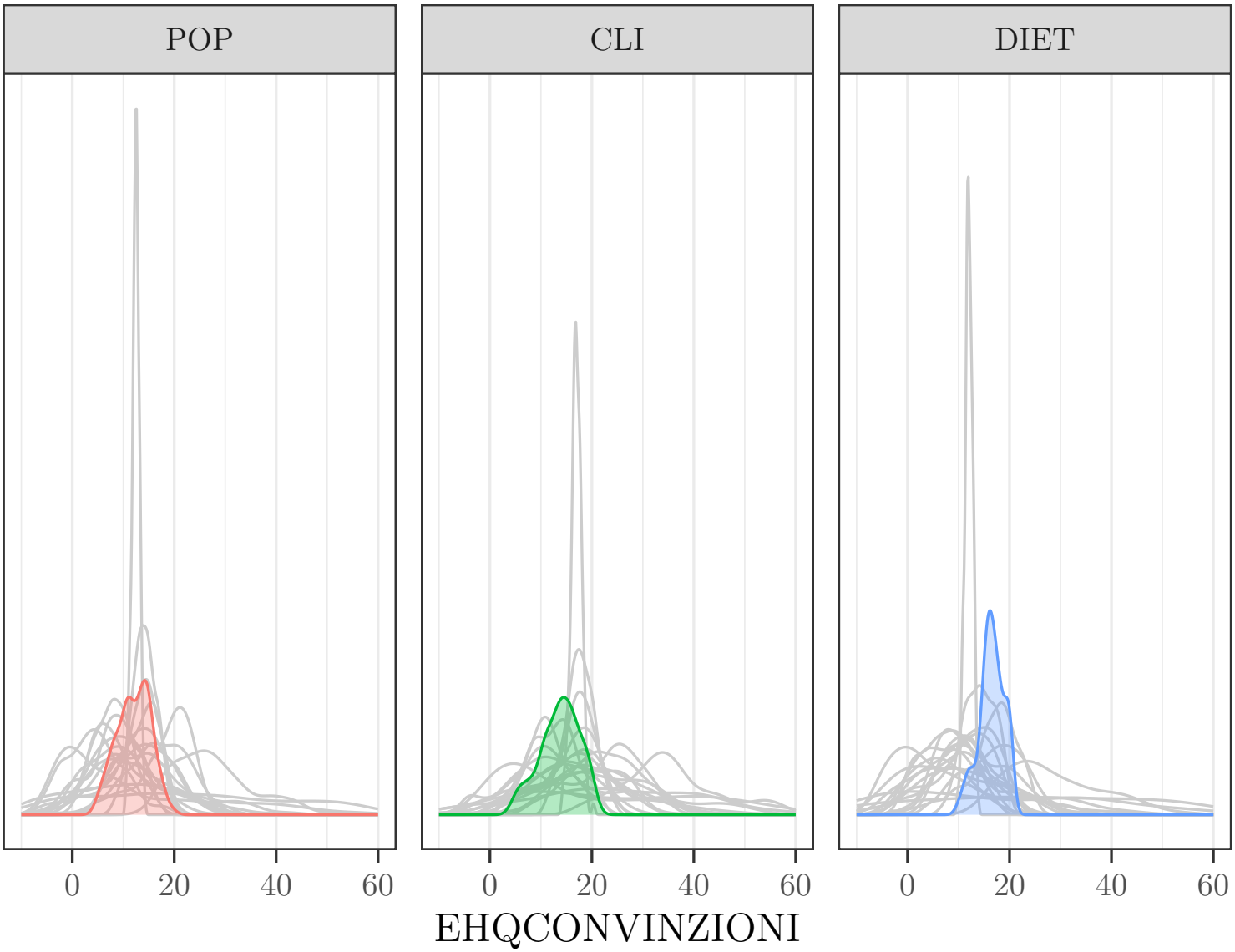
Prior Predictive Check

```
fit_prior <- brm(
      bf( EHQ_PROBLEMI ~ GROUP + MPS_concern +
            MPS_striving + OCI_TOT ) +
      bf( EHQ_CONVINZIONI ~ GROUP + MPS_concern +
            MPS_striving + OCI_TOT ) +
      bf( EHQ_EMOZIONI ~ GROUP + MPS_concern +
            MPS_striving + OCI_TOT ) +
        set_rescor( TRUE ),
  data = Z, prior = myPrior,
        sample_prior = "only", seed = 1 )
```

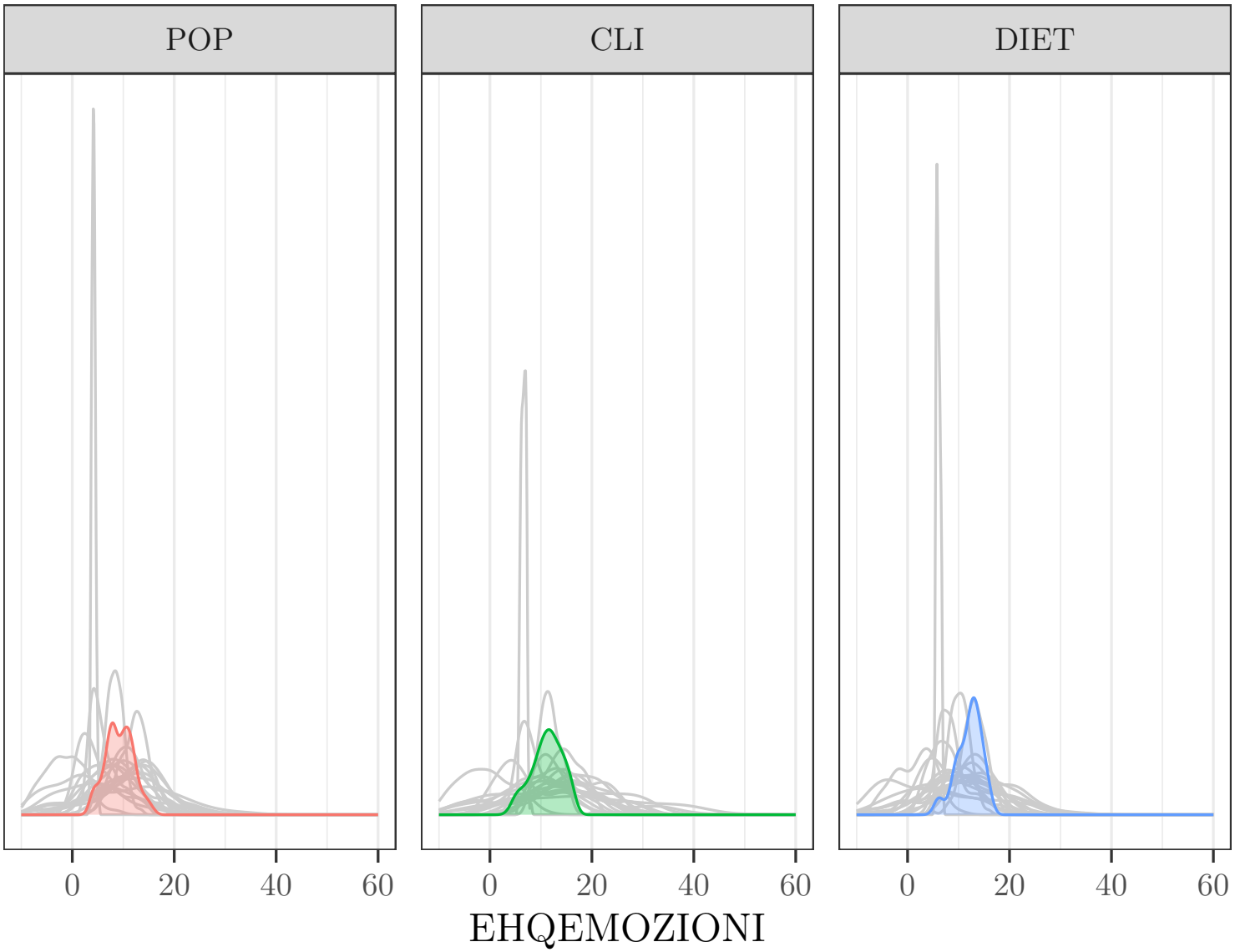
Prior Predictive Check



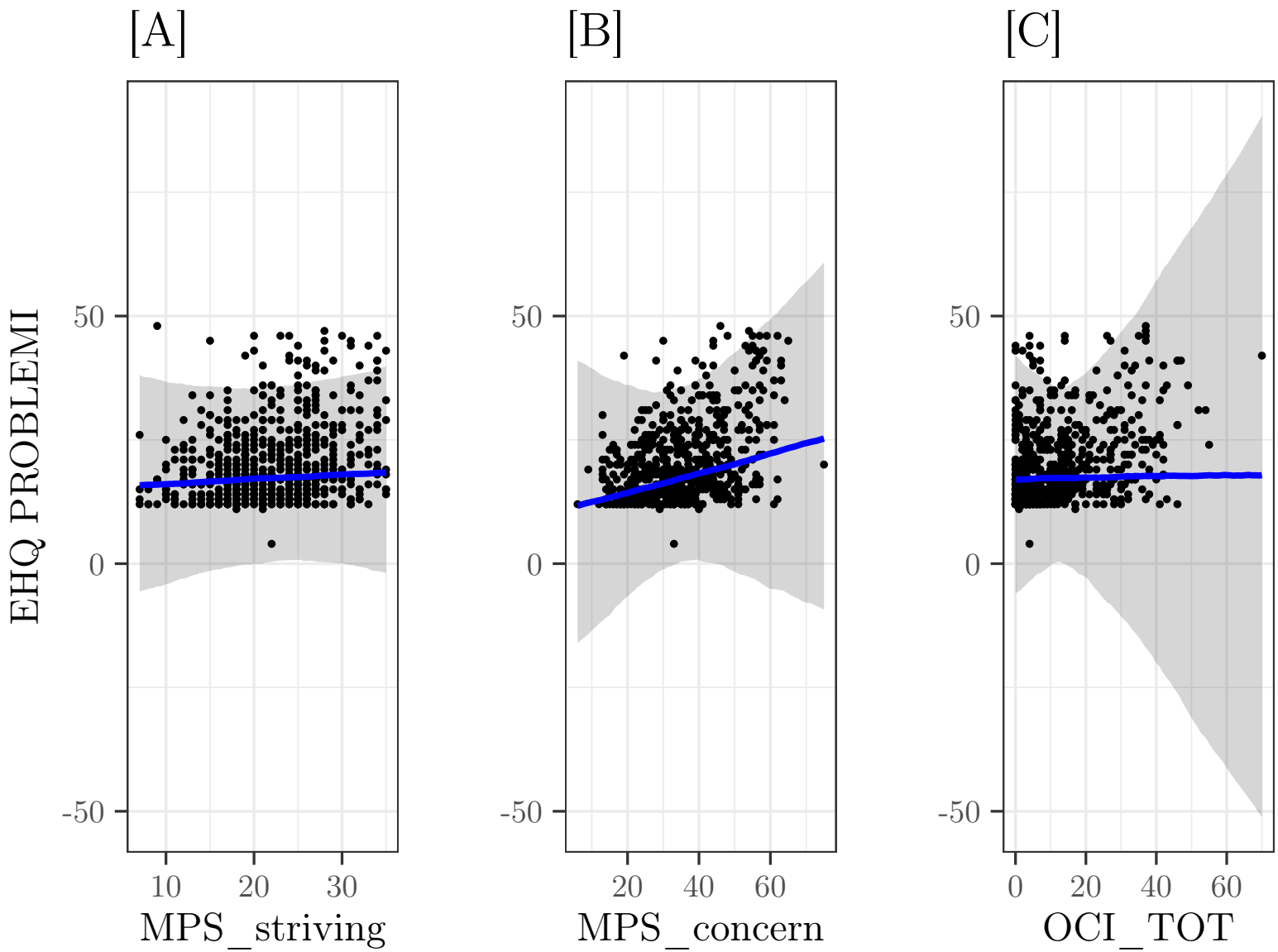
Prior Predictive Check



Prior Predictive Check



Prior Predictive Check



Modello finale (solo su PROBLEMI)

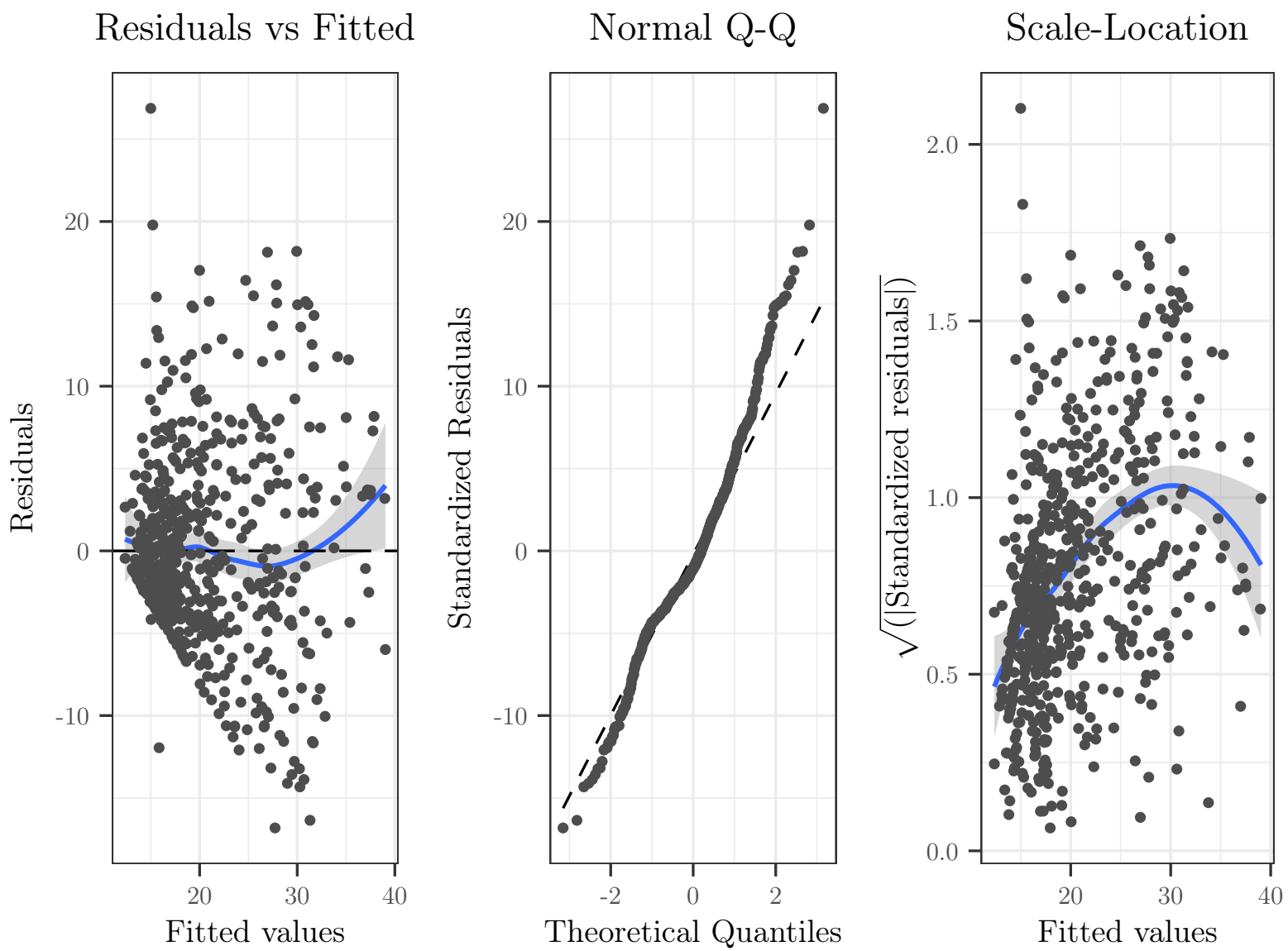
Le prior

	prior	class	coef
26	student_t(3, 4, 0.1)	b	GROUPCLI
27	student_t(3, 0, 0.5)	b	GROUPCLI:MPS_concern
28	student_t(3, 0, 0.5)	b	GROUPCLI:MPS_striving
29	student_t(3, 0, 0.5)	b	GROUPCLI:OCI_TOT
30	student_t(3, 0, 0.1)	b	GROUPDIET
31	student_t(3, 0, 0.5)	b	GROUPDIET:MPS_concern
32	student_t(3, 0, 0.5)	b	GROUPDIET:MPS_striving
33	student_t(3, 0, 0.5)	b	GROUPDIET:OCI_TOT
34	student_t(3, 0.2, 0.3)	b	MPS_concern
35	student_t(3, 0.1, 0.3)	b	MPS_striving
36	student_t(3, 0, 0.5)	b	OCI_TOT
39	student_t(3, 18, 7.4)	Intercept	
43	student_t(3, 0, 7.4)	sigma	

Modello finale (solo su PROBLEMI)

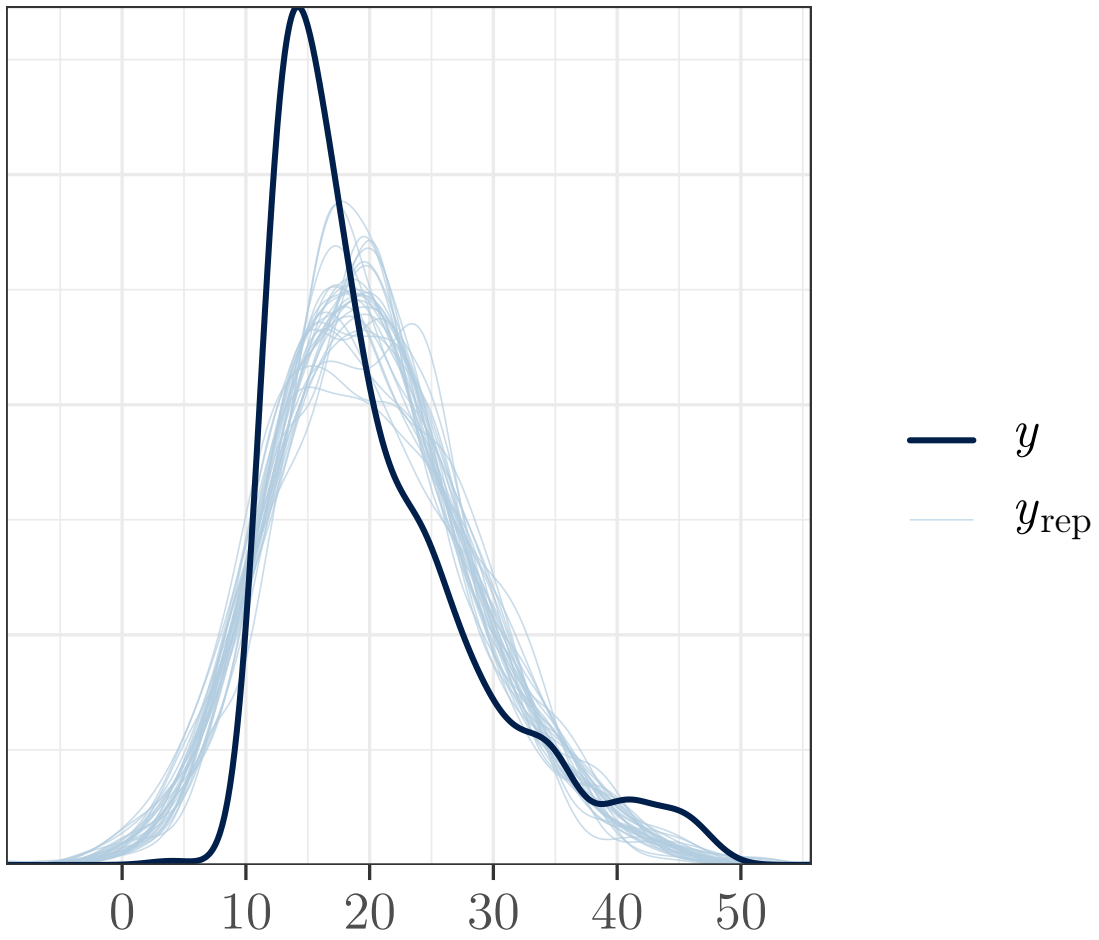
```
fit2 <- brm( EHQ_PROBLEMI ~ GROUP * MPS_concern +
               GROUP * MPS_striving + GROUP * OCI_TOT,
             data  = Z, prior = myPrior, cores = 4  )
```

Diagnostics

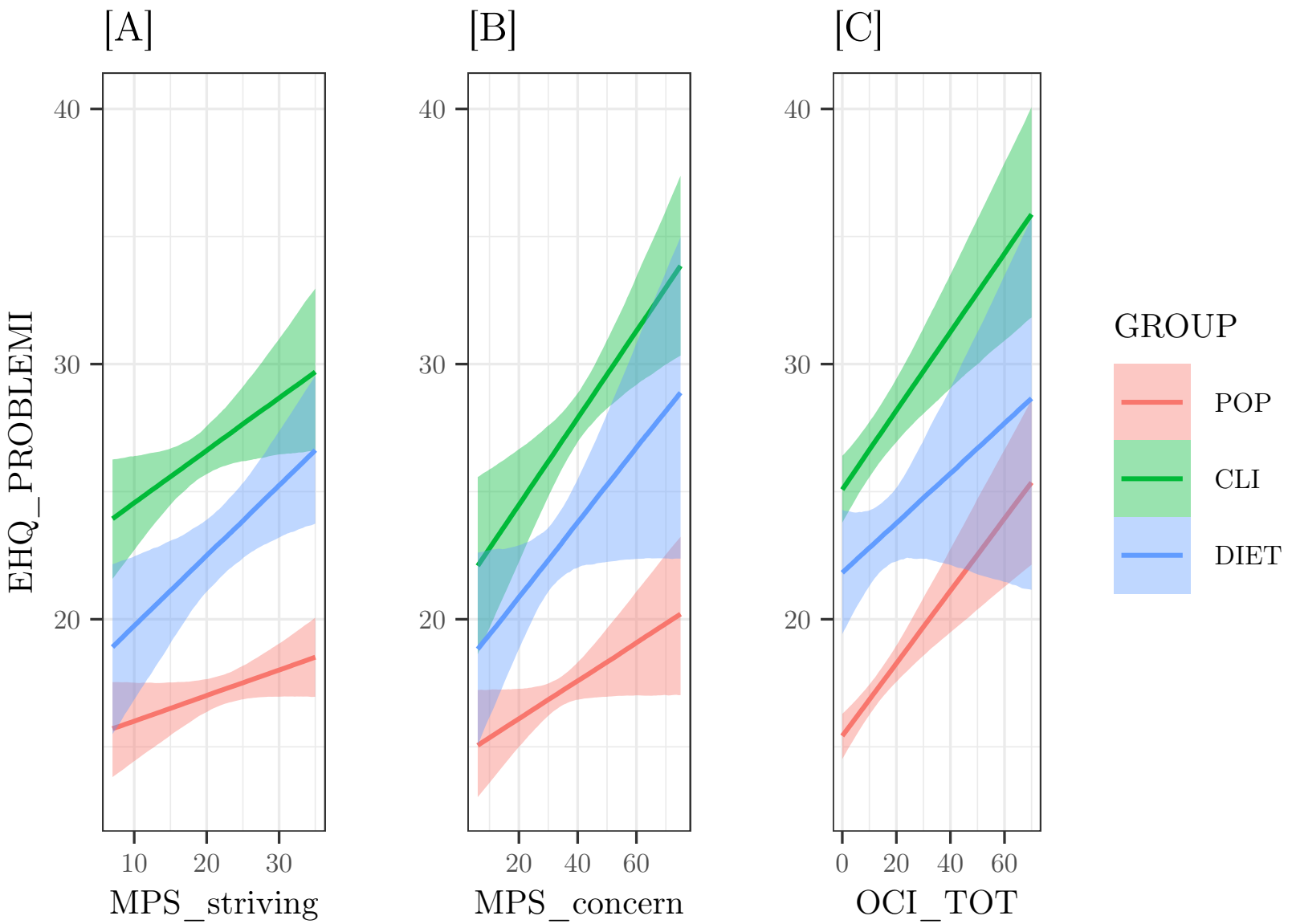


Posterior Predictive Check

```
pp_check( fit2, ndraws = 30 )
```



Model predictions



Analisi delle ipotesi

```
ll <- 15

newData <- expand.grid (
 MPS_striving =
  seq(min(Z$MPS_striving), max(Z$MPS_striving), length=ll),
 MPS_concern =
  seq(min(Z$MPS_concern), max(Z$MPS_concern), length=ll),
 OCI_TOT =
  seq(min(Z$OCI_TOT), max(Z$OCI_TOT), length=ll),
 GROUP = levels(Z$GROUP)
)

PP <- posterior_predict( fit2, ndraws = NDRAWS,
                         newdata = newData )
```

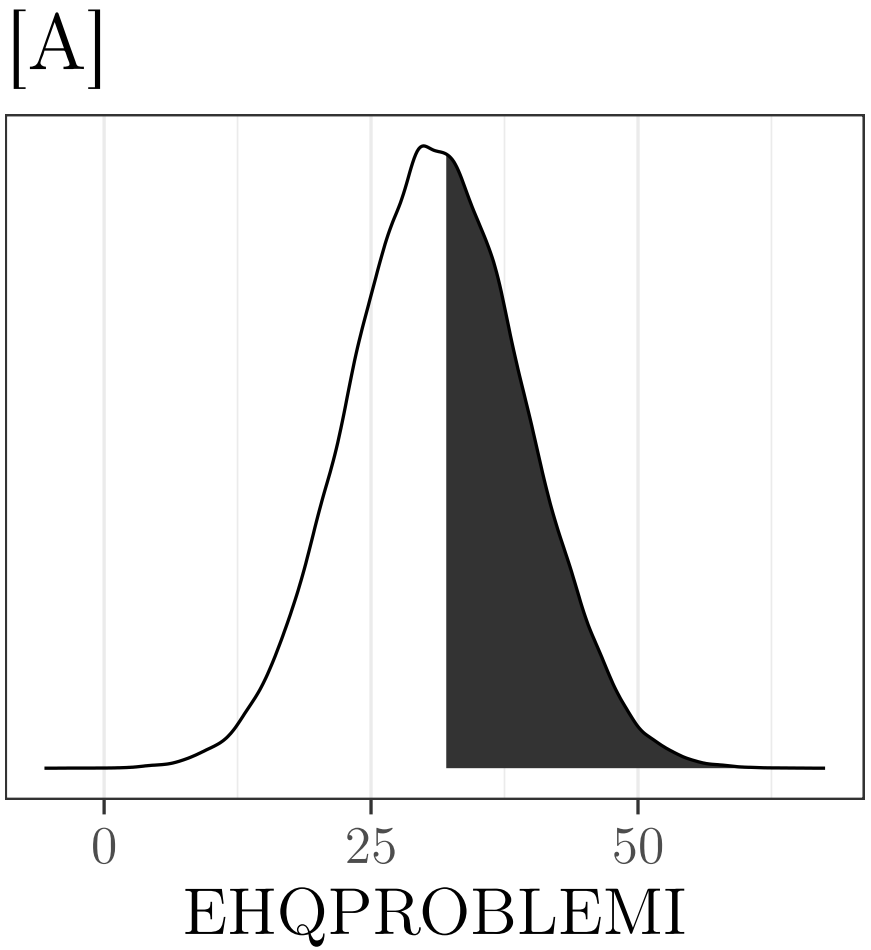
[...]

Analisi delle ipotesi

- Per i clinici attendiamo punteggi superiori ai 32 punti.

Analisi delle ipotesi

- Per i clinici attendiamo punteggi superiori ai 32 punti.
- *Posterior predictive distribution* dei punteggi del gruppo clinico:
  $Pr(x > 32) = 0.46$

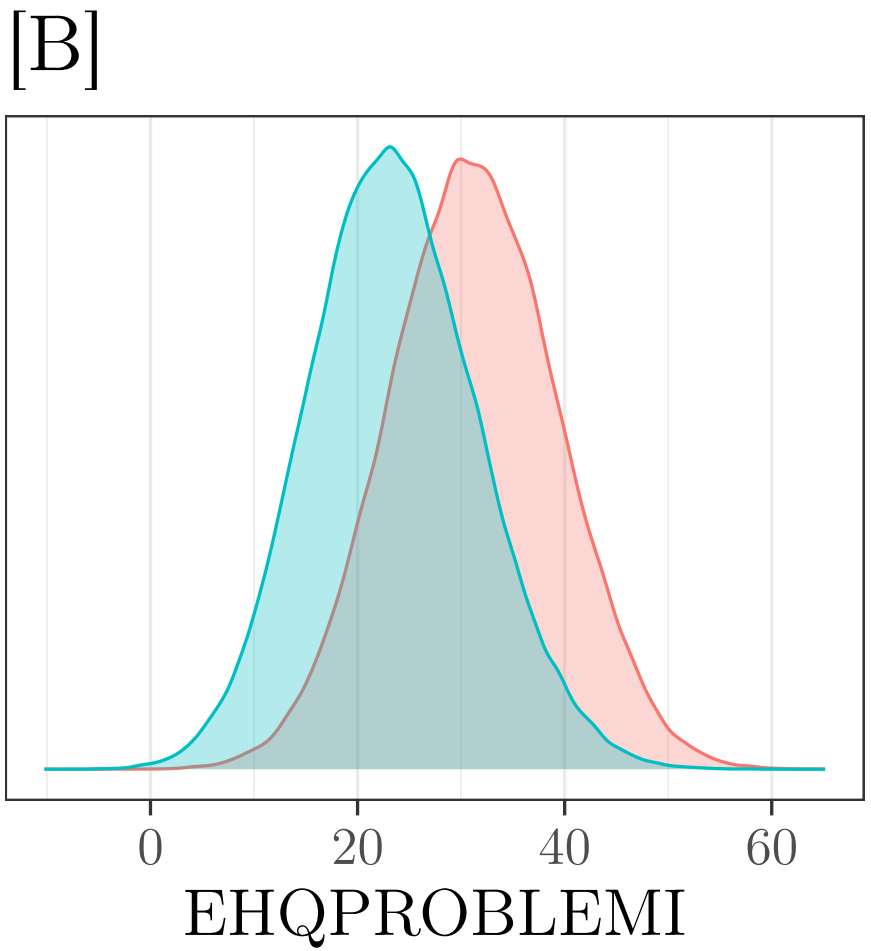


Analisi delle ipotesi

- Per i clinici attendiamo punteggi superiori ai soggetti degli altri due gruppi.

Analisi delle ipotesi

- Per i clinici attendiamo punteggi superiori ai soggetti degli altri due gruppi.
- *Posterior predictive distributions* dei punteggi del gruppo clinico (in rosso) vs gli altri due gruppi: $Pr(x_{\text{CLI}} > x_{\text{others}}) = 0.25$



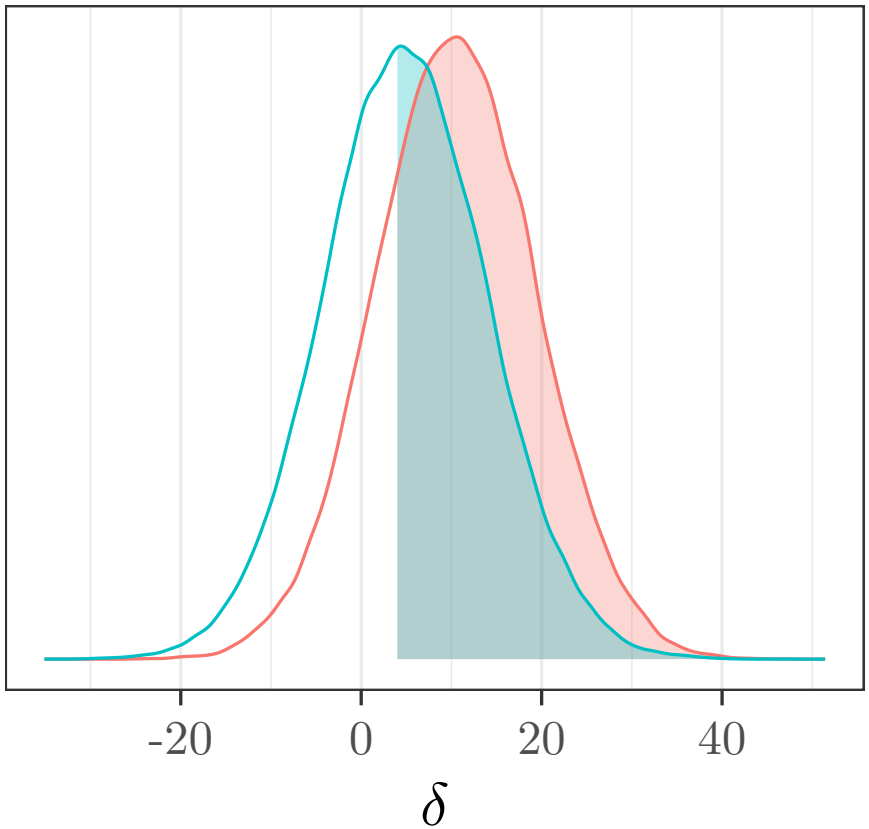
Analisi delle ipotesi

- Per i clinici attendiamo punteggi superiori di circa 4 punti rispetto a quelli degli altri due gruppi.

Analisi delle ipotesi

- Per i clinici attendiamo punteggi superiori di circa 4 punti rispetto a quelli degli altri due gruppi.
- *Posterior predictive distributions* delle differenze tra i punteggi CLI-POP (in rosso) e tra i punteggi CLI-DIET (in azzurro); $Pr(\delta_{\text{CLI-POP}} > 4) = 0.77$, $Pr(\delta_{\text{CLI-DIET}} > 4) = 0.54$

[C]

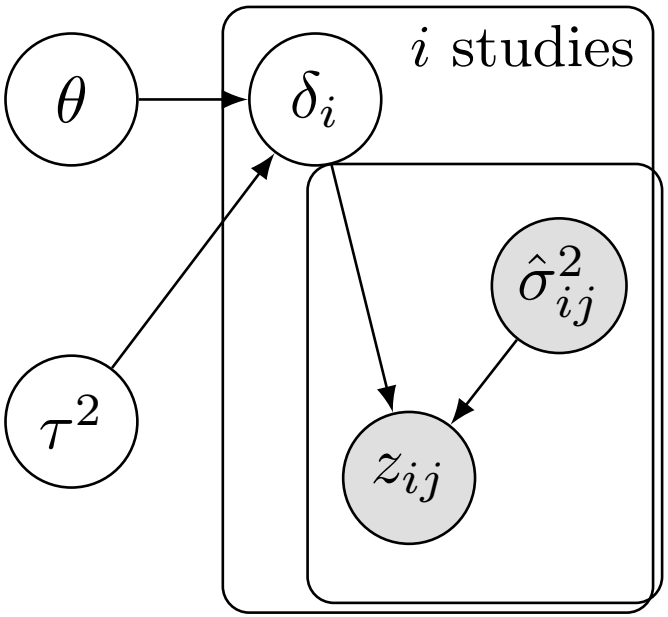


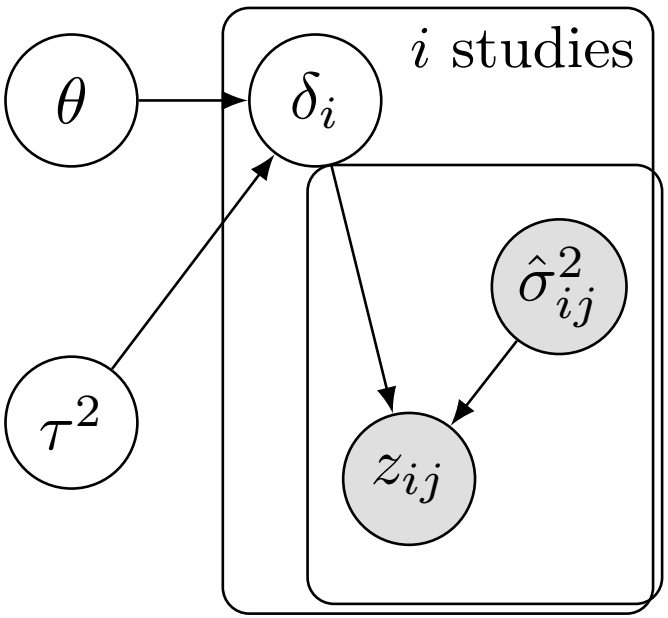
Meta analisi

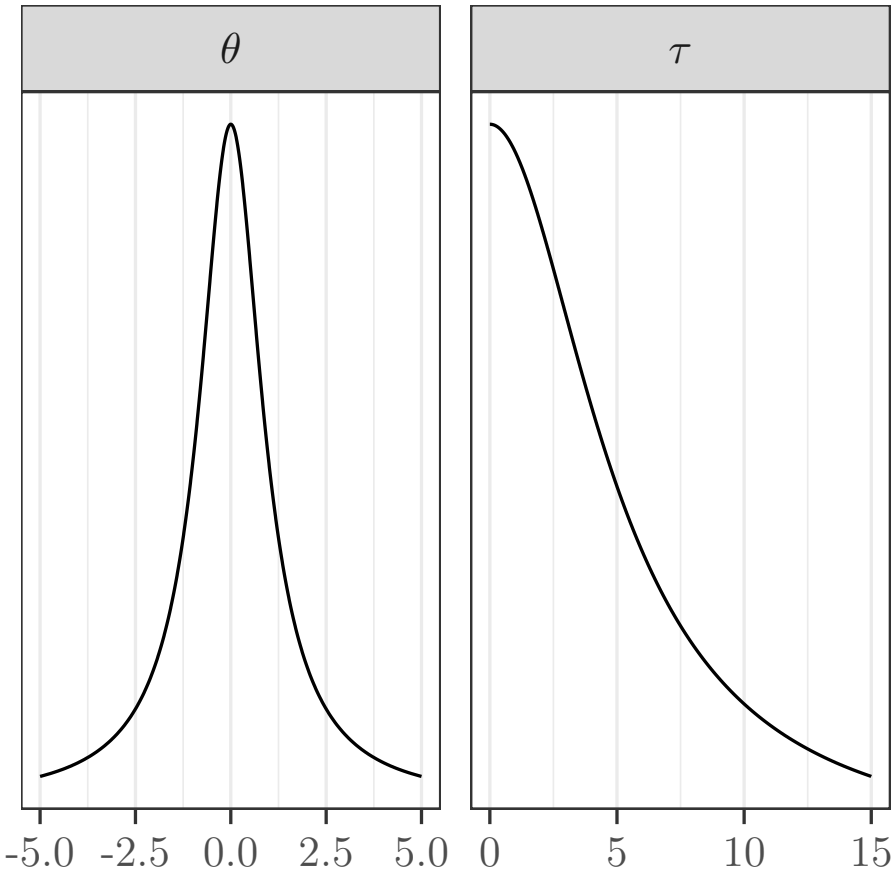
Abstract

In two Bayesian meta-analyses, we investigated associations between Sensory Processing Sensitivity (SPS) and the Big Five personality traits (MA1) as well as both Positive and Negative Affect (MA2). Moderators were age and the three SPS subscales. In MA1 (8 papers, 6790 subjects), SPS in children correlated with Neuroticism ($r = 0.42$) but did not with Extraversion, Openness, Agreeableness or Conscientiousness. In adults, SPS correlated with Openness ($r = 0.14$) and Neuroticism ($r = 0.40$) but did not with Extraversion, Agreeableness or Conscientiousness. In MA2 (19 papers, 5326 subjects), SPS in children correlated with Negative ($r = 0.29$) and Positive Affect ($r = 0.21$), but only with Negative Affect ($r = 0.34$) in adults. Developmental and measurement aspects are discussed.

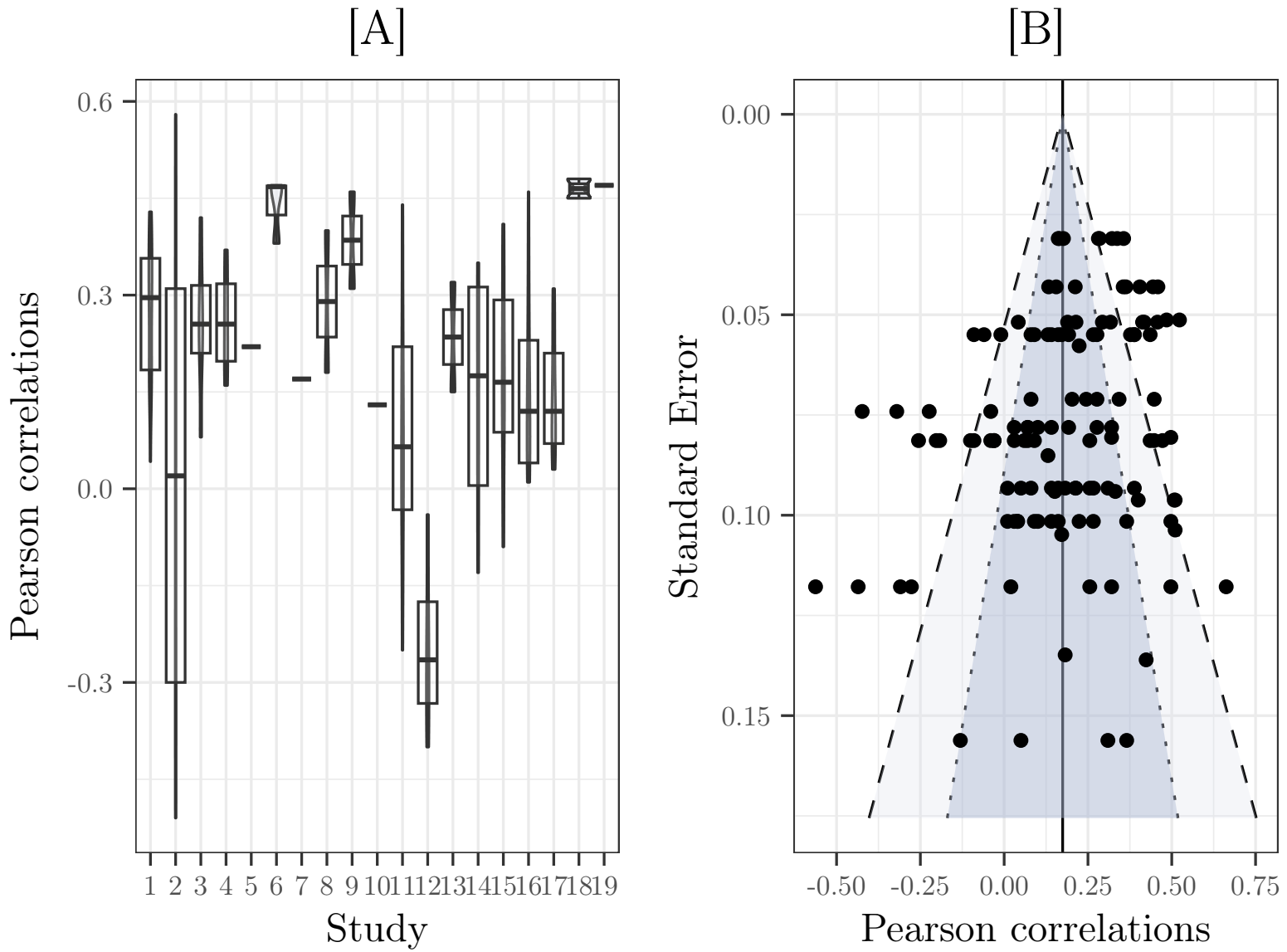
Lionetti, F., Pastore, M., Moscardino, U., Nocentini, A., Pluess, K., Pluess, M. (2019). Sensory Sensory Processing Sensitivity and its Association with Personality Traits and Affect: A Meta-Analysis. *Journal of Research in Personality*, 81, 138-152.


i studies
θ
δ_i
$\hat{\sigma}^2_{ij}$
τ^2
z_{ij}


θ
δ_i
i studies
$\hat{\sigma}^2_{ij}$
τ^2
z_{ij}

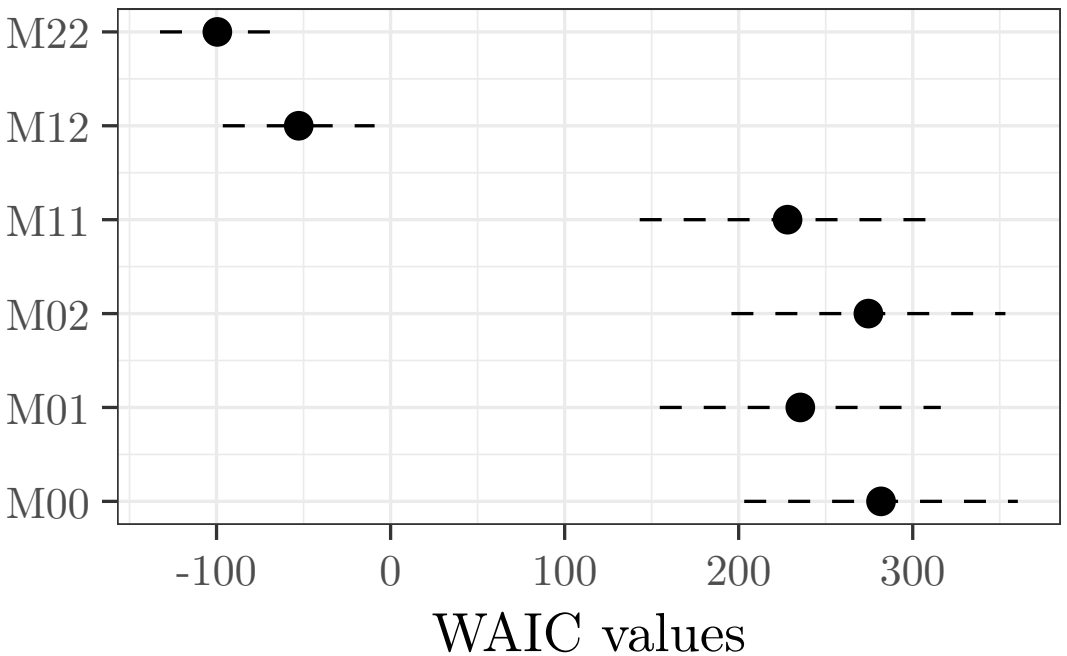

θ
τ
-5.0
-2.5
0.0
2.5
5.0
0
5
10
15

	authors	r	mean	sd	Zmean	Zsd	Nmean	Nsd
1	1 Nocentini	24	0.28	0.11	0.29	0.12	654.33	290.75
2	10 Pluess	1	0.13		0.13		141.00	
3	11 Sobocko	24	0.08	0.19	0.09	0.20	142.00	17.34
4	12 Booth	4	-0.24	0.15	-0.25	0.16	185.00	0.00
5	13 Meredith	2	0.23	0.12	0.24	0.13	116.00	0.00
6	14 Kiellgren	4	0.14	0.22	0.15	0.23	44.00	0.00
7	15 Pluess	16	0.18	0.16	0.18	0.16	334.00	0.00
8	16 Moscardino	12	0.16	0.14	0.16	0.15	100.00	0.00
9	17 Listou	8	0.15	0.10	0.15	0.10	167.00	0.00
10	18 Benham	2	0.46	0.02	0.50	0.03	383.00	0.00
11	19 Aron	1	0.47		0.51		96.00	
12	2 Evers	9	0.01	0.40	0.02	0.44	75.00	0.00
13	3 Liss	6	0.26	0.12	0.27	0.12	201.00	0.00
14	4 Ahadi	4	0.26	0.09	0.27	0.10	118.00	0.00
15	5 Gearhart	1	0.22		0.22		303.00	
16	6 Bakker	3	0.44	0.05	0.47	0.06	111.00	0.00
17	7 Wachs	1	0.17		0.17		94.00	
18	8 Jonsson	2	0.29	0.16	0.30	0.17	57.50	0.71
19	9 Brindle	2	0.39	0.11	0.41	0.12	157.00	0.00
20	TOTAL	126	0.17	0.21	0.18	0.22	259.17	243.51

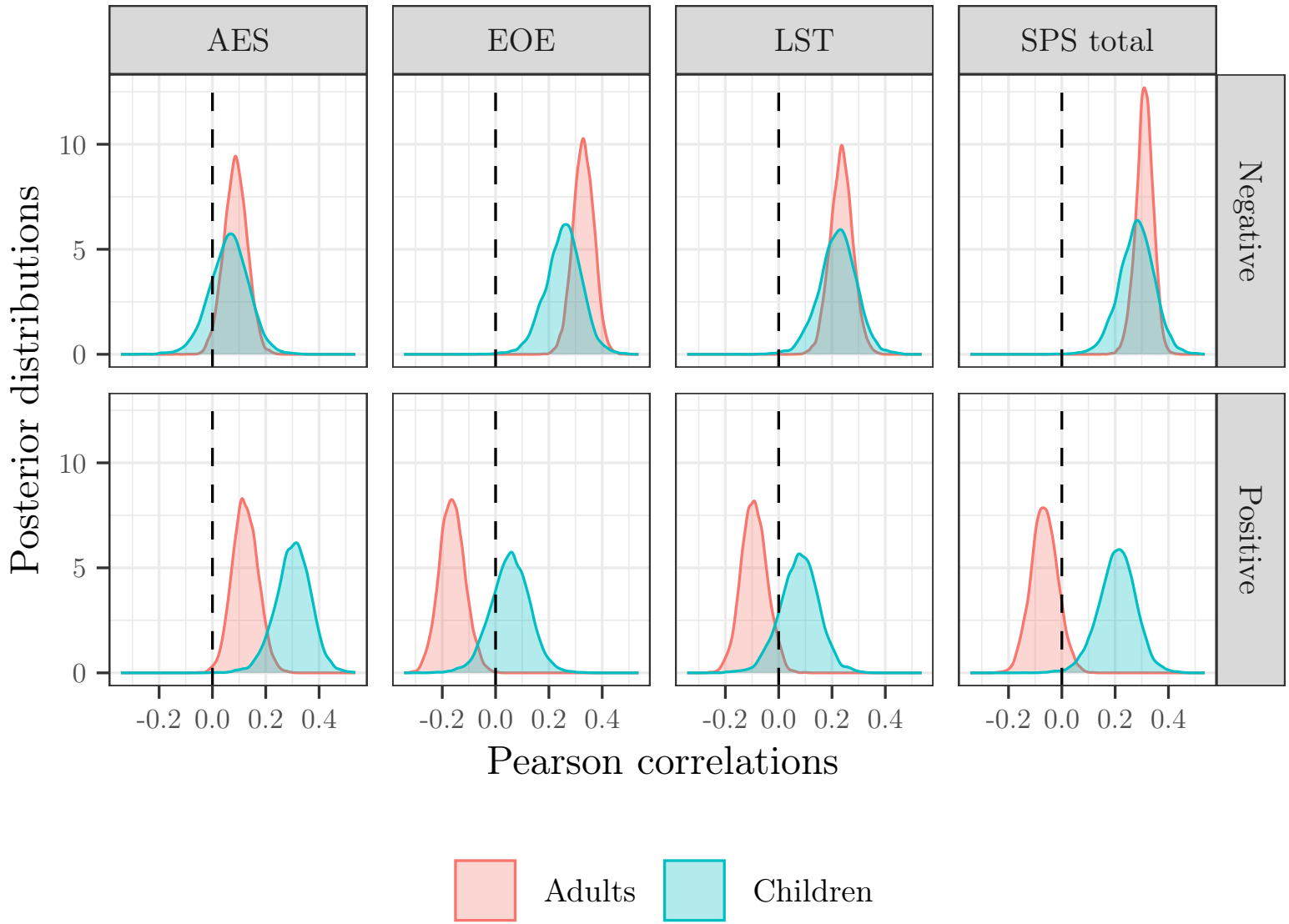

[A]
[B]
Pearson correlations
0.6
0.3
0.0
-0.3
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
Study
Standard Error
0.00
0.05
0.10
0.15
-0.50 -0.25 0.00 0.25 0.50 0.75
Pearson correlations

Model comparison

	n	WAIC	se	delta	weight
M00	126	282	79	-381	0.00
M01	126	235	81	-335	0.00
M02	126	275	79	-374	0.00
M11	126	228	85	-328	0.00
M12	126	-53	44	-47	0.00
M22	126	-100	33	0	0.99



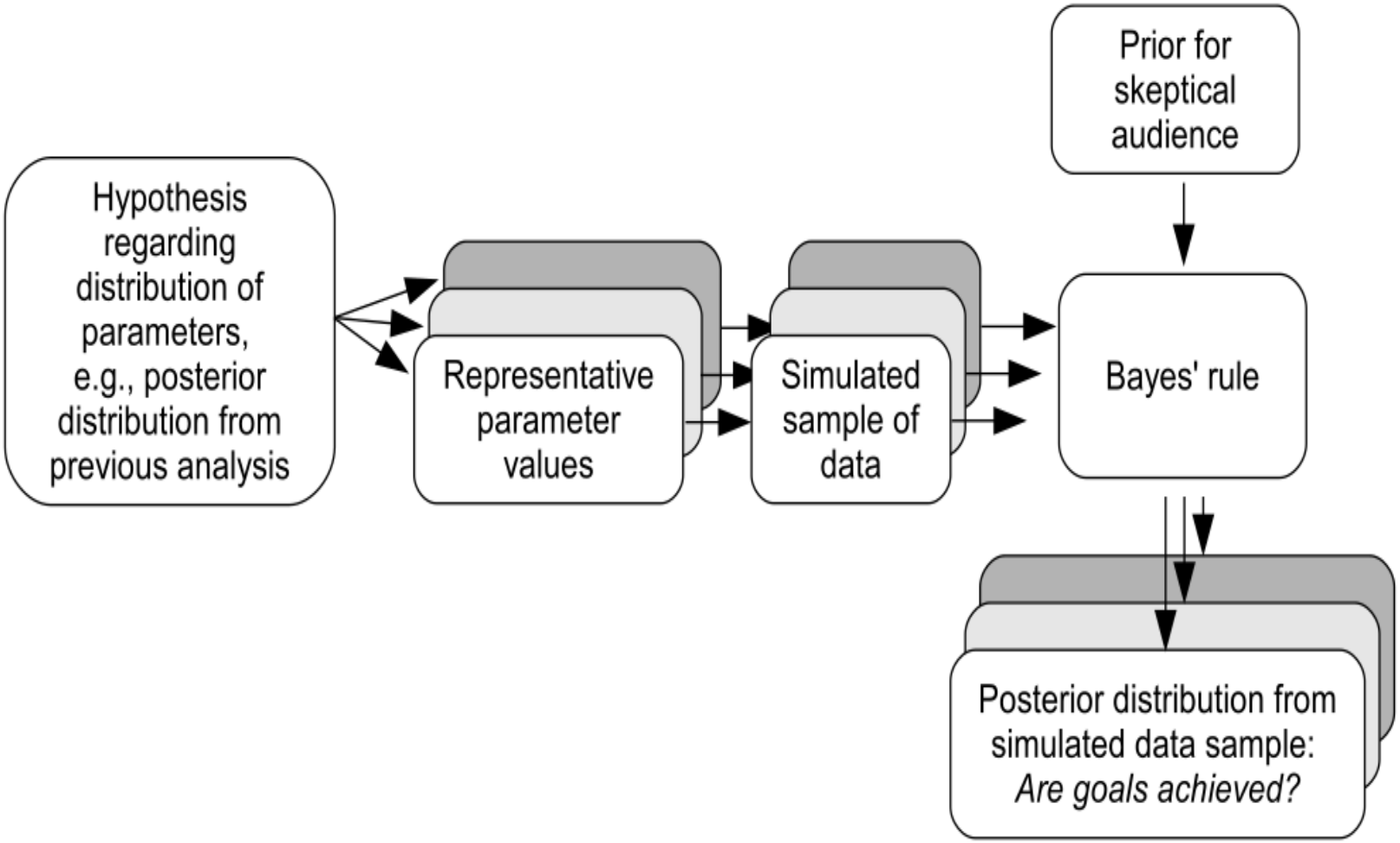
Best model


AES
EOE
LST
SPS total
Negative
Positive
Posterior distributions
10
5
0
-0.2
0.0
0.2
0.4
Pearson correlations
Adults
Children

mtype	spstype	age	mean	hpd.lb	hpd.ub	range	d
neg	AES	children	0.06	-0.05	0.18	0.23	0.13
neg	EOE	children	0.25	0.14	0.36	0.21	0.53
neg	LST	children	0.22	0.11	0.33	0.22	0.45
neg	total	children	0.28	0.17	0.38	0.21	0.58
neg	AES	adults	0.09	0.02	0.15	0.14	0.17
neg	EOE	adults	0.33	0.27	0.39	0.12	0.70
neg	LST	adults	0.24	0.17	0.30	0.13	0.48
neg	total	adults	0.31	0.26	0.37	0.10	0.65
pos	AES	children	0.30	0.20	0.41	0.21	0.63
pos	EOE	children	0.05	-0.06	0.16	0.23	0.11
pos	LST	children	0.08	-0.03	0.20	0.23	0.16
pos	total	children	0.21	0.10	0.32	0.22	0.42
pos	AES	adults	0.12	0.05	0.20	0.15	0.25
pos	EOE	adults	-0.16	-0.24	-0.09	0.15	-0.33
pos	LST	adults	-0.09	-0.17	-0.01	0.15	-0.18
pos	total	adults	-0.07	-0.15	0.01	0.16	-0.13

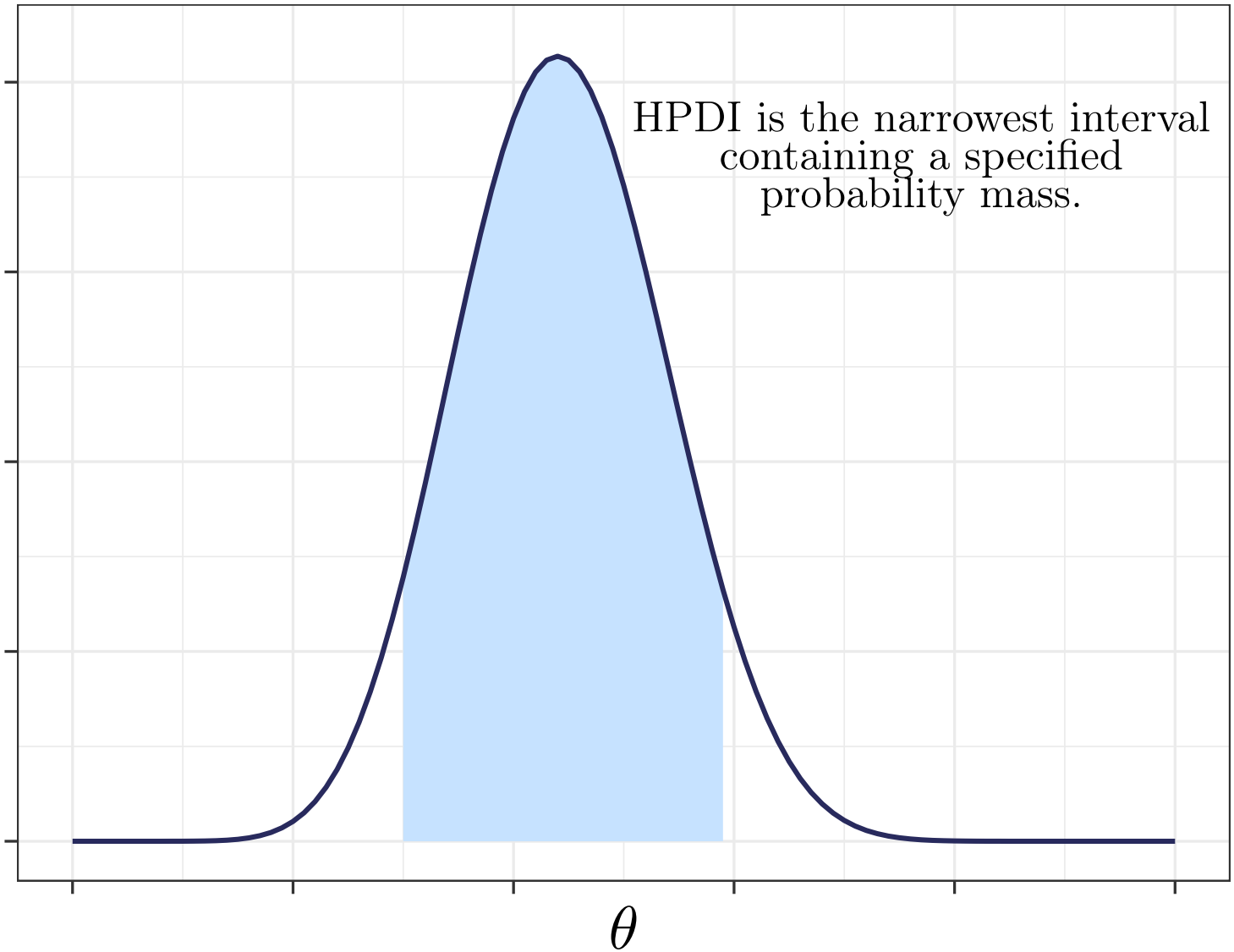
Power analysis

Esempio 4: (Bayesian) power analysis

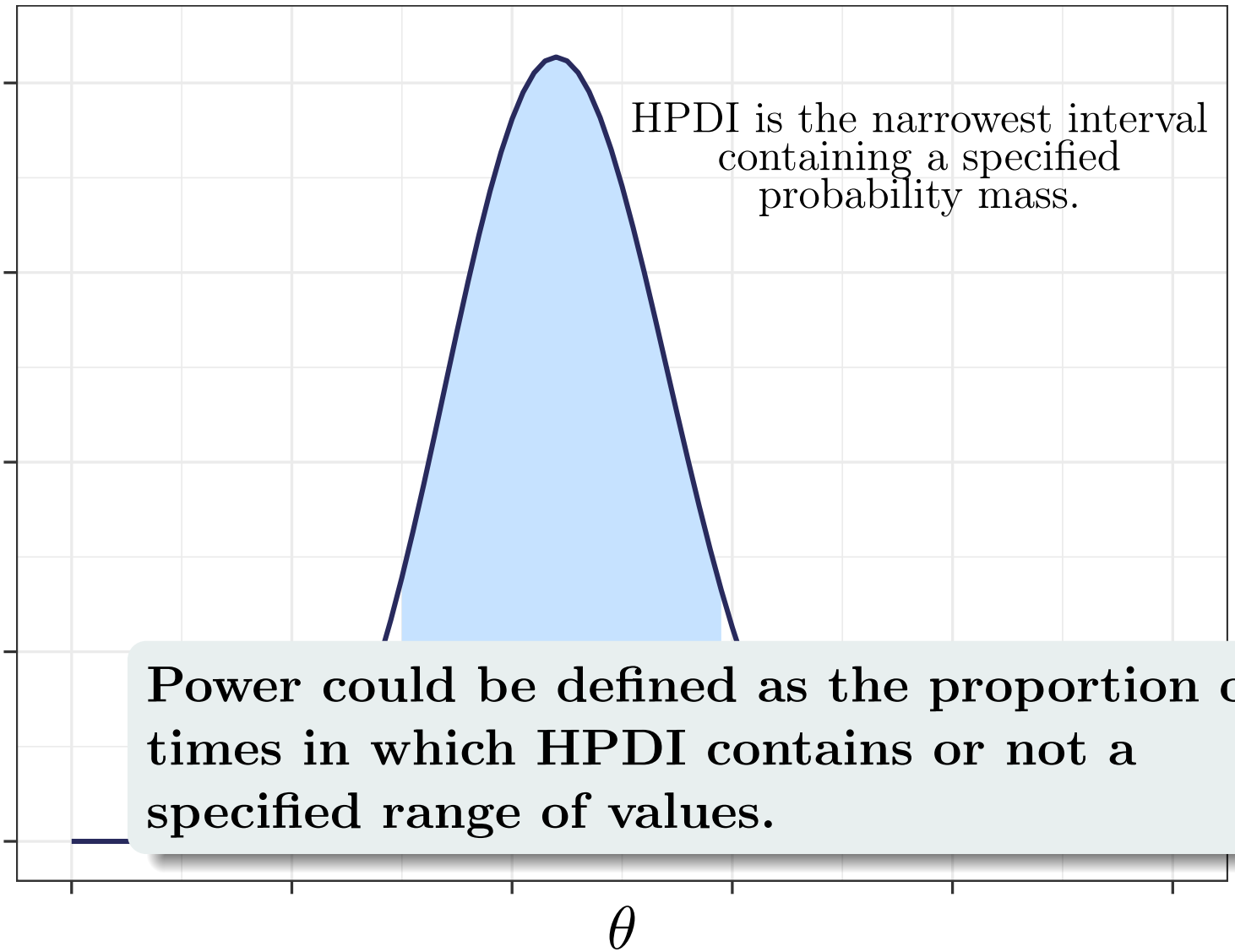


Kruschke, J. K. (2015). Doing Bayesian Data Analysis. Academic Press, UK.

Highest Posterior Density Interval



Highest Posterior Density Interval



Power could be defined as the proportion of times in which HPDI contains or not a specified range of values.

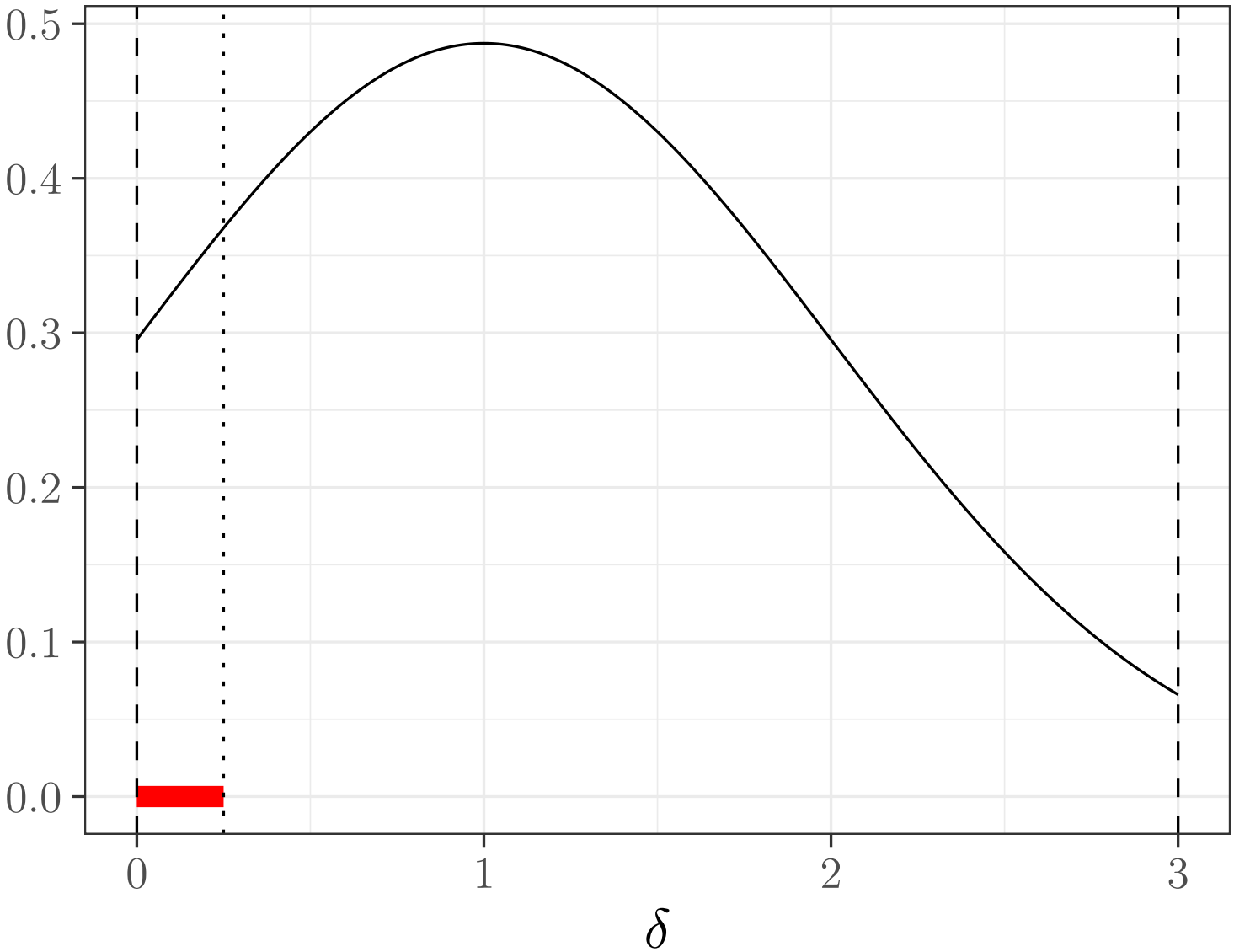
Example 4.1

- Let's suppose to want to test the mean scores standardized difference (δ) between two groups (Control and Experimental).
- A difference less than 0.25 is considered trivial in our particular application, i.e. null difference; consequently, the interval $[0, 0.25]$ represents the *Region of Practical Equivalence* (ROPE; Kruschke, 2015).
- How many subjects (per group) do we need for detecting a difference greater than 0.25?

Example 4.1: prior definition

- We know that, empirically, the difference could range between 0 to 3.
- We also believe that the most plausible value of the difference should be around 1.
- Consequently, we assume for the difference a truncated normal distribution in the interval $[0, 3]$, with mean equal to 1 and a variance value related to our degree of beliefs (for example $\sigma^2 = 1$).

Example 4.1

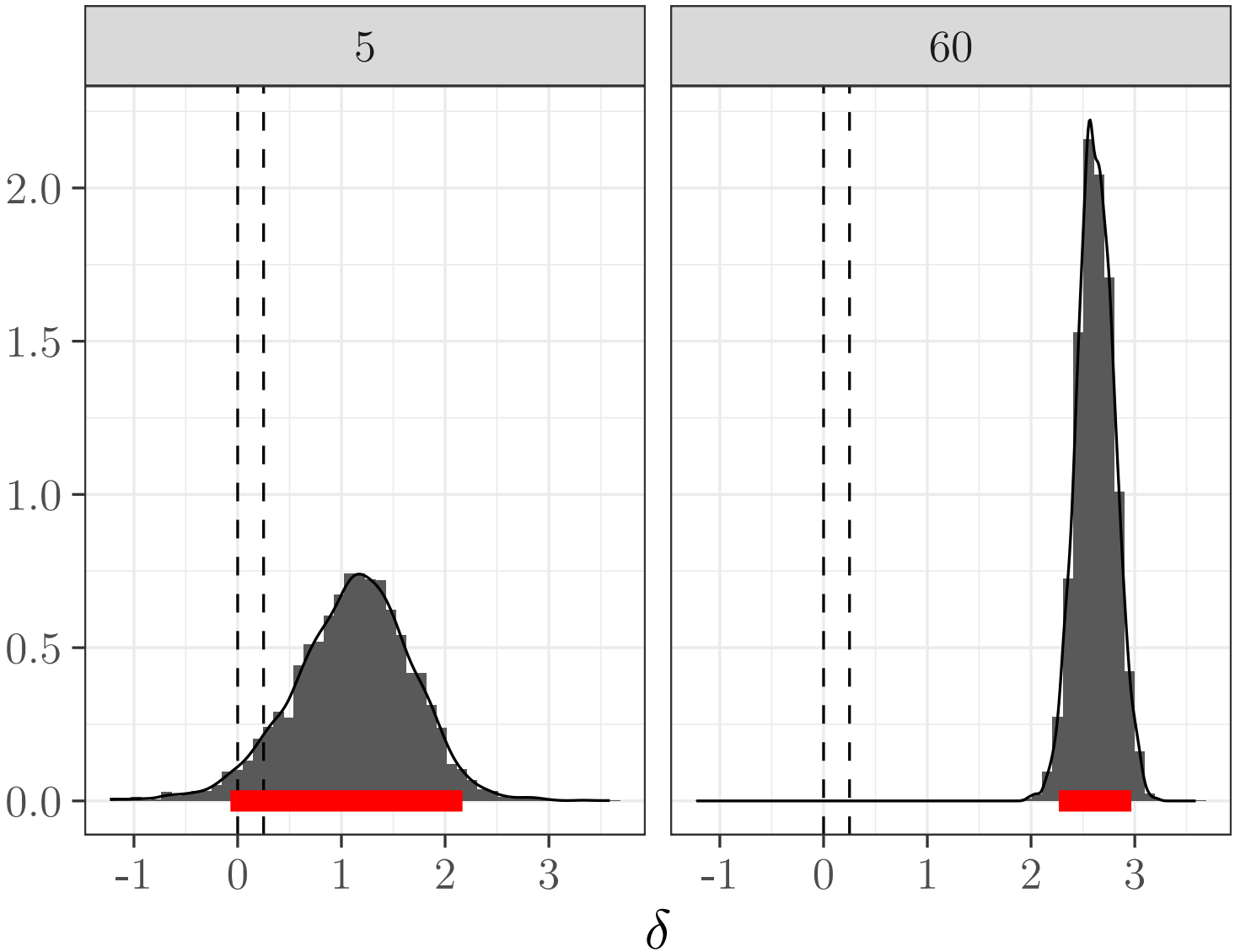


Example 4.1

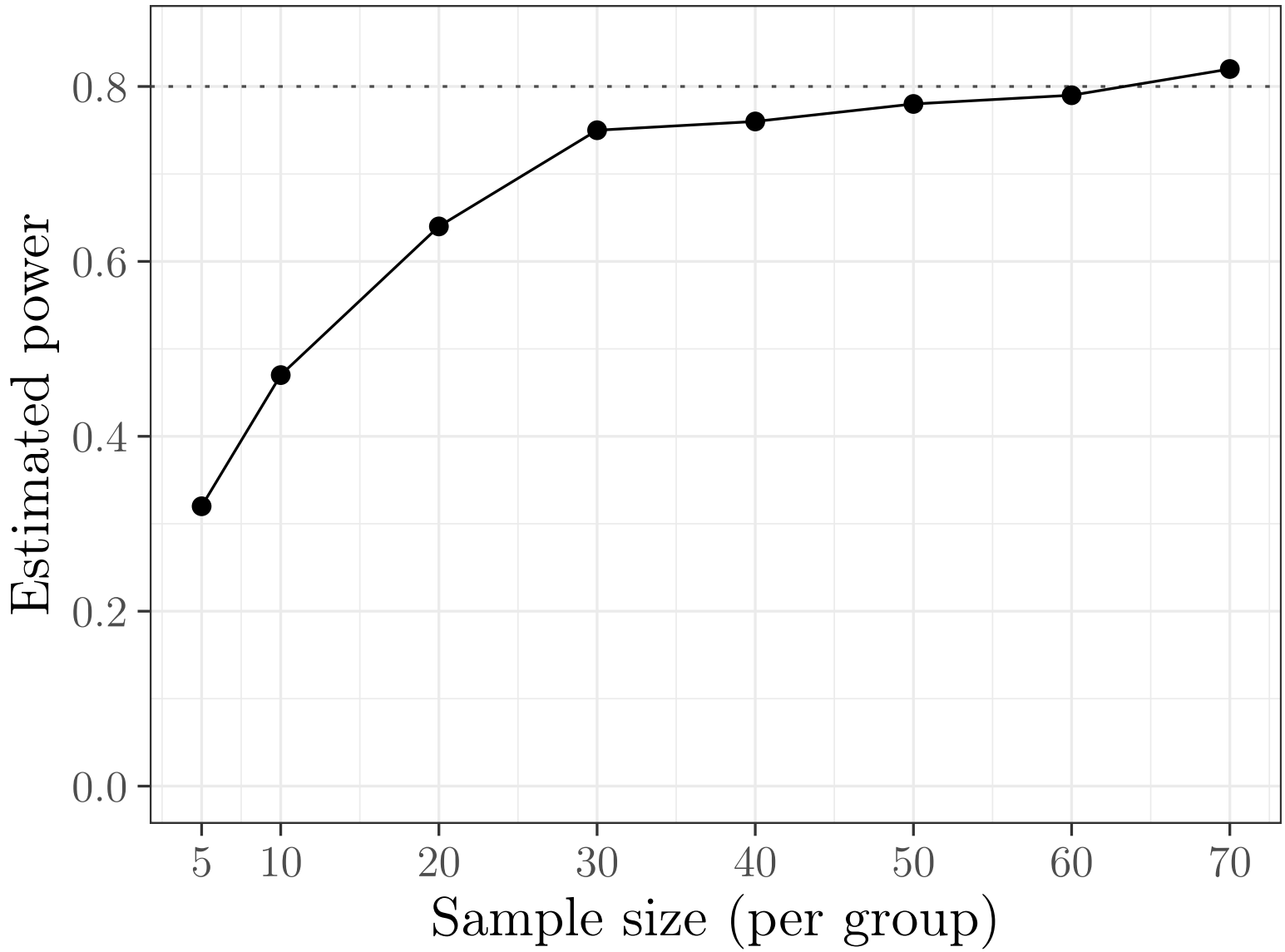
The following steps were repeated 1000 times for each specified value of sample size $n = (5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100)$.

1. Sample a δ value from a TruncNorm(0, 3, $\mu = 1$, $\sigma = 1$).
2. Generate a couple of independent samples with sizes n from two populations in which the true difference $\mu_1 - \mu_2$ is δ.
3. Estimate posterior distribution of $\hat{\delta} = \overline{x}_2 - \overline{x}_1$ by using a skeptical prior – e.g. $\delta \sim \text{Normal}(0, 1)$.
4. Compute the 89% HPDI, and the number of cases in which the value 0.25 is outside the interval.

Example 4.1: results



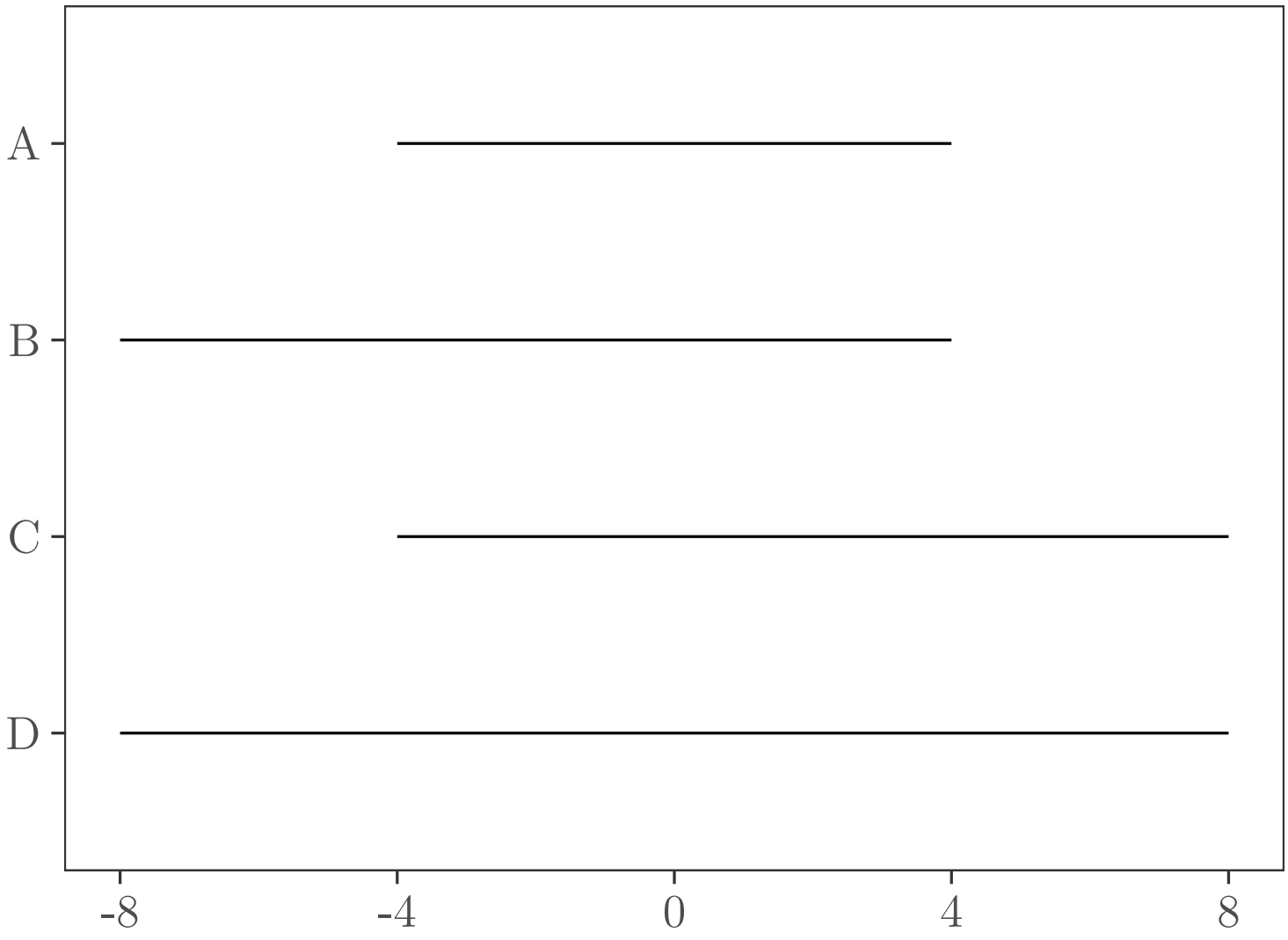
Example 4.1: results



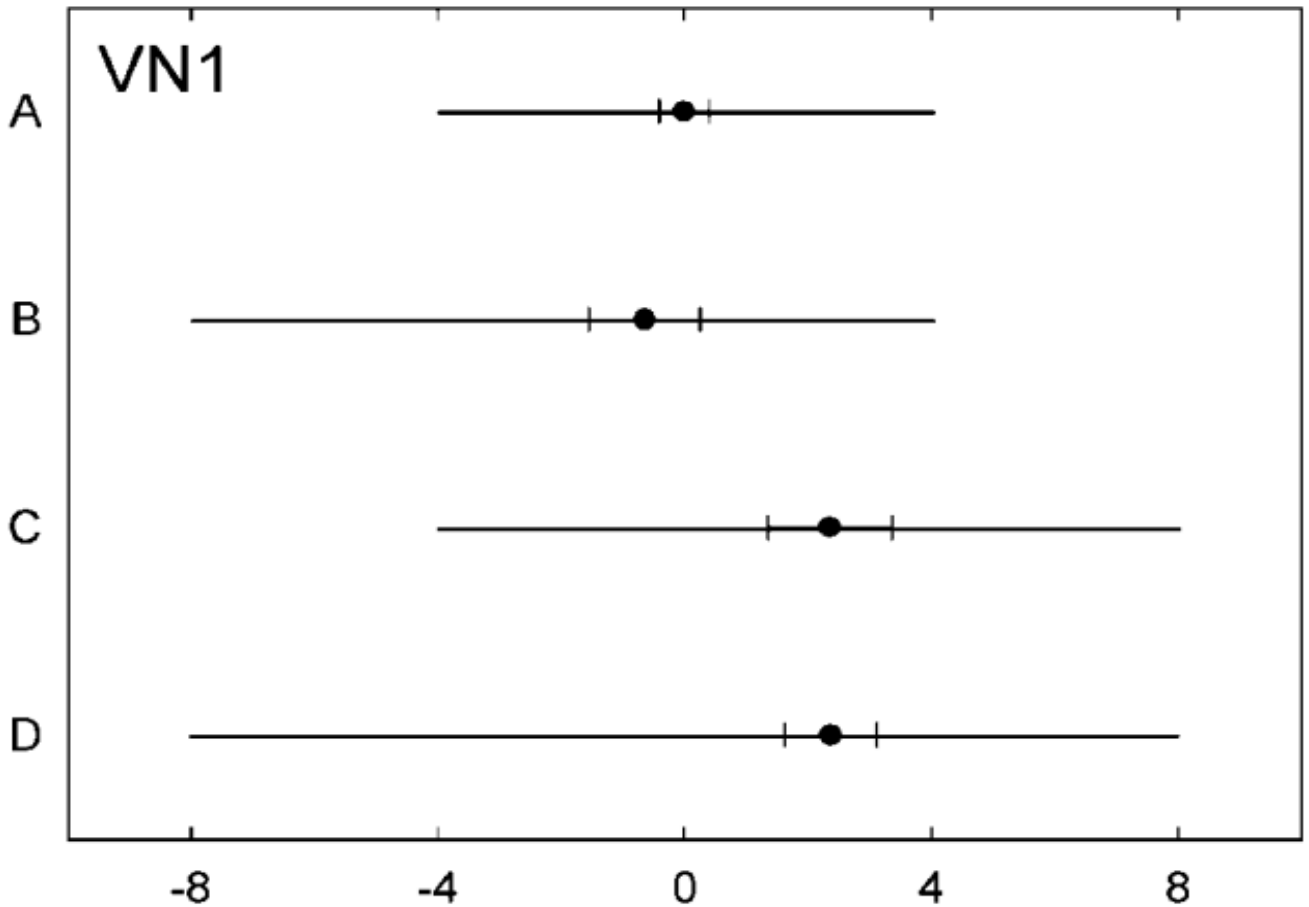
Example 4.2

- In a recent work, McIntosh et al. (2005, 2017) have shown how line bisection in neglect can be analysed by looking at the effect that the positions of the left and the right endpoints exert on the position marked by the patient, rather than by looking at the directional error.
- We consider a 2×2 factorial design in which the Left endpoint is either 40 or 80 mm to the left of the display centre, and the Right endpoint is either 40 or 80 mm to the right of the display centre, and the dependent measure is the response position (P), coded with respect to the page midline.
- Overall, 40 trials are given (defined as Block) – 10 repeats for each of the four stimulus lines, in pseudo-random order.

Example 4.2: The task



Example of Individual data for a patient



McIntosh, R.D., Schindler, I., Birchall, D., Milner, A.D. (2005). Weights and measures: A new look at bisection behaviour in neglect. *Cognitive Brain Research*, 25, 833–850.

Example 4.2: The model

$$\begin{cases} P \sim \mathcal{N}(\mu, \sigma) & \text{general model} \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{cases}$$

Example 4.2: The model

$$\begin{cases} P \sim \mathcal{N}(\mu, \sigma) & \text{general model} \\ \mu = k + dP_L \times L + dP_R \times R & \mu \text{ model} \end{cases}$$

Example 4.2: The model

$$
\begin{cases}
P \sim \mathcal{N}(\mu, \sigma) & \text{general model} \\
\\
\mu = k + dP_L \times L + dP_R \times R & \mu \text{ model} \\
\sigma = exp(\alpha + \beta \times (R - L)) & \sigma \text{ model}
\end{cases}
$$

Example 4.2: The model

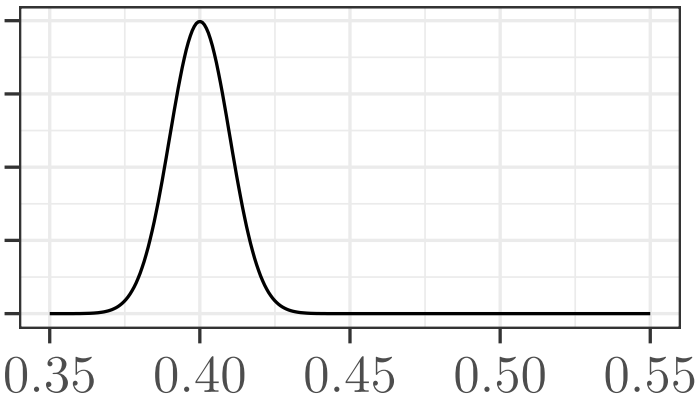
$$
\begin{cases}
P \sim \mathcal{N}(\mu, \sigma) & \text{general model} \\
\\
\mu = k + dP_L \times L + dP_R \times R & \mu \text{ model} \\
\sigma = exp(\alpha + \beta \times (R - L)) & \sigma \text{ model} \\
\\
k = k_0 + k_B[Block] & \text{elements of } \mu \\
dP_L = dP_{L_0} + dP_{L_B}[Block] & \\
dP_R = dP_{R_0} + dP_{R_B}[Block] & \\
\\
\alpha = \alpha_0 + \alpha_B[Block] & \text{elements of } \sigma \\
\beta = \beta_0 + \beta_B[Block] &
\end{cases}
\tag{1}
$$

Example 4.2: The question

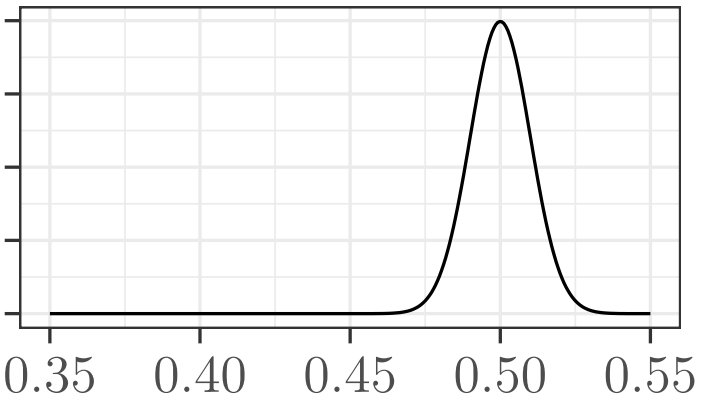
- We are interested in evaluating the experimental effects (for semplicity only for the mean), measured from the parameters dP_L and dP_R.
- In this case, we suppose that values lower than 0.2 can be considered equivalent to a NULL value.
- Unfortunately, we cannot have more than one patient and consequently our power analysis does not focus on "*how many subjects*" but instead on "*how many Blocks*" do we need.

Representative parameter values

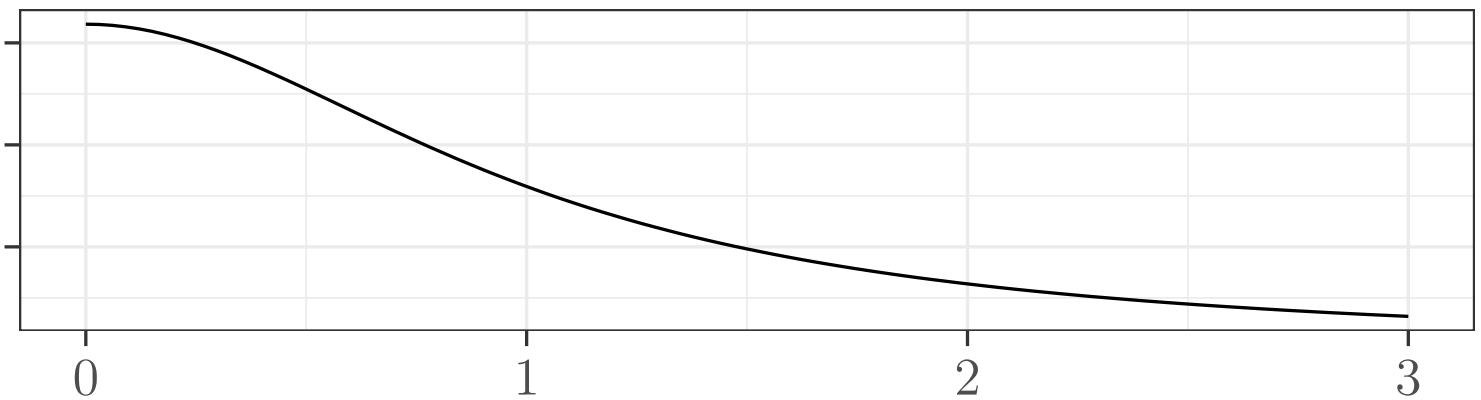
$dP_{L_B} = \mathcal{N}(0.4, 0.01)$



$dP_{R_B} = \mathcal{N}(0.5, 0.01)$



$\sigma_{dP_L} = \sigma_{dP_R} = \text{HalfCauchy}(0, 1)$

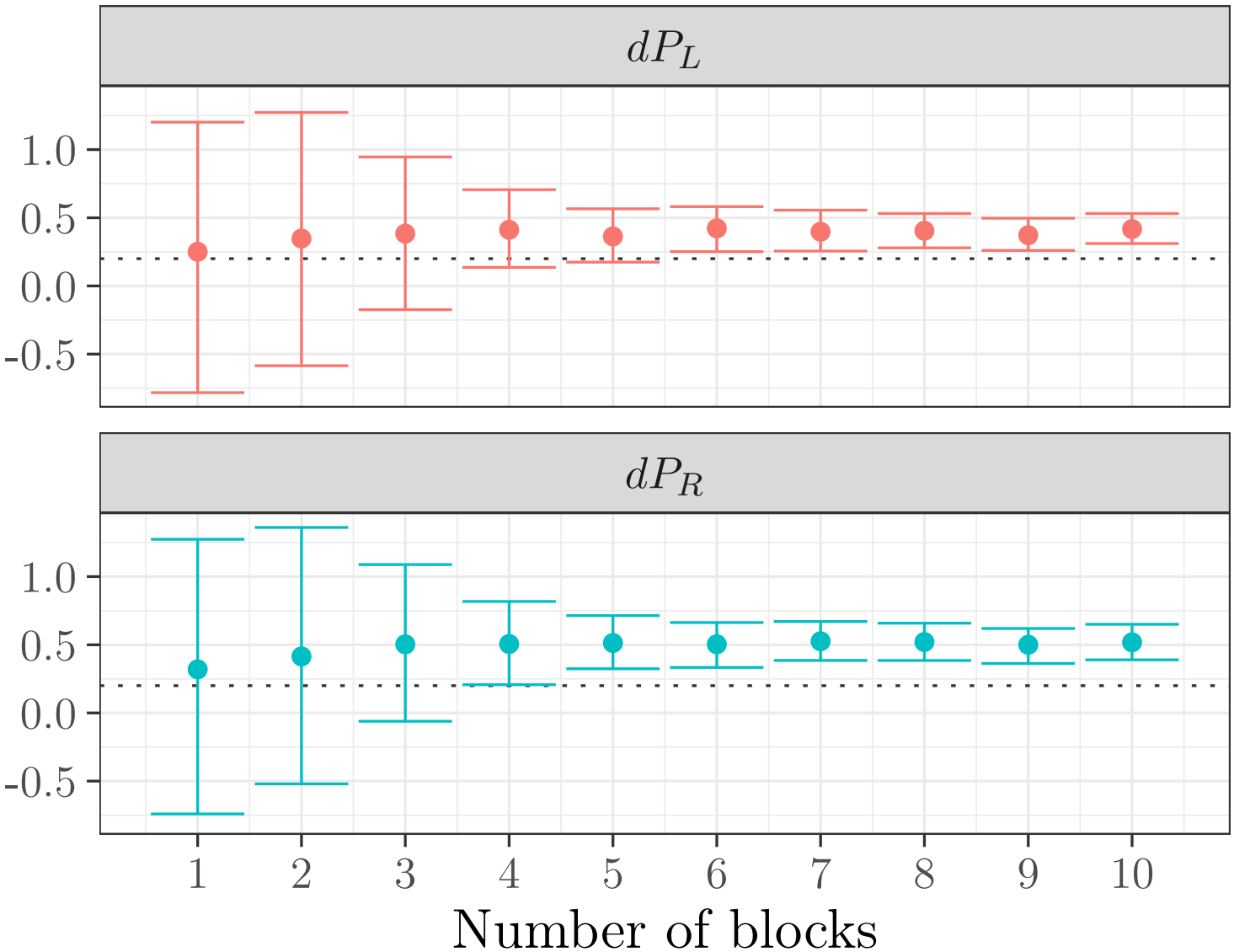


Example 4.2

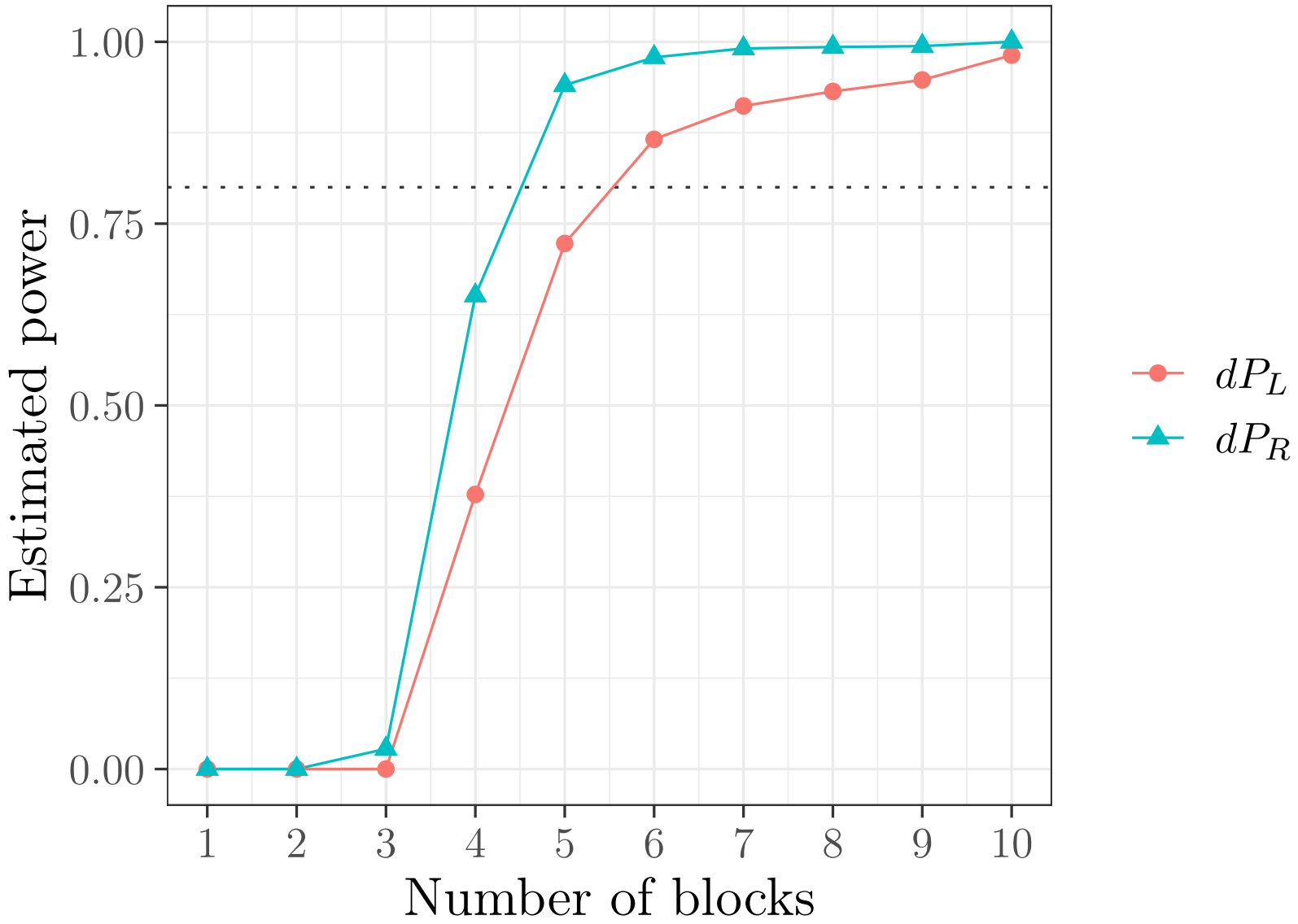
We replicated at least 1000 times for each specified number of Blocks (from Inf to -Inf) the following steps:

1. Generate 40 values of P for each Block.
2. Estimate posterior distributions of dP_L and dP_R.
3. Compute the 89% HPDI, and the number of cases in which the value 0.2 is outside the interval.

Posterior distributions (89% HPDI)



Estimated power

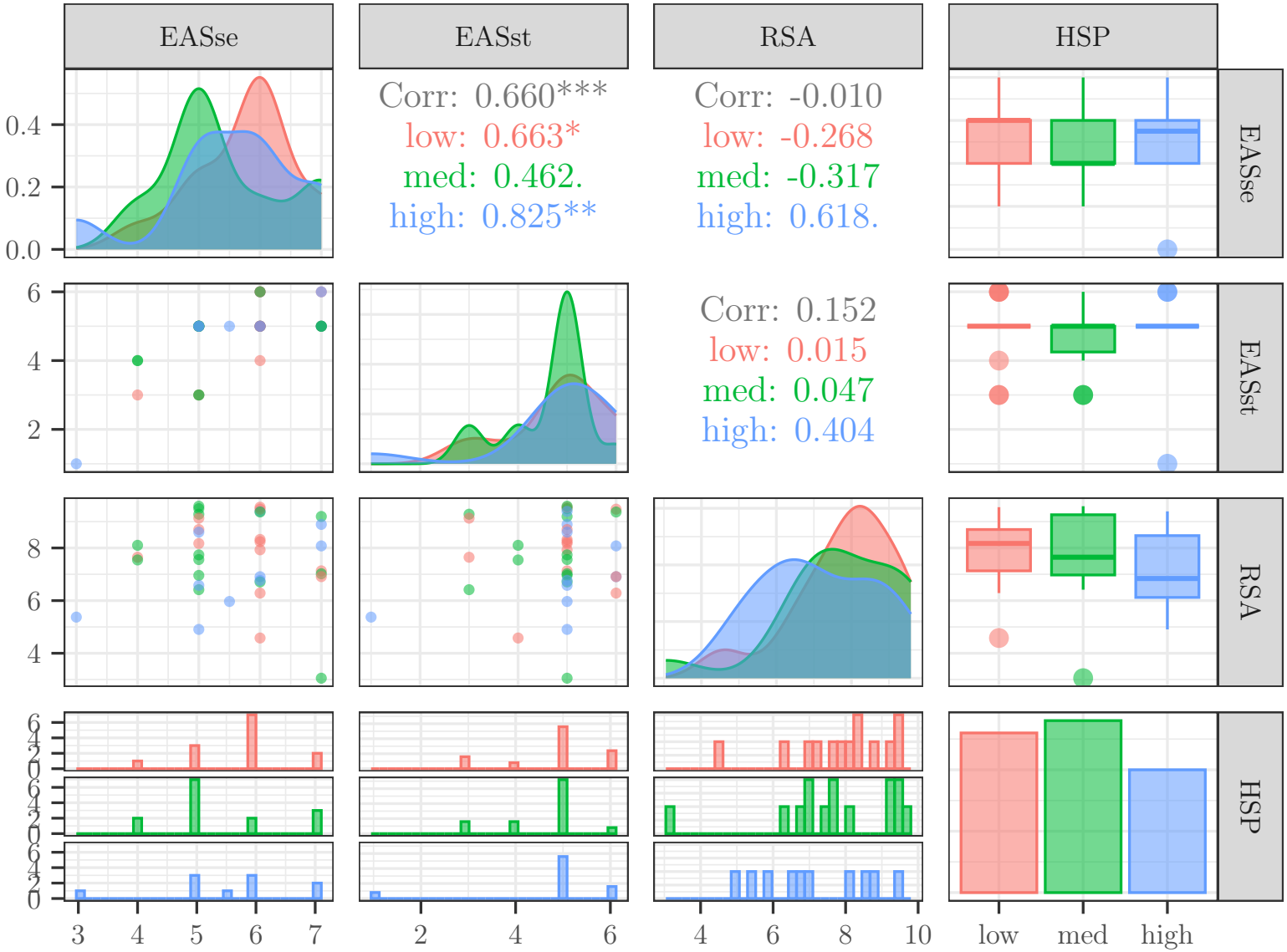


The dark side of the priors

Esempio 5: A Correlation Case

- In a study of the association between environmental sensitivity, parenting behavior and physiological arousal, 37 mothers with 3-month-old infants were recruited.
- The aim of the study was to examine whether the association between environmental sensitivity and parenting behavior is moderated by physiological arousal.
- Three variables were considered:
  1. Environmental Sensitivity (ES, measured by the Highly Sensitive Person scale, Pluess et al., 2023)
  2. Parenting Behavior (measured by the Emotional Availability Scales, EAS, Biringen, 2008)
  3. Physiological Arousal (respiratory sinus arrhythmia, RSA)

Pluess, M., Lionetti, F., Aron, E. N., & Aron, A. (2023). People differ in their sensitivity to the environment: An integrated theory, measurement and empirical evidence. *Journal of Research in Personality*, 104, 104377.
Biringen, Z. (2008). *The Emotional Availability (EA) Scales Manual*, 4th Edn. Boulder, CO: International Center for Excellence in Emotional Availability.


EASse
EASst
RSA
HSP
Corr: 0.660***
low: 0.663*
med: 0.462.
high: 0.825**
Corr: -0.010
low: -0.268
med: -0.317
high: 0.618.
Corr: 0.152
low: 0.015
med: 0.047
high: 0.404
low
med
high

Hypothetical correlation values

Pairs		SPS	r	μ
RSA	EASse	Low	[-0.2; 0.2]	0.00
		Medium	[0; 0.3]	0.15
		High	[0.2; 0.4]	0.30
RSA	EASst	Low	[-0.2; 0.2]	0.00
		Medium	[0; 0.3]	0.15
		High	[0.2; 0.4]	0.30

- Once we have defined the target intervals, we need to formalise the priors.

Hypothetical correlation values

Pairs		SPS	r	μ
RSA	EASse	Low	[-0.2; 0.2]	0.00
		Medium	[0; 0.3]	0.15
		High	[0.2; 0.4]	0.30
RSA	EASst	Low	[-0.2; 0.2]	0.00
		Medium	[0; 0.3]	0.15
		High	[0.2; 0.4]	0.30

- Since the correlations are between -1 and 1, the truncated Student's t is a good choice.

Hypothetical correlation values

Pairs		SPS	r	μ
RSA	EASse	Low	[-0.2; 0.2]	0.00
		Medium	[0; 0.3]	0.15
		High	[0.2; 0.4]	0.30
RSA	EASst	Low	[-0.2; 0.2]	0.00
		Medium	[0; 0.3]	0.15
		High	[0.2; 0.4]	0.30

- We set the degrees of freedom to 3 and the mean of the distribution to the mid-points of the target intervals (i.e. 0, 0.15, 0.3).

Hypothetical correlation values

Pairs		SPS	r	μ
RSA	EASse	Low	[-0.2; 0.2]	0.00
		Medium	[0; 0.3]	0.15
		High	[0.2; 0.4]	0.30
RSA	EASst	Low	[-0.2; 0.2]	0.00
		Medium	[0; 0.3]	0.15
		High	[0.2; 0.4]	0.30

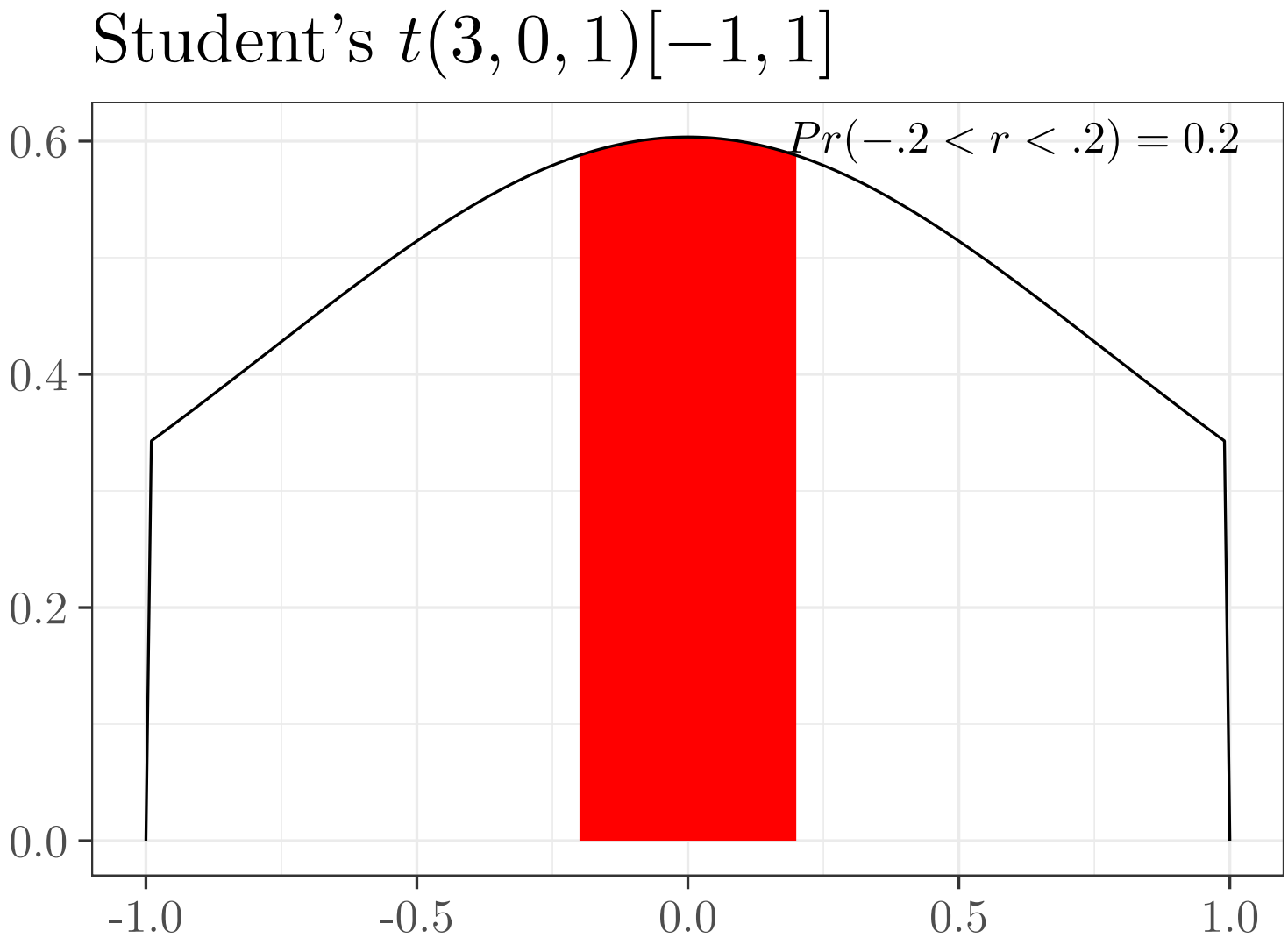
- We then need to determine the variance based on the knowledge/belief of the researcher.

Hypothetical correlation values

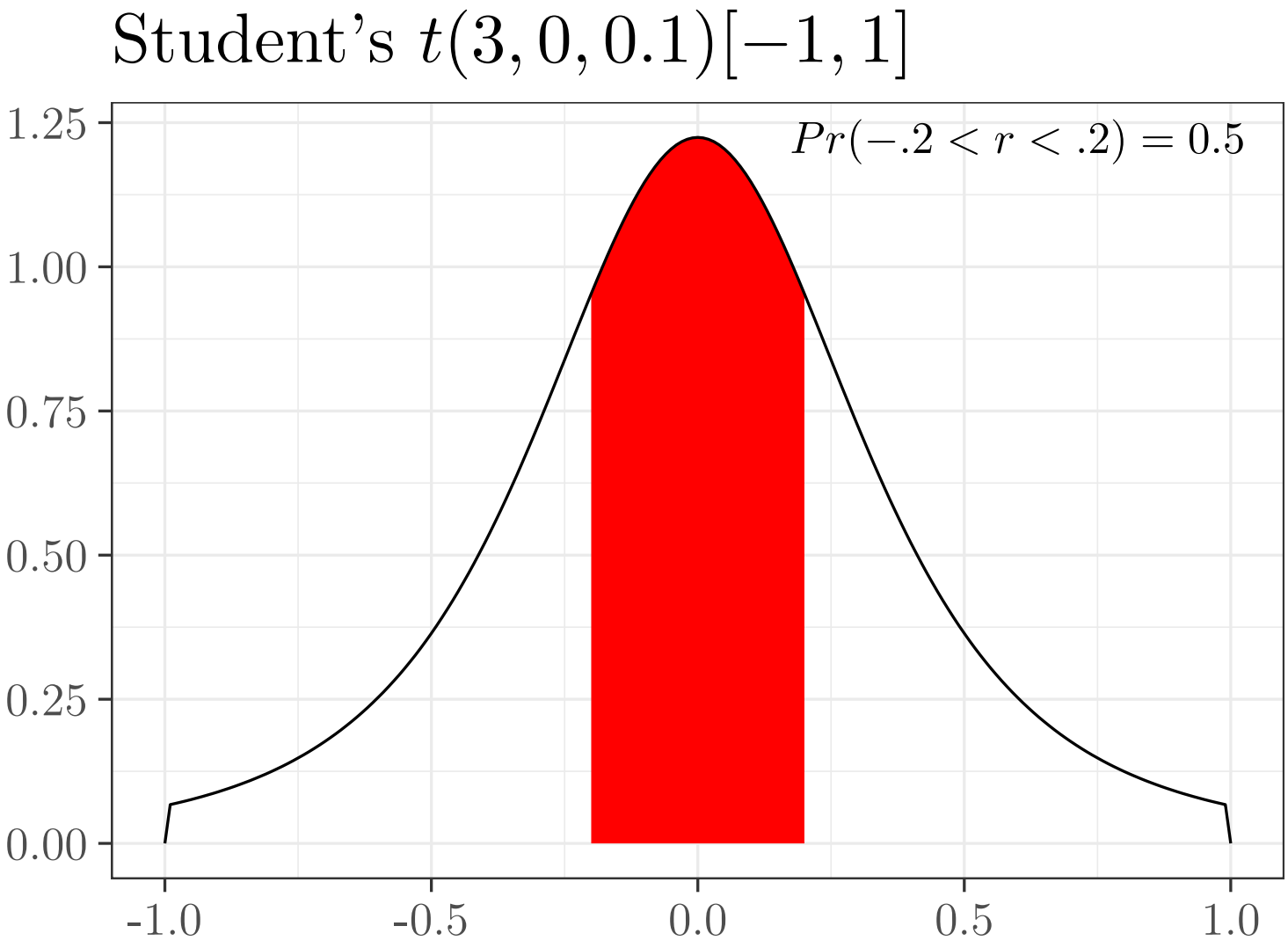
Pairs		SPS	r	μ
RSA	EASse	Low	[-0.2; 0.2]	0.00
		Medium	[0; 0.3]	0.15
		High	[0.2; 0.4]	0.30
RSA	EASst	Low	[-0.2; 0.2]	0.00
		Medium	[0; 0.3]	0.15
		High	[0.2; 0.4]	0.30

- A simple way of doing this is to ask the expert for the subjective probability that the correlation is in the target interval, and use this probability to calculate the appropriate variance of the distribution.

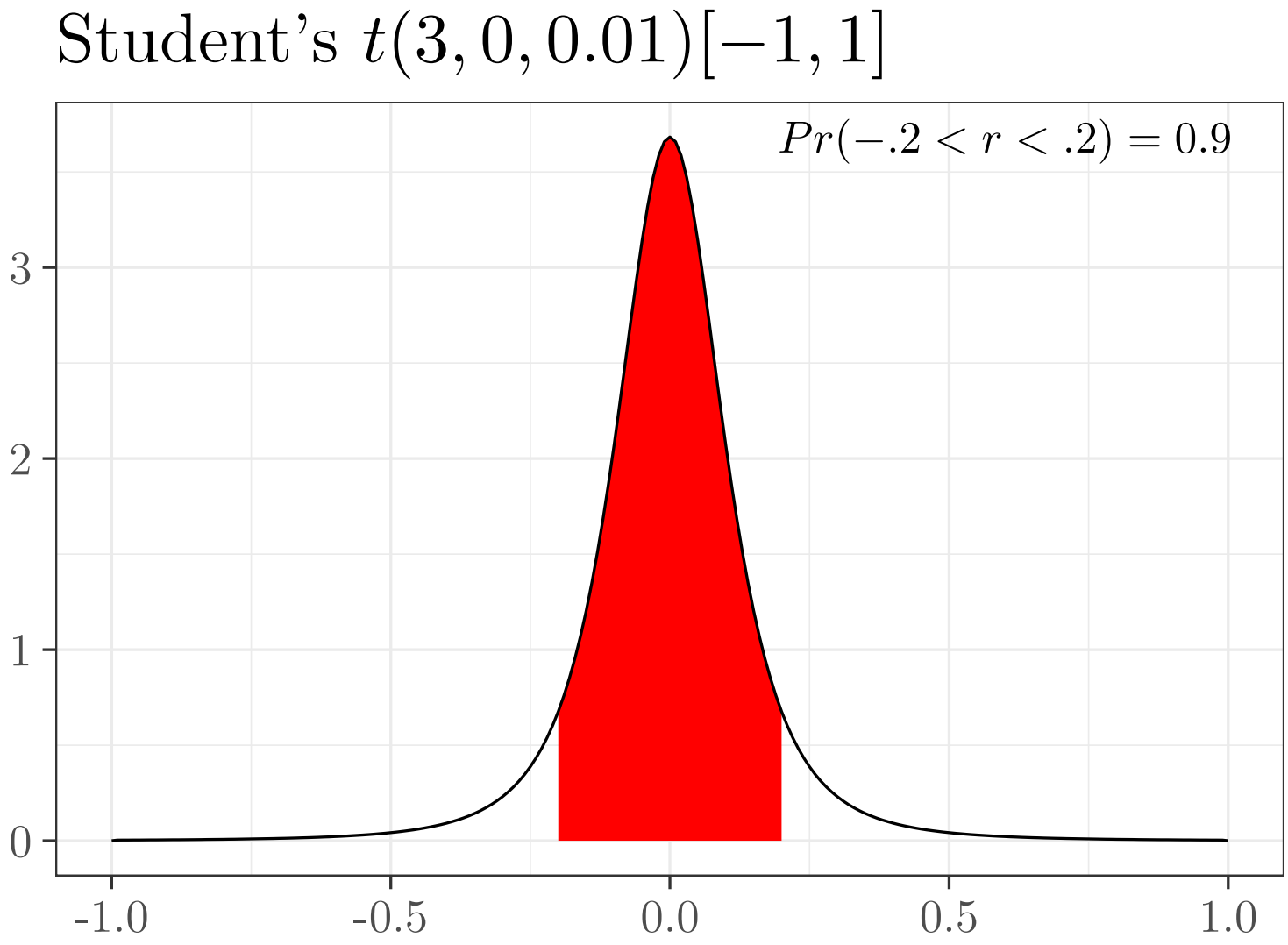
Truncated Student's t



Truncated Student's t

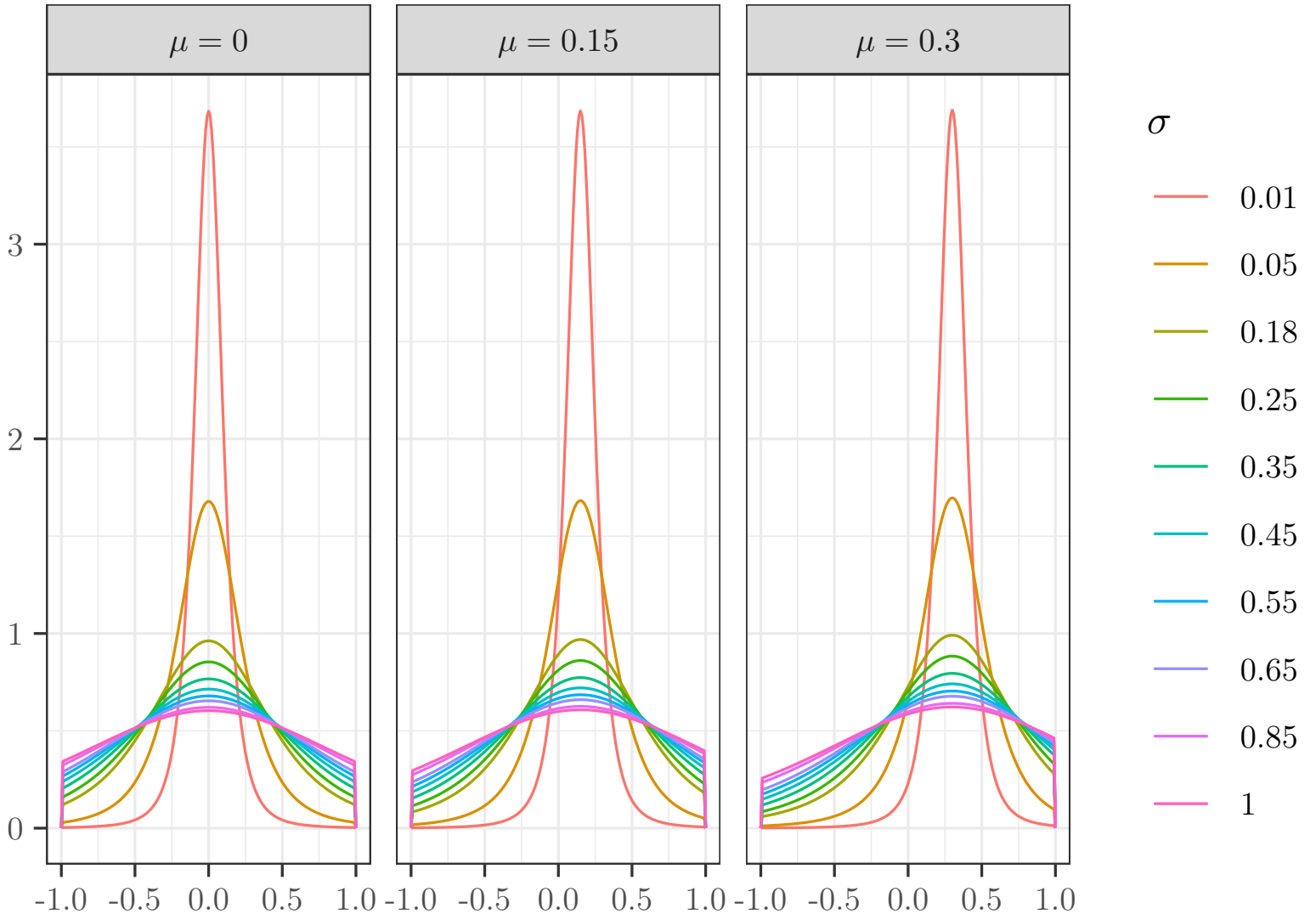


Truncated Student's t



Sensitivity Analysis

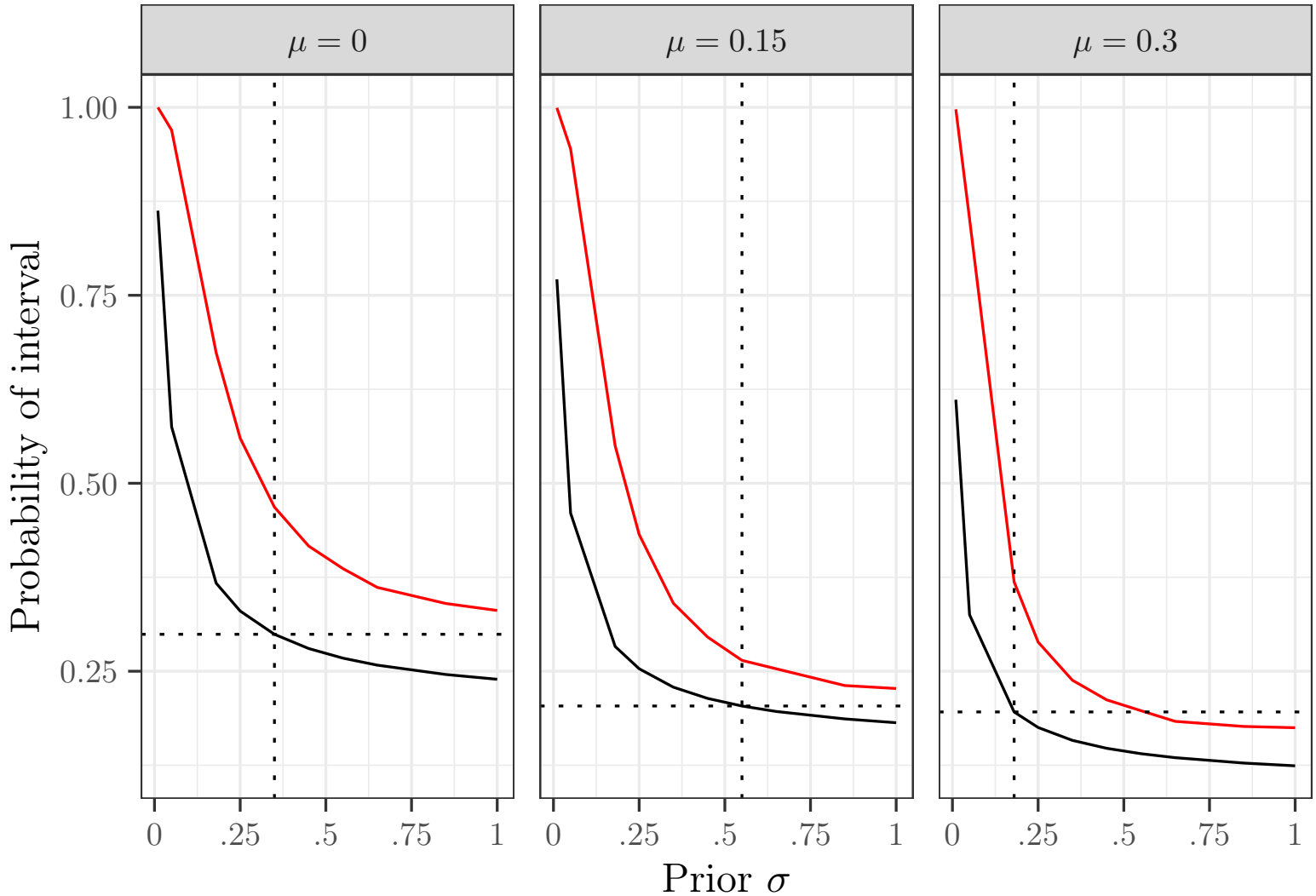
- Given the small sample size, the choice of hypothesized intervals and related prior probabilities is crucial. In particular, we must consider that prior variance cannot be too small because – with small sample size – this means to obtain results depending only on prior distribution.
- Consequently we performed a sensitivity analysis for the three groups identified with the SPS scores. The objective of this analysis was to identify an optimal prior variance value that would enable us to evaluate our hypotheses without excessively influencing the results.



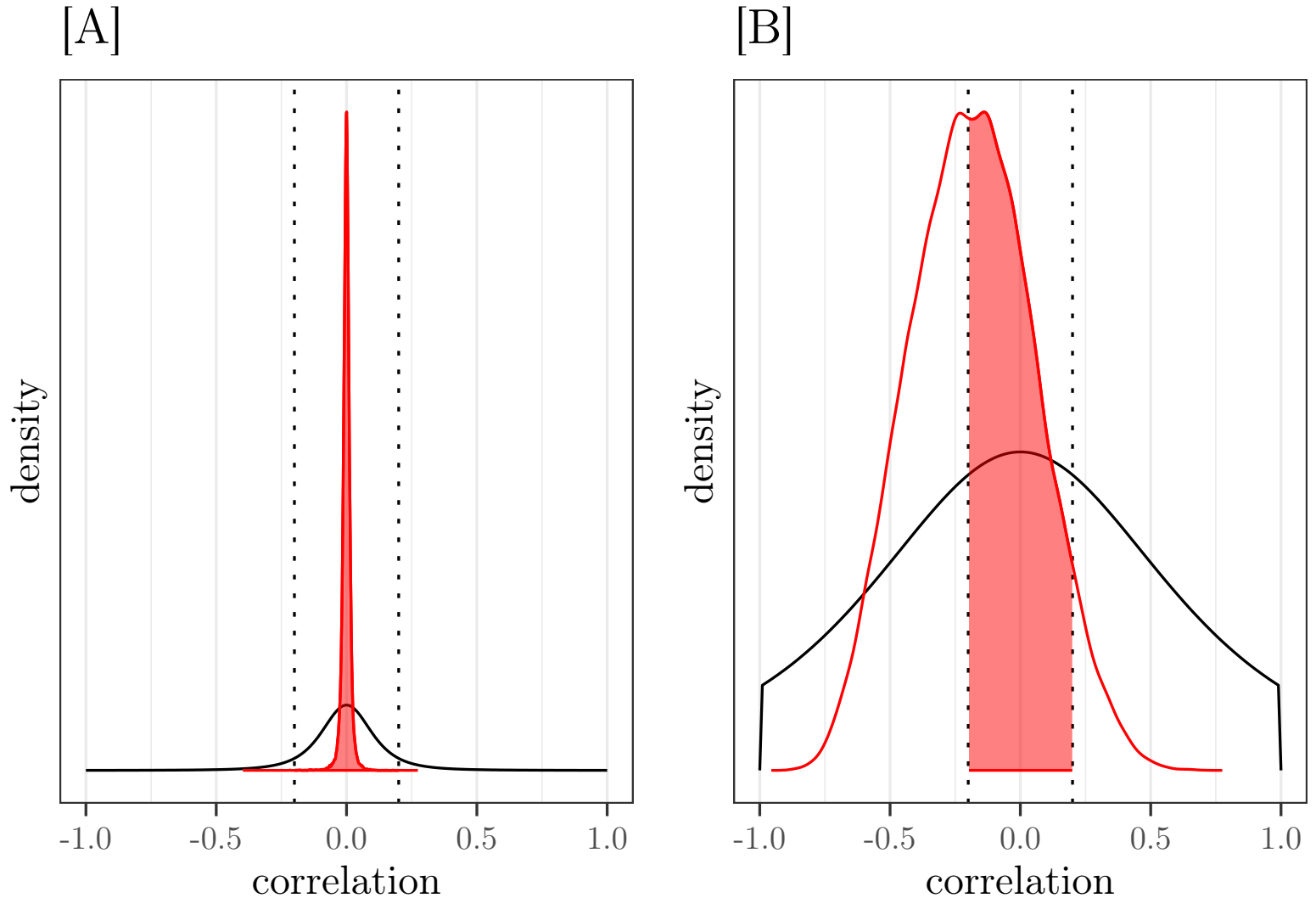
Truncated Student's t priors distributions in different scenarios. Panels refers to different means, colored lines to different standard deviations for fixed 3 degrees of freedom.

Estimation

- For each scenario, we repeated the correlation estimation using these priors – the mean remains fixed across the groups, specifically at 0, 0.15, 0.3, while σ varies based on the 10 selected values – with four chains of 15000 effective replications each.
- Additionally, we calculated the prior and posterior probabilities for the target intervals, namely [-0.2; 0.2], [0; 0.3], and [0.2; 0.4] respectively.



On the x-axis, the chosen values of σ are plotted. The black lines represent the prior probability of the target intervals, while the red lines represent the posterior probability. The vertical dashed lines indicate the selected values of σ for the analyses, while the horizontal dashed lines represent their corresponding prior probabilities.

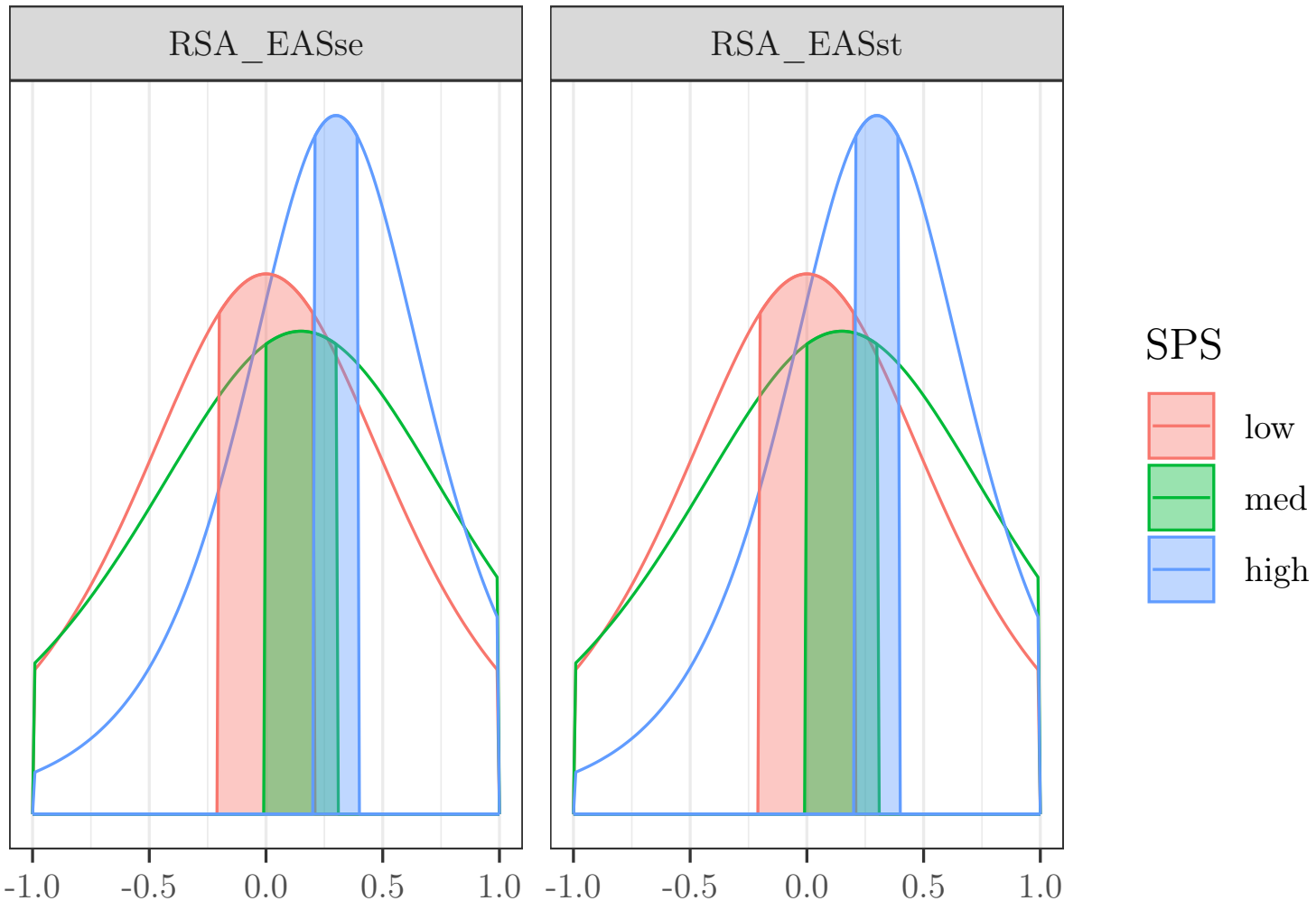


Examples of Scenarios for the case with $\mu = 0$; the black curves represent the priors, and the red curves represent the posteriors. [A] Priors is a truncated Student's $t(3, 0, 0.01)[-1, 1]$. $Pr(\rho \in [-.2, .2]) = 0.86$; $Pr(\rho \in [-.2, .2]|D) = 1$. [B] Priors is a truncated Student's $t(3, 0, 0.35)[-1, 1]$. $Pr(\rho \in [-.2, .2]) = 0.3$; $Pr(\rho \in [-.2, .2]|D) = 0.47$.

Correlation analysis

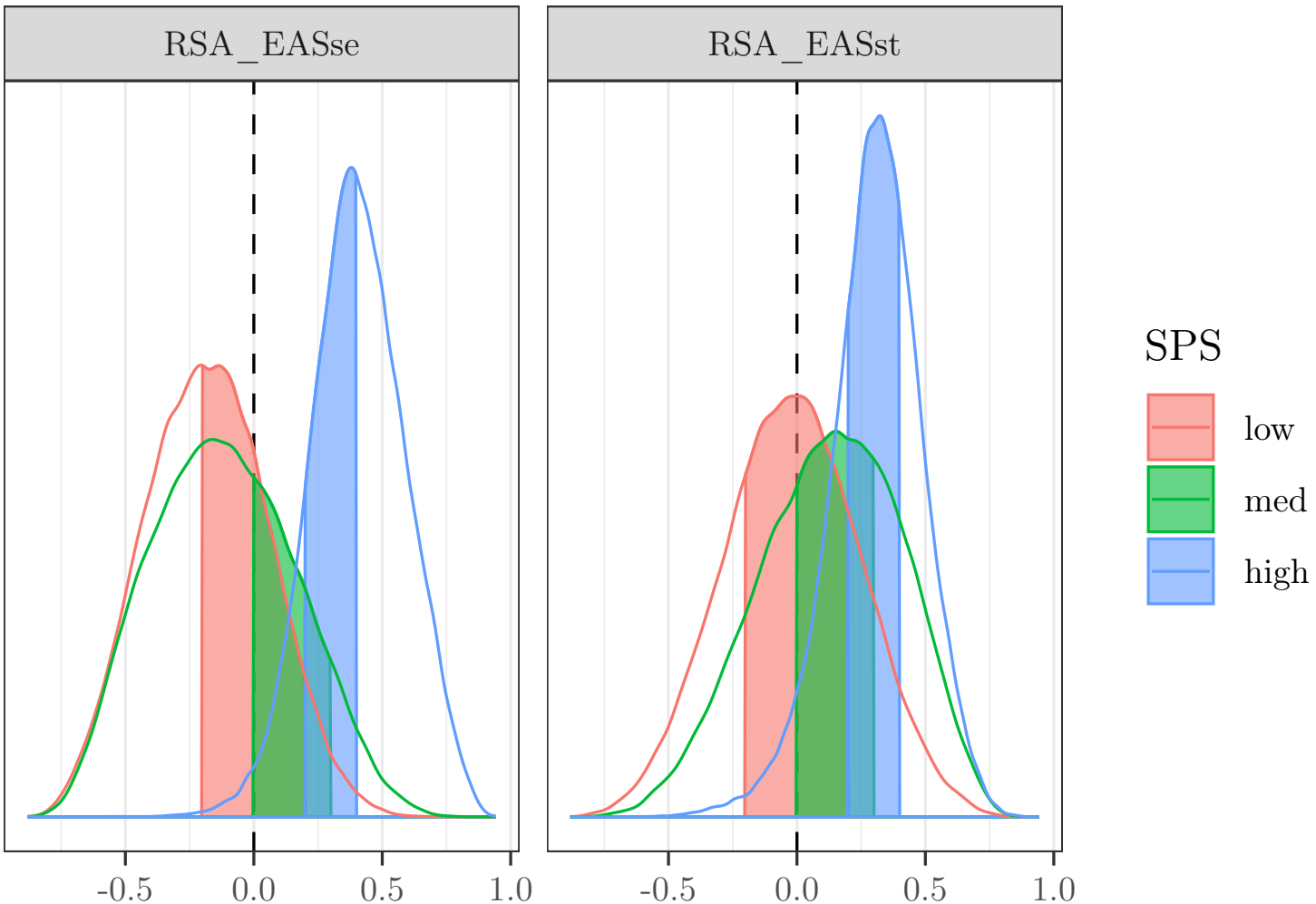
Pairs		SPS	r	μ	σ	Prob.
RSA	EASse	Low	[-0.2; 0.2]	0.0	0.3	0.30
		Medium	[0; 0.3]	0.2	0.6	0.20
		High	[0.2; 0.4]	0.3	0.2	0.20
RSA	EASst	Low	[-0.2; 0.2]	0.0	0.3	0.30
		Medium	[0; 0.3]	0.2	0.6	0.20
		High	[0.2; 0.4]	0.3	0.2	0.20

Priors



Priors distributions of correlations. Colors indicate the three SPS groups. Filled areas represent the prior probability of the target interval.

Posteriors



Posterior distributions of correlations. Colors refer to the SPS groups. Filled areas represent the posterior probability that correlation lies in the target interval.

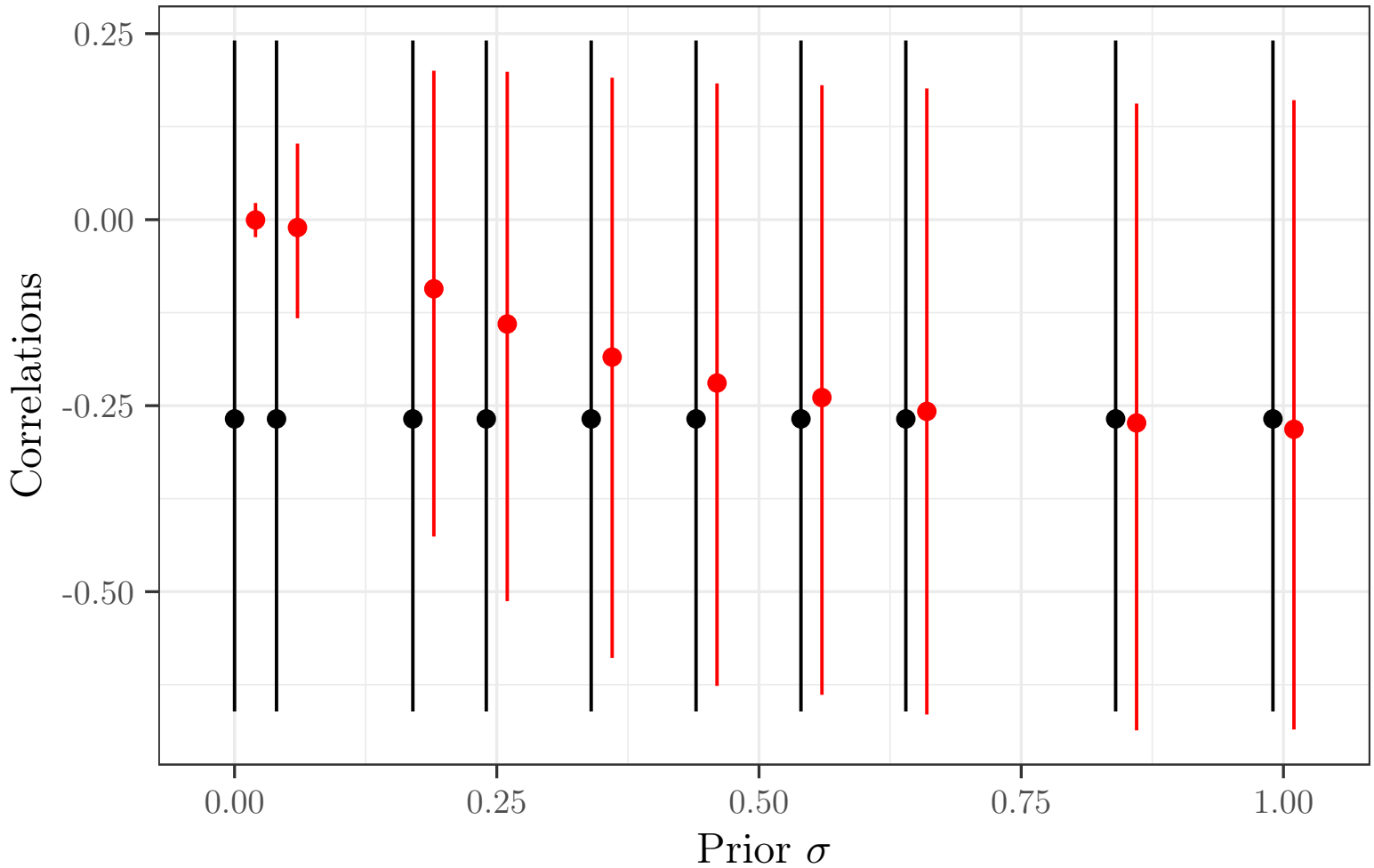
Pairs	HSP	n	Est r	89% HDI	Interval	Prior Pr.	Post Pr.
RSA EASse	low	13	-0.18	[-0.58; 0.2]	[-0.2; 0.2]	0.30	0.47
	med	14	-0.13	[-0.57; 0.33]	[0; 0.3]	0.20	0.27
	high	10	0.40	[0.12; 0.72]	[0.2; 0.4]	0.20	0.37
RSA EASst	low	13	-0.02	[-0.46; 0.42]	[-0.2; 0.2]	0.30	0.54
	med	14	0.14	[-0.32; 0.58]	[0; 0.3]	0.20	0.39
	high	10	0.31	[0.02; 0.61]	[0.2; 0.4]	0.20	0.45

Posterior table. Pairs = paired variables; HSP = SPS group; n = sample size; Est r = estimated correlation, i.e. the median of posterior distribution; 89% HDI = Highest Density Interval; Interval = target interval; Prior Pr. = prior probability of target interval; Post Pr. = posterior probability of target interval.

A frequentist remark

- To evaluate the advantages of the Bayesian approach over the simple frequentist method, we compare the frequentist confidence intervals with the Bayesian credible intervals.

A frequentist remark



Comparison of 89% confidence intervals (black lines) and 89% credibility intervals (red lines) as a function of the chosen values of σ in the prior.

- Obviously, the frequentist estimates (Pearson r) do not change – as the prior is not used – and the intervals are always quite wide, averaging 0.9, which is almost half of the available range $[-1, 1]$.
- On the other hand, the HDIs increase as σ increases, and when this parameter is too low – indicating a very strong prior – the HDIs become too narrow to be credible.
- In other words, if we were to adopt a frequentist approach (without using a prior), we would necessarily rely solely on the observed correlations and therefore have very wide confidence intervals. As a result, the credibility of the results would be certainly lower.
- Conversely, by adopting priors, we can reason in terms of probabilities (prior and posterior) and assess hypotheses regarding the differences between groups. This approach provides a level of credibility for the results that is certainly not less than that of the frequentist approach.

Summarizing

- The Bayesian approach allows incorporating prior information (priors) and updating it with the newly observed data to obtain the posterior distribution.
- With small sample sizes, the choice of the prior distribution is crucial and can significantly influence the results. Sensitivity analysis on the priors allows evaluating the impact of the prior choice on the results.
- In cases of small samples, the Bayesian approach can be more informative compared to frequentist confidence intervals alone. However, the use of overly informative priors can lead to results excessively influenced by the prior assumptions rather than the observed data.

Used R packages

- `bayesplot`. Gabry J, Mahr T (2025).
- `brms`. Bürkner P (2017).
- `cmdstanr`. Gabry J, Češnovar R, Johnson A, Bronder S (2025).
- `cowplot`. Wilke C (2024).
- `gamlss.dist`. Stasinopoulos M, Rigby R (2023).
- `GGally`. Schloerke B, Cook D, Larmarange J, Briatte F, Marbach M, Thoen E, Elberg A, Crowley J (2024).
- `ggdist`. Kay M (2024).
- `ggplot2`. Wickham H (2016).
- `HDInterval`. Meredith M, Kruschke J (2022).
- `knitr`. Xie Y (2025).
- `mnormt`. Azzalini A, Genz A (2022).
- `posterior`. Bürkner P, Gabry J, Kay M, Vehtari A (2025).
- `R`. R Core Team (2025).
- `Rcpp`. Eddelbuettel D, Francois R, Allaire J, Ushey K, Kou Q, Russell N, Ucar I, Bates D, Chambers J (2025).
- `report`. Makowski D, Lüdecke D, Patil I, Thériault R, Ben-Shachar M, Wiernik B (2023).
- `rethinking`. McElreath R (2024).
- `rstan`. Stan Development Team (2025).
- `rstanarm`. Goodrich B, Gabry J, Ali I, Brilleman S (2024).
- `StanHeaders`. Stan Development Team (2020).
- `truncnorm`. Mersmann O, Trautmann H, Steuer D, Bornkamp B (2023).
- `xtable`. Dahl D, Scott D, Roosen C, Magnusson A, Swinton J (2019).

massimiliano.pastore@unipd.it
https://psicostat.dpss.psy.unipd.it/

LaTeX