

TESTING ANOVA REPLICATIONS BY MEANS OF THE PRIOR PREDICTIVE P -VALUE

MARIËLLE ZONDERVAN-ZWIJNENBURG

PROMOTOR: HERBERT HOIJTINK
COPROMOTOR: RENS VAN DE SCHOOT



Universiteit Utrecht

Example

ANOVA

$$y_{ijd} = \mu_{jd} + \epsilon_{ijd}$$

$$\epsilon_{ijd} \sim \mathcal{N}(0, \sigma^2_d)$$

Study	Group 1		Group 2		Group 3	
	n	M (SD)	n	M (SD)	n	M (SD)
Original ¹	19	1.88 (1.38)	19	2.54 (1.95)	29	0.02 (2.38)
Replication ²	20	0.98 (1.20)	27	0.02 (1.88)	28	0.27 (1.72)

¹Monin, B., Sawyer, P. J., & Marquez, M. J. (2008). The rejection of moral rebels: Resenting those who do the right thing. *Journal of Personality and Social Psychology*, 95(1), 76. doi: 10.1037/0022-3514.95.1.76

²Frank, M. C., & Holubar, T. (2015, Dec). *Replication of Monin, Sawyer, & Marquez (2008, JPSP 95(1), exp. 4)*. Open Science Framework. Retrieved from osf.io/pz0my

01

Calculating the prior predictive *p*-value

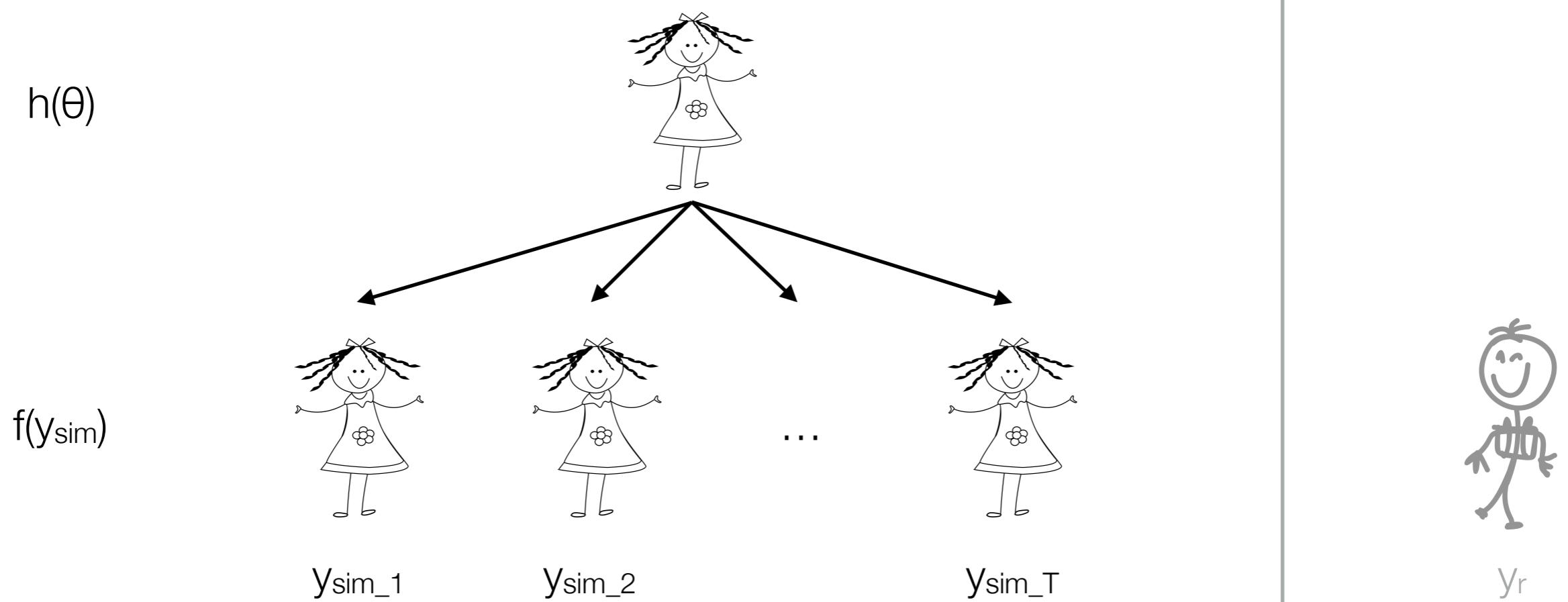
Box, 1980. $H_0: \boldsymbol{\theta} \sim h(\boldsymbol{\theta})$

Step 1: Prior Predictive Distribution of \mathbf{y}_o

$$h(\boldsymbol{\theta}_r) = g(\boldsymbol{\theta}_o | \mathbf{y}_o)$$

$$\begin{aligned} f(\mathbf{y}_{\text{sim}} | \mathbf{y}_o) &= \int f(\mathbf{y}_{\text{sim}} | \boldsymbol{\theta}_o) g(\boldsymbol{\theta}_o | \mathbf{y}_o) d\boldsymbol{\theta}_o \\ &= \int f(\mathbf{y}_{\text{sim}} | \boldsymbol{\theta}_r) h(\boldsymbol{\theta}_r) d\boldsymbol{\theta}_r = f(\mathbf{y}_{\text{sim}}) \end{aligned}$$

Prior predictive p -value: Step 1



Step 2: Hypotheses & Test Statistic

Study	Obedient		Rebel Self-Affirmation		Rebel Bogus	
	n	M (SD)	n	M (SD)	n	M (SD)
Original	19	1.88 (1.38)	19	2.54 (1.95)	29	0.02 (2.38)

What to replicate?

Step 2: Hypotheses & Test Statistic

Study	Obedient		Rebel Self-Affirmation		Rebel Bogus	
	n	M (SD)	n	M (SD)	n	M (SD)
Original	19	1.88 (1.38)	19	2.54 (1.95)	29	0.02 (2.38)

What to replicate?

$$H_01: (\mu_{ob}, \mu_{rsa}) > \mu_{rb}$$

Step 2: Hypotheses & Test Statistic

Study	Obedient		Rebel Self-Affirmation		Rebel Bogus	
	n	M (SD)	n	M (SD)	n	M (SD)
Original	19	1.88 (1.38)	19	2.54 (1.95)	29	0.02 (2.38)

What to replicate?

$$H_{01}: (\mu_{ob}, \mu_{rsa}) > \mu_{rb}$$

$$H_{02}: \mu_{ob} - \mu_{rb} \geq 1.50, \mu_{rsa} - \mu_{rb} \geq 2.00$$

Step 2: Hypotheses & Test Statistic

Study	Obedient		Rebel Self-Affirmation		Rebel Bogus	
	n	M (SD)	n	M (SD)	n	M (SD)
Original	19	1.88 (1.38)	19	2.54 (1.95)	29	0.02 (2.38)

What to replicate?

$$H_{01}: (\mu_{ob}, \mu_{rsa}) > \mu_{rb}$$

$$H_{02}: \mu_{ob} - \mu_{rb} \geq 1.50, \mu_{rsa} - \mu_{rb} \geq 2.00$$

$$H_{03}: d_{ob,rb} \geq .80, d_{rsa,rb} \geq .80$$

Step 2: Hypotheses & Test Statistic

Study	Obedient		Rebel Self-Affirmation		Rebel Bogus	
	n	M (SD)	n	M (SD)	n	M (SD)
Original	19	1.88 (1.38)	19	2.54 (1.95)	29	0.02 (2.38)

What to replicate?

$$H_01: (\mu_{ob}, \mu_{rsa}) > \mu_{rb}$$

$$H_02: \mu_{ob} - \mu_{rb} \geq 1.50, \mu_{rsa} - \mu_{rb} \geq 2.00$$

$$H_03: d_{ob,rb} \geq .80, d_{rsa,rb} \geq .80$$

$$H_04: \mu_{ob} = 1.88, \mu_{rsa} = 2.54, \mu_{rb} = 0.02$$

Step 2: Hypotheses & Test Statistic

Study	Obedient		Rebel Self-Affirmation		Rebel Bogus	
	n	M (SD)	n	M (SD)	n	M (SD)
Original	19	1.88 (1.38)	19	2.54 (1.95)	29	0.02 (2.38)

What to replicate?

$$H_01: (\mu_{ob}, \mu_{rsa}) > \mu_{rb}$$

$$H_02: \mu_{ob} - \mu_{rb} \geq 1.50, \mu_{rsa} - \mu_{rb} \geq 2.00$$

$$H_03: d_{ob,rb} \geq .80, d_{rsa,rb} \geq .80$$

$$H_04: \mu_{ob} = 1.88, \mu_{rsa} = 2.54, \mu_{rb} = 0.02$$

$$\bar{F}_{\mathbf{y}_d} = \frac{\text{RSS}_{d,H_{0z}} - \text{RSS}_{d,H_u}}{S_d^2}$$

Step 2: Hypotheses & Test Statistic

Study	Obedient		Rebel Self-Affirmation		Rebel Bogus	
	n	M (SD)	n	M (SD)	n	M (SD)
Original	19	1.88 (1.38)	19	2.54 (1.95)	29	0.02 (2.38)

What to replicate?

$$H_01: (\mu_{ob}, \mu_{rsa}) > \mu_{rb}$$

$$H_02: \mu_{ob} - \mu_{rb} \geq 1.50, \mu_{rsa} - \mu_{rb} \geq 2.00$$

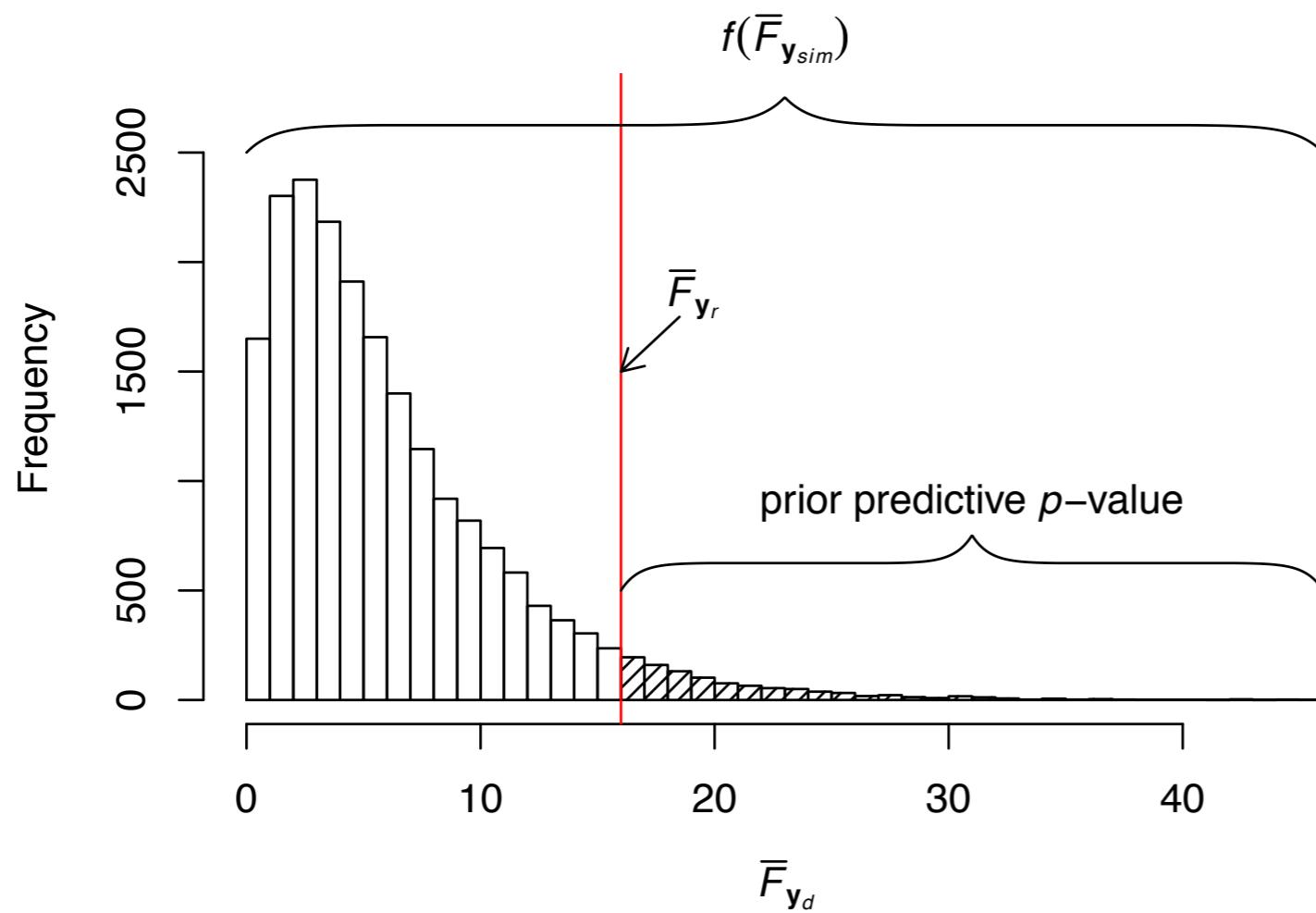
$$H_03: d_{ob,rb} \geq .80, d_{rsa,rb} \geq .80$$

$$H_04: \mu_{ob} = 1.88, \mu_{rsa} = 2.54, \mu_{rb} = 0.02$$

$$\left. \begin{array}{l} \\ \\ \\ \end{array} \right\} \bar{F} \longrightarrow f(\bar{F}_{y_sim}), \bar{F}_{y_r}$$

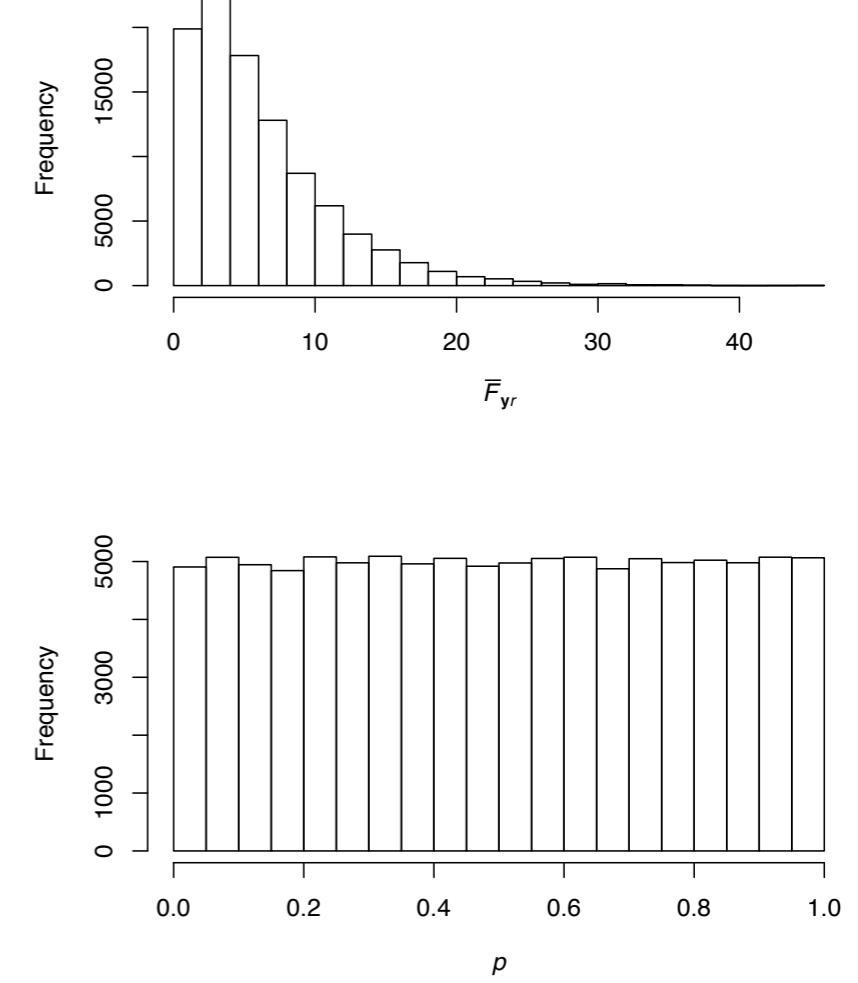
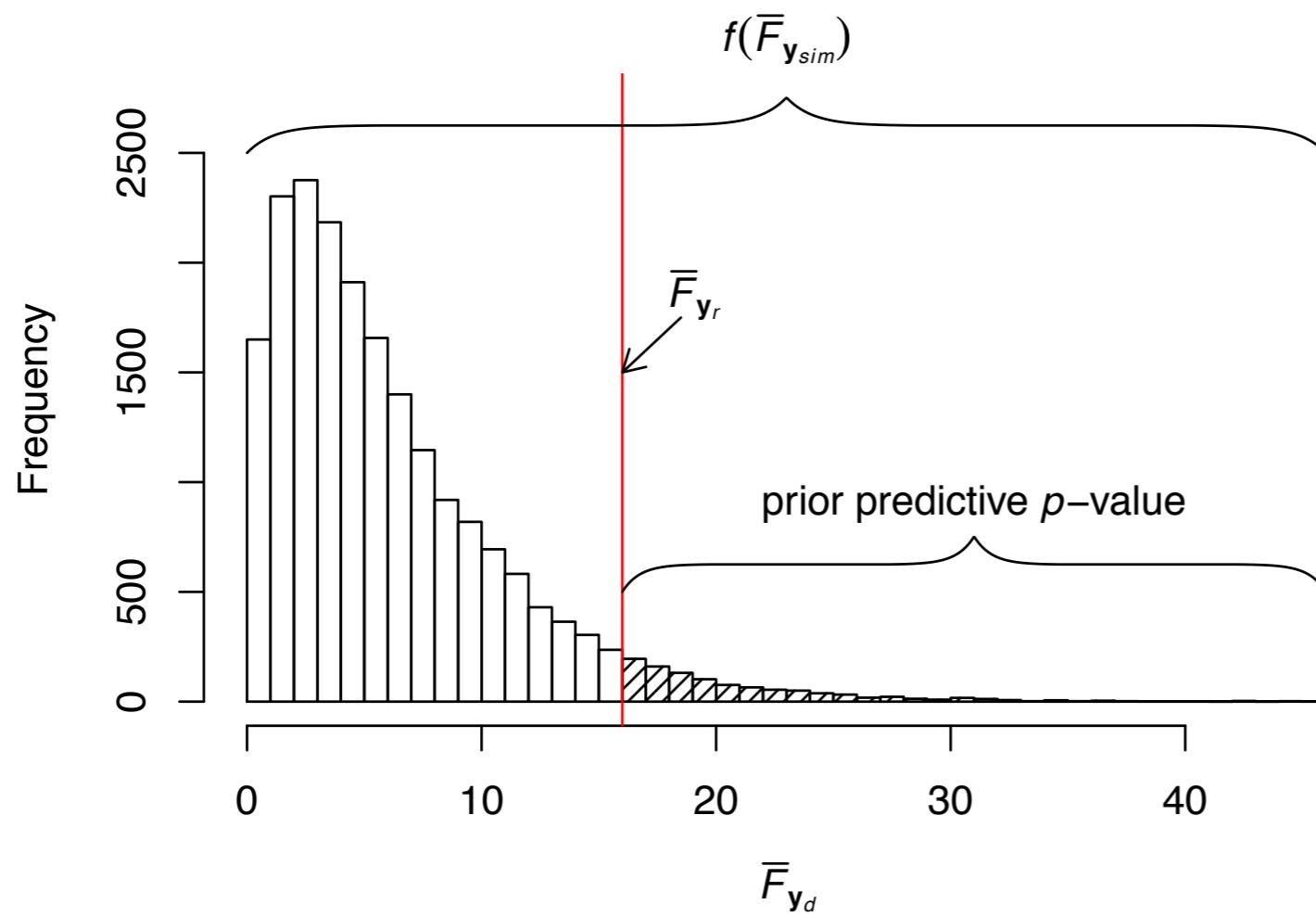
Step 3: p -value

$$P(\bar{F}_{\mathbf{y}_{\text{sim}}} \geq \bar{F}_{\mathbf{y}_r} | H_{0z})$$



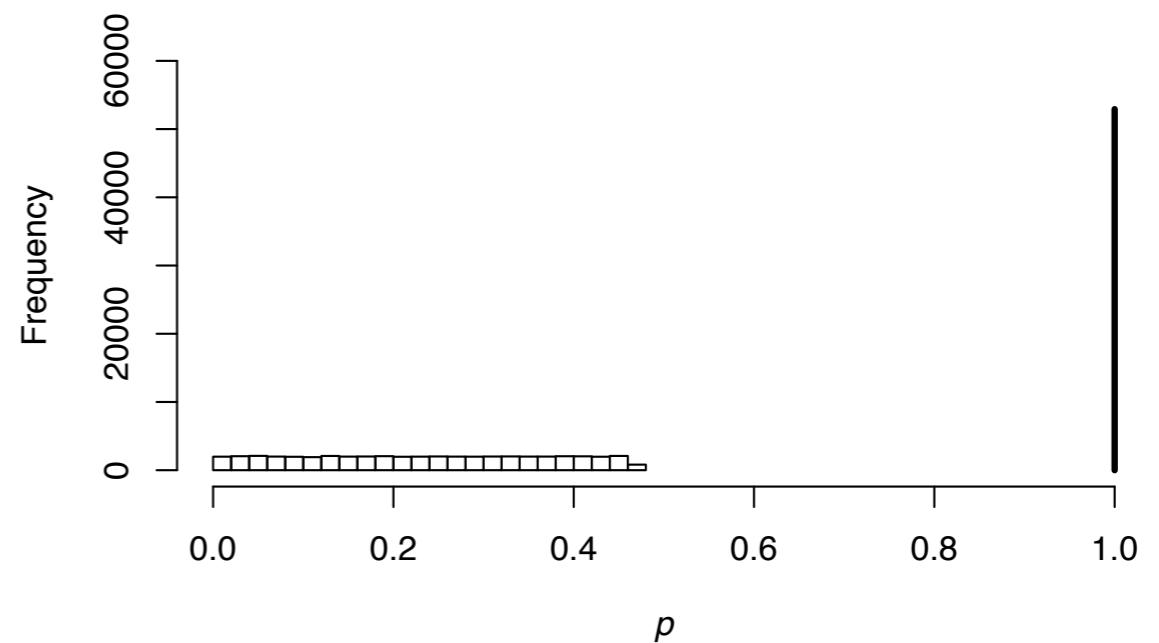
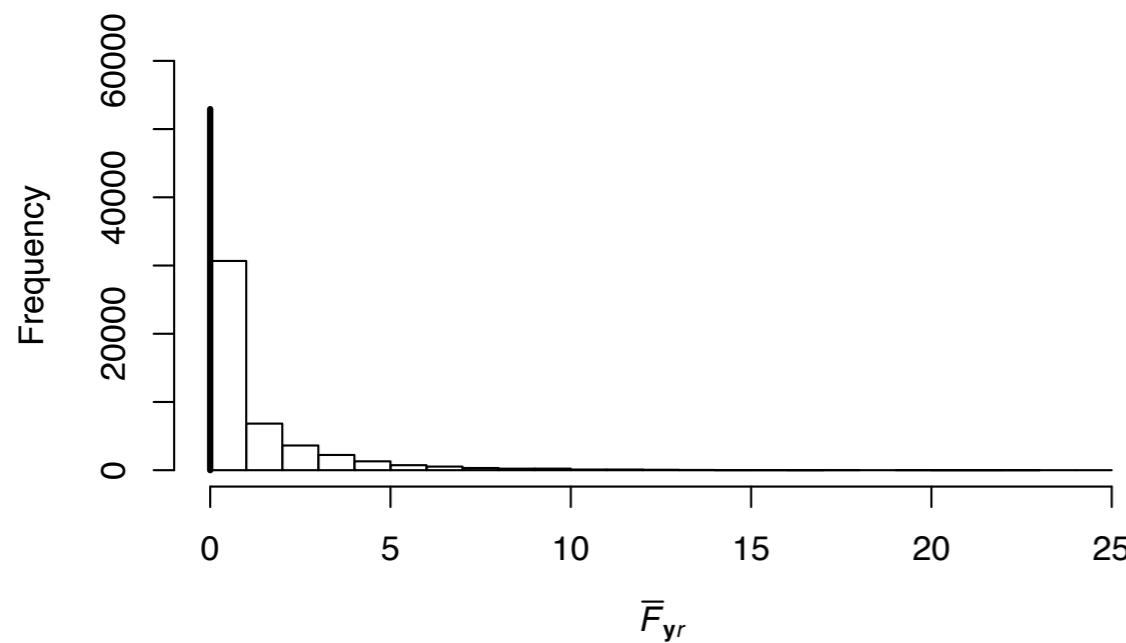
Step 3: p -value - Uniformity

$$f(p \leq \alpha | H_0) \leq \alpha \text{ for all } \alpha \in [0, 1]$$



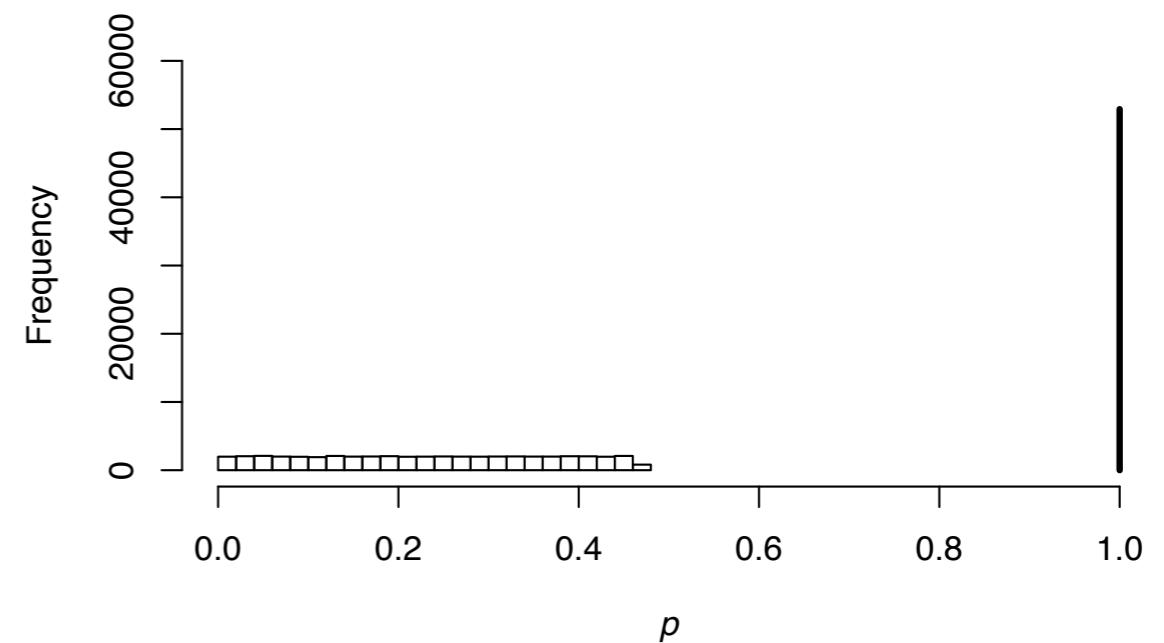
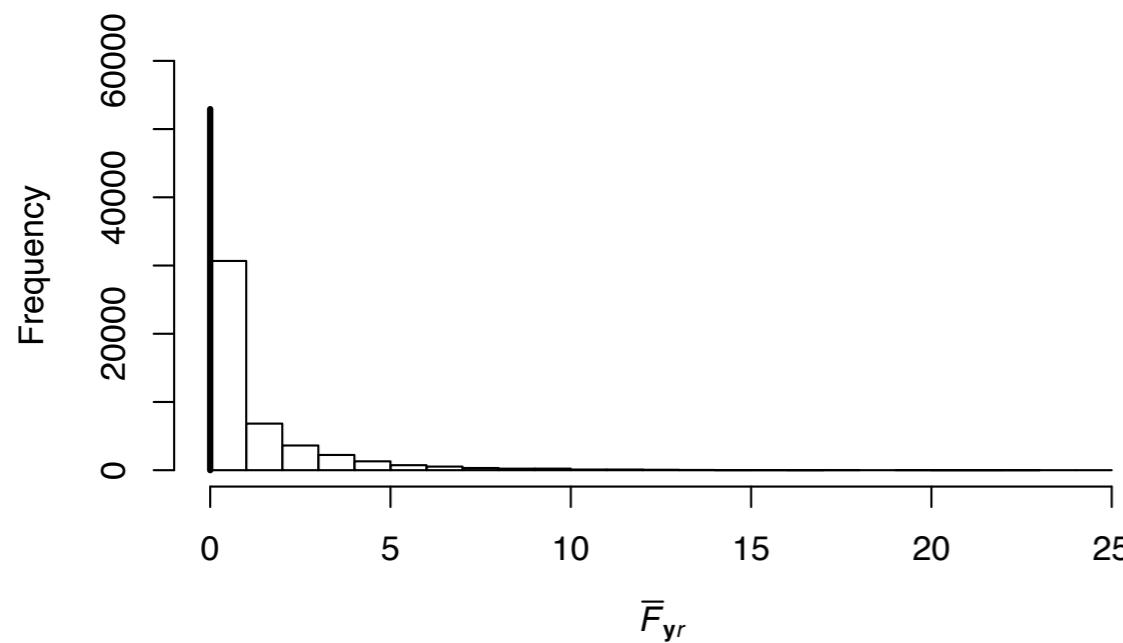
Step 3: p -value - Uniformity

$$f(p \leq \alpha | H_0) \leq \alpha \text{ for all } \alpha \in [0, 1]$$



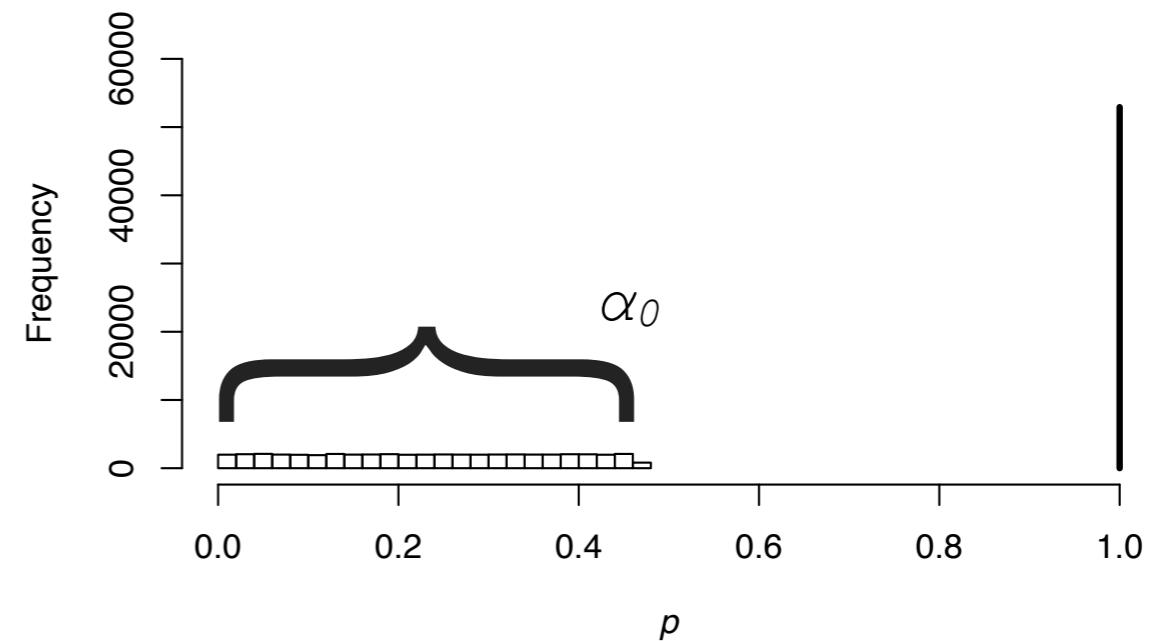
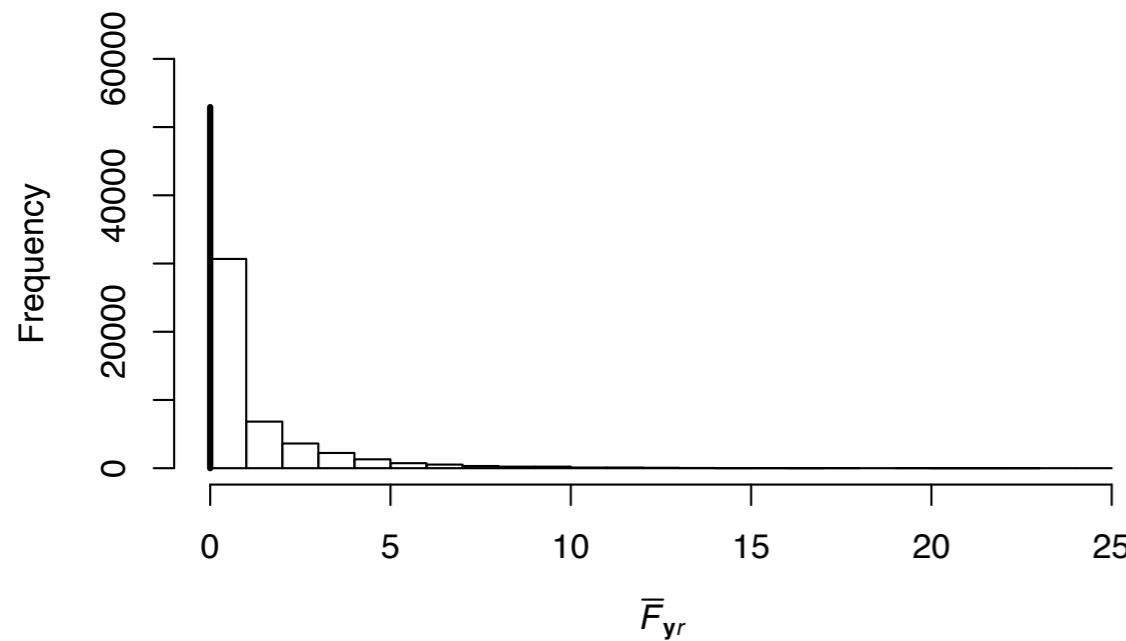
Step 3: p -value - Uniformity

$$f(p \leq \alpha | H_0) \leq \alpha \text{ for all } \alpha \in [0, 1]$$



Step 3: p -value - Uniformity

$$f(p \leq \alpha | H_0) \leq \alpha \text{ for all } \alpha \in [0, \alpha_0]$$



02

Interactive application

Zondervan-Zwijnenburg, M., Van de Schoot, R., & Hoijtink, H. (2017, May 15). ANOVA. <http://doi.org/10.17605/OSF.IO/6H8X3>

M. A. J. Zondervan-Zwijnenburg (2017). ANOVArePLICATION: Test ANOVA Replications by Means of the Prior Predictive p-Value. R package version 1.0.0. <https://CRAN.R-project.org/package=ANOVArePLICATION>
osf.io/6h8x3

ANOVA Replication App

This application is associated with the paper: Zondervan-Zwijnenburg, M.A.J., Van de Schoot, R., and Hoijtink, H. (2017). With the application the replication of specific ANOVA features can be tested by means of a sampling-based prior predictive check. Additionally, the total sample size can be calculated to reject replication for populations with equal means.

By using this app you agree to be bound by the Terms of Usage.

Original Study New Study Replication Test Sample Size Calculator

As a first step to test replication, a summary of the parameters based on the original data needs to be obtained. This summary is a posterior distribution. 1. Submit the original data, or provide descriptive statistics and click generate data to simulate the original data. 2. Click the button below to run a Gibbs sampler that results in the required posterior distribution. Afterwards, you can continue to upload (summary statistics of) the New Study in the next tab, or calculate the required sample size for a new study with the Sample Size Calculator

Type of input original study

Provide original data
 Provide original data descriptives

Number of groups:

2 **3** 10

Mean group 1 SD group 1 Sample size group 1
1.88 1.38 19

Mean group 2 SD group 2 Sample size group 2
2.54 1.95 19

Mean group 3 SD group 3 Sample size group 3
0.02 2.38 29

Generate data

Original Data New Data Replication Test Results Sample Size Output

Summary **Data**

The total sample size is: 67

Sample size per group

1	2	3
19	19	29

Mean by group

Group	Mean
1	1 1.88
2	2 2.54
3	3 0.02

Bayesian summary

As a first step to test replication, a summary of the parameters based on the original data needs to be obtained. This summary is a posterior distribution. 1. Submit the original data, or provide descriptive statistics and click generate data to simulate the original data. 2. Click the button below to run a Gibbs sampler that results in the required posterior distribution. Afterwards, you can continue to upload (summary statistics of) the New Study in the next tab, or calculate the required sample size for a new study with the Sample Size Calculator

Type of input original study

- Provide original data
- Provide original data descriptives

Number of groups:

3

Mean group 1

1.88

SD group 1

1.36

Sample size group 1

19

Mean group 2

2.54

SD group 2

1.95

Sample size group 2

19

Mean group 3

0.02

SD group 3

2.38

Sample size group 3

29

[Generate data](#)

Obtain posterior distribution

Burnin iterations will be discarded. A Bayesian sampler typically runs several thousand post-burnin iterations. We recommend to start with 5000 post-burnin iterations for each of the two chains. Trace plots will appear in the main panel. If convergence is not obtained, increase the number of iterations below

Number of burnin iterations to run per chain Number of (post-burnin) iterations to run per chain To obtain fixed results, set a seed value other than 0.

500

5000

1

[Obtain the posterior](#)

The total sample size is: 67

Sample size per group

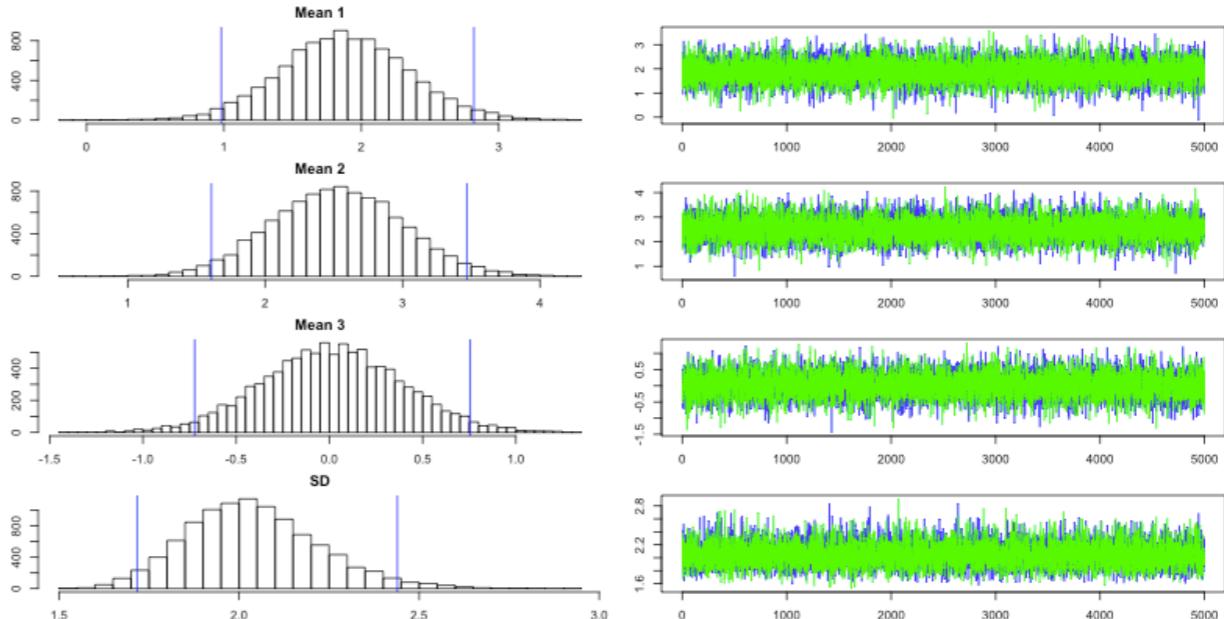
1	2	3
19	19	29

Mean by group

Group	Mean
1	1 1.88
2	2 2.54
3	3 0.02

Bayesian summary

The plots below help you to examine the convergence of your model. The right column shows traceplots. The traceplots should look like fat caterpillars, that is, they should have stable means and variances. If the traceplots do not have stable means and variances, you should increase the number of iterations to obtain the posterior distribution. The plots on the left are the posterior distributions: the outcomes of the analysis.



The table below gives the output of the Bayesian analysis on the original data. In the first column the means of the marginal posterior distributions are given, while the second column provides the standard deviation of these distributions. The third and fourth column show the limits of the 95% credible interval.

ANOVA Replication App

This application is associated with the paper: Zondervan-Zwijnenburg, M.A.J., Van de Schoot, R., and Hoijtink, H. (2017). With the application the replication of specific ANOVA features can be tested by means of a sampling-based prior predictive check. Additionally, the total sample size can be calculated to reject replication for populations with equal means.

By using this app you agree to be bound by the Terms of Usage.

Original Study New Study Replication Test Sample Size

Submit the data of the new study (sometimes called replication study), or provide descriptive statistics to simulate the new data. This input is not required for sample size calculations.

Type of input new study

Provide new data
 Provide new data descriptives

Upload a csv file with in the first column the dependent variable, and in the second column a categorical variable indicating group membership. Use the options below to let the app read your file correctly. You can check your upload in the panel on the right.

Choose CSV File

Choose File Upload complete

Header

Separator

Comma
 Semicolon
 Tab

Documents

data_o.csv
 data_r.csv

2	2	0.01851852
3	3	0.27380952

Cancel Open

1. Specify the features of the original study that are relevant to replication. Base your hypothesis on the original study. 2. Run the prior predictive check to obtain the posterior predictive p-value.

Type of hypothesis

- Inequality constraint(s)
- Exact mean values

Number of constraints:

1 2 3 4 5 6 7 8 9 10

Specify your pairwise contrasts below. The data consists of a number of groups, which each have a mean. If the mean of group 1 is expected to be larger than that of group 2 based on the original data, your contrast is: 1>2. The larger mean of the contrast is to be put on the left, while the smaller mean of the contrast is to put on the right side of the > symbol.

Constraint 1

2 > 3

Constraint 2

1 > 3

Do you want to add minimum differences between means?

- Yes, minimum values
- Yes, minimum effect sizes
- No

Provide the minimum difference for each constraint below.

Difference 1

2.00

Difference 2

1.50

To obtain fixed results, set a seed value other than 0.

1

Note that running the analysis may take a couple of minutes.

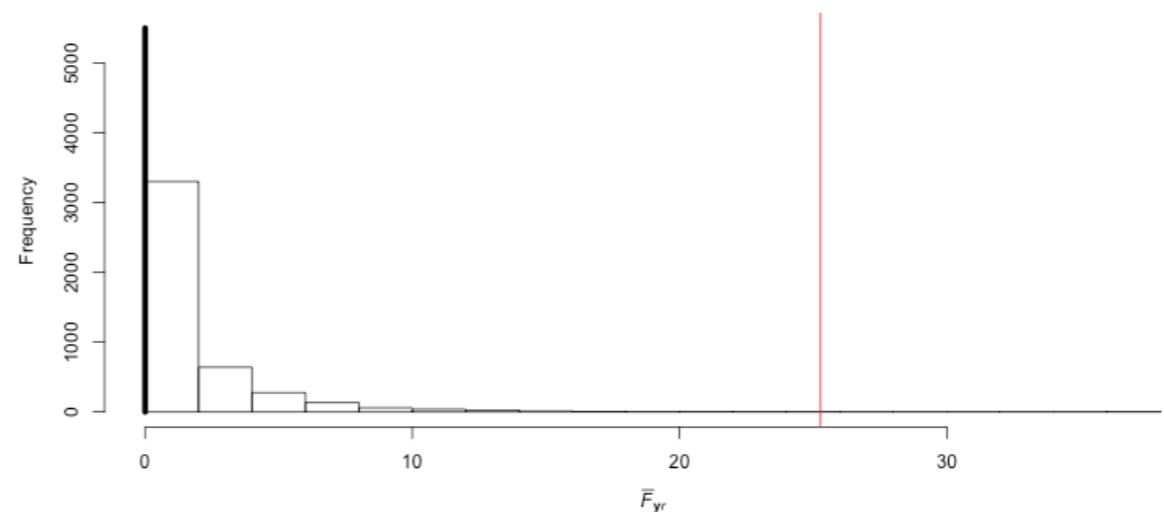
[Run the replication test](#)

Below you can find the results of the prior predictive check. The summary of the distribution of F-bar given the original data (also plotted in the histogram) demonstrates likely F-bar values for future studies given the parameter estimates from the original study. 'F-bar replication data' is the F-bar value obtained in the replication study (indicated with a red line in the histogram). The prior predictive p-value indicates how extreme the obtained F-bar in the replication study is compared to the F-bar values that are expected based on the original study.

```
$`distribution F-bar given original data`
V1
Min. : 0.0000
1st Qu.: 0.0000
Median : 0.0000
Mean   : 0.7612
3rd Qu.: 0.5667
Max.   : 36.3477

$`F-bar replication data`
[1] 25.27561

$prior predictive p-value
[1] 2e-04
```



[Download simulated F values](#)

[Download histogram](#)

Here you can determine the sample size for a new study with sufficient power to reject replication when all means are equal.

Upload or provide descriptives of the original data first in the tab 'Original study' and obtain the posterior distribution.

Type of hypothesis to be replicated

- Inequality constraint(s)
- Exact mean values

Type the exact values to replicate below separated by commas.

Exact mean values

1.88,2.54,0.02

Note that running the analysis may take a while. A trial run with max two iterations (option below) may give an indication for a good starting n (option below). In the right upper corner of the results tab, you see a progress indicator when the calculations are running. You can limit computation time by providing a better starting value for n, limiting the number of iterations in the Bayesian analysis (Original Data tab), limiting the number of iterations below, or by limiting the maximum total sample size below. Note that fewer iterations in the Bayesian analysis and in sample size calculations may reduce the quality of the calculations.

Target power Margin for target power Alpha level

.825 .025 .05

n per group to start calculations with

20

Maximum number of iterations

10

Maximum total sample size

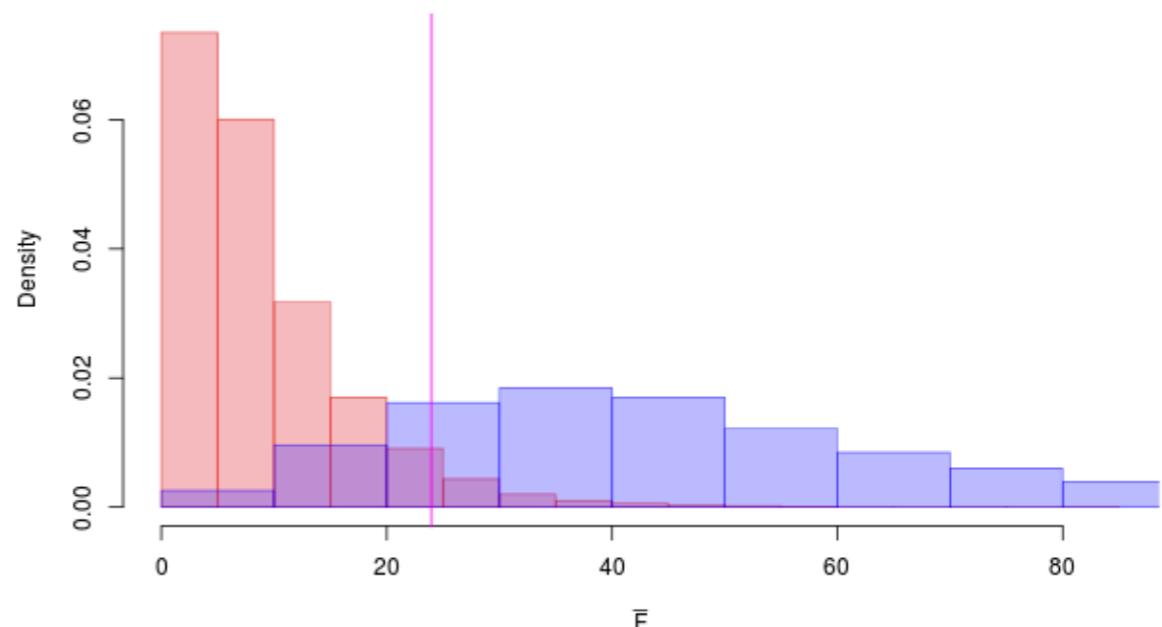
600

Run sample size calculations

Below you can find the results of the sample size calculator for the prior predictive check. First the output provides the reason to stop the iterative sample size calculations. Subsequently, the matrix with the output is given. The number in column n per group is the sample size per group, the value on the right is the associated power. The histogram is based on information for the last sample size and power calculated. The red distribution is composed of F-bars for datasets from a population in which replication holds. The blue distribution shows F-bars from a population with equal means for which replication should be rejected. The vertical line indicates the critical value located at 95% of the null distribution of replication. The proportion of the alternative distribution at the right side of the critical value is the statical power.

```
[[1]]
[1] "The target power level has been reached."
```

```
[[2]]
  n per group Power
 [1,]      20  0.66
 [2,]      26  0.73
 [3,]      29  0.76
 [4,]      42  0.82
```



03

Current developments

- More (e.g., SEM) models
- + R-package



Universiteit Utrecht