PRDA package: Prospective and Retrospective Design Analysis.

Andreella, A., Bertoldo, G., Vesely, A., Altoe, G. and Callegher, C.
June 19, 2020

The power ritual

The power ritual

- **Replication crisis** in social sciences and medicine $\rightarrow$ Starting to **promoting** large-scale replication effort.
- **However**, they found smaller effects than originals $\rightarrow$ **Decline effect**:
  - **What does not kill statistical significance makes it stronger**;
  - **Winner curse**.
- **Statistical inference** is often viewed as an **isolated procedure** $\rightarrow$ from **null ritual** to **power ritual**! :(
- From Gelman and Carlin (2014), effect size estimation and statistical significance are **closely related**:

1. **Type M (magnitude) error** (the exaggeration ratio): indicates the predictable average **overestimation** of an statistically significant effect;
2. **Type S (sign) error** (the sign error): indicates the probability to find a statistically significant effect in the **opposite direction** to the one considered plausible.

- **Design Analysis**: **Retrospective** and **Prospective**

The PRDA package

Introduction

How to install it?

```
#devtools::install_github("masspastore/PRDA")
library(PRDA)
```

Two implementations:

- **Pearson correlation**: $H_0 : \rho = 0$;
- **Cohen's d**: $H_0 : d = 0$ (one sample and two samples).

Retrospective Design Analysis

```
retrospective(sample_n1, effect_size, sample_n2 = NULL,
              effect_type = c("cohen_d","correlation"),
              alternative = c("two.sided","less","greater"),
              tl = -Inf, tu = Inf)
```

INPUT:

- `sample_n1`: sample size
- `effect_size`: value of the effect size or function indicating the hypothetical population effect size.
- `sample_n2`: sample size if two samples test is used
- `effect_type`: Which effect size do you want to analyze?
- `alternative`: Which alternative hypothesis?

OUTPUT:

- `power`, `typeM`, and `typeS` and some `Study characteristics`

Prospective Design Analysis

```
prospective(effect_size, power, ratio_n = 1,
            effect_type = c("cohen_d","correlation"),
            alternative = c("two.sided","less","greater"),
            tl = -Inf, tu = Inf)
```

INPUT:

- `effect_size`: effect size value of the study or function indicating the hypothetical population effect size
- `power`: power value of the study
- `ratio_n2`: ratio between `sample_n1` and `sample_n2`
- `effect_type`: Which effect size do you want to analyze?
- `alternative`: Which alternative hypothesis?

OUTPUT:

- `power`, `typeM`, `typeS`, and `Study characteristics`

Case Studies

Pearson correlation

We consider the study by Eisenberger et al. (2003) entitled: “Does Rejection Hurt? An fMRI Study of Social Exclusion”.

- The Anterior Cingulate Cortex (ACC) is involved in the experience of physical pain. **Could pain from social stimuli, such as social exclusion, share similar neural underpinnings?**
- Each of the 13 participants plays a virtual game with other **two fictitious players** while undergoing functional Magnetic Resonance Imaging (fMRI);
- Players had to toss a virtual ball among each other in three conditions: **social inclusion**, **explicit social exclusion** and **implicit social exclusion**;
- Each participant completed a **self-report measure** regarding their perceived distress;
- Correlation coefficient between perceived distress and activity in the ACC, $r = .88$, $p < .005$ and power equals $.13$.

Pearson correlation - Retrospective Design Analysis

We consider a **plausible effect size** equals $\rho = .25$ (Vul and Pashler, 2017):

```
retrospective(sample_n1 = 13, effect_size = 0.25, effect_type = "correlation", alternative = "two.sided",
              sig_level = 0.05)
```

```
##
##  Design Analysis
##
## Hypothesized effect:  rho = 0.25
##
## Study characteristics:
##    test_method   sample_n1   sample_n2   alternative   sig_level   df
##    pearson       13          13          two.sided     0.05        11
##
## Inferential risks:
##    power   typeM   typeS
##    0.13    2.599   0.023
##
## Critical value(s): rho  =  ± 0.553
```

Pearson correlation - Prospective Design Analysis

We want the 80% probability to detect a **plausible effect size** of at least $\rho = .25$ (Vul and Pashler, 2017):

```
prospective(effect_size = 0.25, power = 0.8, effect_type = "correlation", alternative = "two.sided")
```

```
##
##  Design Analysis
##
## Hypothesized effect:  rho = 0.25
##
## Study characteristics:
##    test_method   sample_n1   sample_n2   alternative   sig_level   df
##    pearson       122         122         two.sided     0.05        120
##
## Inferential risks:
##    power   typeM   typeS
##    0.801   1.121   0
##
## Critical value(s): rho  =  ± 0.178
```

Cohen's d

We consider the study by Kay et al., 2014 entitled "A functional basis for structure-seeking: Exposure to structure promotes willingness to engage in motivated action".

- 67 partecipants randomly assign to read two different types of text: natural phenomena are **unpredictable and random** or **predictable and systematic**.

- The outcome measure was the inclination to work towards a goal that each participant chose as their most important. The expected result was a **higher score in structure condition than random**.

- **Cohen's d** regarding the difference between the high score means under the two conditions equals 0.5, with t-test equals 2 and p-value 0.05.

Cohen's d - Retrospective Design Analysis

We consider a **plausible effect size** equals $d = .35$ (Open Science Collaboration, 2015):

```
retrospective(sample_n1 = 34, sample_n2 = 33, effect_size = 0.2, effect_type = "cohen_d",
              alternative = "two.sided", sig_level = 0.05)
```

```
##
##  Design Analysis
##
## Hypothesized effect:  cohen_d = 0.2
##
## Study characteristics:
##    test_method   sample_n1   sample_n2   alternative   sig_level   df
##    welch         34          33          two.sided     0.05        64.94
##
## Inferential risks:
##    power   typeM   typeS
##    0.123   3.084   0.023
##
## Critical value(s): cohen_d  =  ± 0.488
```

Cohen's d - Prospective Design Analysis

We want the 80% probability to detect a **plausible effect size** of at least $d = .35$ (Open Science Collaboration, 2015):

```
prospective(effect_size = 0.35, power = 0.8, effect_type = "cohen_d", alternative = "two.sided")
```

```
##
##  Design Analysis
##
## Hypothesized effect:  cohen_d = 0.35
##
## Study characteristics:
##    test_method   sample_n1   sample_n2   alternative   sig_level   df
##    welch         126         126         two.sided     0.05        250
##
## Inferential risks:
##    power   typeM   typeS
##    0.791   1.132   0
##
## Critical value(s): cohen_d  =  ± 0.248
```

Cohen's d - Prospective Design Analysis

We can also explore inferential risk:

```
prospective(effect_size = 0.35, power = 0.6, effect_type = "cohen_d", alternative = "two.sided")
```

```
##
##  Design Analysis
##
## Hypothesized effect:  cohen_d = 0.35
##
## Study characteristics:
##    test_method   sample_n1   sample_n2   alternative   sig_level   df
##    welch         80          80          two.sided     0.05        158
##
## Inferential risks:
##    power   typeM   typeS
##    0.596   1.302   0
##
## Critical value(s): cohen_d  =  ± 0.312
```

Take home messages

Design Analysis:

- **Type M** and **Type S** errors quantify the **inferential risks** in terms of effect size estimation $\longrightarrow$ **surpass power ritual**
- Contribute to planning more **robust and replicable studies**;
- Use information **outside the data** at hand! Rather than focusing only on a single pilot or published study;
- Contribute **planning** to planning more robust and replicable studies (**PROSPECTIVE**) and to **evaluate already conducted studies** (**RETROSPECTIVE**);
- Exploration of different **scenarios**;
- **Further directions**: other effect sizes, Bayesian approach (Bayes Factor).

Slides and references on https://github.com/angeella/eRum_2020 and Package repository on https://github.com/masspastore/PRDA.

"Accept uncertainty. Be thoughtful, open, and modest.

Remember ATOM."

Wasserstein et al. (2019, p. 2)