3 Data collection & samples: Assignment

Instructions

Submit your answers to the 8 first exercises through Aula Global. Upload the remaining exercise concerning peer review and your pseudonym through Aula Global as well.

Describing samples

Sample 1

> You want to study whether speed of second language acquisition changes depending on the first language of learners. To this end, you recruit a sample of 50 monolingual first year students enrolled at UPF who just started their first class of Catalan.

1. Is Sample 1 a representative sample? Is it a complete sample?

Sample 2

> You want to study whether the speed of reaching B1 in Catalan changes depending on the first language of learners. To this end, you recruit a sample of 50 monolingual first year students enrolled at UPF who just started their first class of Catalan.

2. Is Sample 2 a representative sample? Is it a complete sample?

Sample 3

> You want to study whether the speed of reaching B1 in Catalan by first year students at UPF changes depending on the first language of learners. To this end, you recruit a sample of 50 monolingual first year students enrolled at UPF who just started their first class of Catalan.

3. Is Sample 3 a representative sample? Is it a complete sample?

Sample 4

> You want to study whether the speed of reaching B1 in Catalan by first year students at UPF changes depending on the first language of learners. To this end, you recruit a sample of all monolingual first year students enrolled at UPF who just started their first class of Catalan.

4. Is Sample 4 representative sample? Is it a complete sample?

Sample 5

> You want to study whether the speed of reaching B1 in Catalan by monolingual first year students at UPF changes depending on the first language of learners. To this end, you recruit a sample made up of all monolingual first year students enrolled at UPF who just started their first class of Catalan.

5. Is Sample 5 representative sample? Is it a complete sample?

_____

Sample 6

You want to study whether the speed of reaching B1 in Catalan by monolingual first year students at UPF changes depending on the first language of learners. To this end, you recruit a sample made up of 50 randomly chosen monolingual first year students enrolled at UPF who just started their first class of Catalan and 50 randomly chosen monolingual first year students enrolled at UPF who just started their first class of German.

6. Is Sample 6 representative sample? Is it a complete sample?

Simulating samples

Here are some simulated samples of the number of gestures that speakers used in a formal vs. informal context an experimental trial.[1] Let's assume that speakers use an average of 55 gestures when speaking with a friend (informal context) and 45 when speaking with a professor (formal), and to be Poisson distributed.

You are tasked to determine what sample sizes, if any, from the six simulated below, give a good approximation of the true mean difference between formal and informal contexts. More precisely, determine which of the samples, if any, gives an estimate at least within 4 average gesture of the true mean difference of the contexts.

Sample 1 (1 subject per condition)

```
set.seed(122) #this makes the random draws be the same each time we run the code

sample1_informal <- rpois(n = 1, lambda = 55)   #1 draw from Poisson with lambda = 55
sample1_formal <-  rpois(n = 1, lambda = 45)     #1 draw from Poisson with lambda = 45
```

Sample 2 (5 subjects per condition)

```
set.seed(11) #this makes the random draws be the same each time we run the code

sample2_informal <- rpois(n = 5, lambda = 55)   #5 draws from Poisson  with lambda = 55
sample2_formal <-  rpois(n = 5, lambda = 45)    #5 draws from Poisson with lambda = 45
```

Sample 3 (15 subjects per condition)

```
set.seed(2) #this makes the random draws be the same each time we run the code

sample3_informal <- rpois(n = 15, lambda = 55)   #15 draws from Poisson with lambda = 55
sample3_formal <-  rpois(n = 15, lambda = 45)    #15 draws from Poisson with lambda = 45
```

_____

[1]We will look at the actual data from this experiment in future sessions Brown et al. (n.d.)

Sample 4 (40 subjects per condition)

```r
set.seed(101) #this makes the random draws be the same each time we run the code

sample4_informal <- rpois(n = 40, lambda = 55)   #40 draws from Poisson  with lambda = 55
sample4_formal <-  rpois(n = 40, lambda = 45)    #40 draws from Poisson  with lambda = 45
```

Sample 5 (100 subjects per condition)

```r
set.seed(101) #this makes the random draws be the same each time we run the code

sample5_informal <- rpois(n = 100, lambda = 55) #100 draws from Poisson  with lambda = 55
sample5_formal <-  rpois(n = 100, lambda = 45)  #100 draws from Poisson  with lambda = 45
```

Sample 6 (1000 subjects per condition)

```r
set.seed(101) #this makes the random draws be the same each time we run the code

sample6_informal <- rpois(n = 1000, lambda = 55) #1000 draws from Poisson  with lambda = 55
sample6_formal <-  rpois(n = 1000, lambda = 45)  #1000 draws from Poisson  with lambda = 45
```

7. Which is the smallest sample that gives an estimate at least within 4 average gestures of the true mean difference?
8. Which sample over- or underestimates the true difference between gestures in formal and informal contexts the most?

Peer review (part I)

Submit answers to these questions concerning your analysis report as a **single page** PDF-file through Aula Global. This submission makes up 50% of your peer-review grade. Remember that **your submission must be anonymous**. That is: do not put any information that can identify you in the PDF nor in the name of the file (like your name).

1. What is your general research question?

2. Why do you think this question is interesting? What does an answer to it tell us?

3. What is your specific research question?

4. What kind of data would you use to address Question 2 if you had unlimited resources?

5. What kind of data are you planning to use to address Question 2 within the scope of this class?

   - How will you obtain it?
   - How much will you collect?
   - Do you think that is enough data to address Question 2? Why (not)?

Elaborate with as much detail as you (currently) have. It is perfectly acceptable if you answer in bullet points and single sentences. Two randomly selected peers will give you feedback on them.

Finally, remember to give yourself a pseudonym in Aula Global (**pseudonym task**). The peer-review will be made through these pseudonyms.

References

Brown, L, H Kim, I Hübscher, and B Winter. n.d. "Iconic Gestures Are Modulated by Social Context: A Study of Multimodal Politeness Across Two Cultures." *Gesture.*