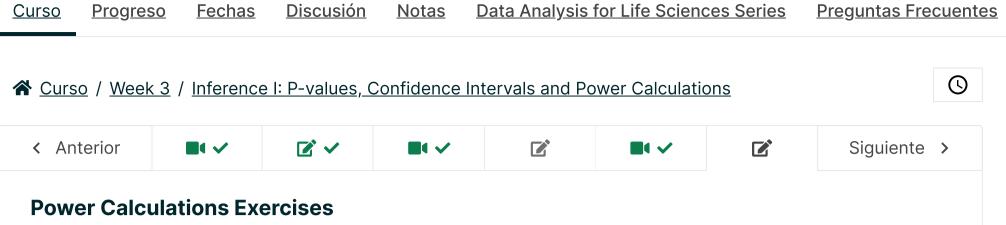


Data Analysis for Life Sciences Series <u>Discusión</u> <u>Progreso</u> <u>Fechas</u> <u>Notas</u>



☐ Marcar esta página

Homework fecha límite Jul 13, 2022 19:00 CEST

For these exercises we will load the babies dataset from babies.txt. We will use this data to review the concepts behind p-values and then test confidence interval concepts.

```
url <- "https://raw.githubusercontent.com/genomicsclass/dagdata/master/inst/extdata/babie
filename <- basename(url)
download(url, destfile=filename)
babies <- read.table("babies.txt", header=TRUE)</pre>
```

This is a large dataset (1,236 cases), and we will pretend that it contains the entire population in which we are interested. We will study the differences in birth weight between babies born to smoking and non-smoking mothers.

First, let's split this into two birth weight datasets: one of birth weights to non-smoking mothers and the other of birth weights to smoking mothers.

```
bwt.nonsmoke <- filter(babies, smoke==0) %>% select(bwt) %>% unlist
bwt.smoke <- filter(babies, smoke==1) %>% select(bwt) %>% unlist
```

Now, we can look for the true population difference in means between smoking and non-smoking birth weights.

```
library(rafalib)
mean(bwt.nonsmoke)-mean(bwt.smoke)
popsd(bwt.nonsmoke)
popsd(bwt.smoke)
```

The population difference of mean birth weights is about 8.9 ounces. The standard deviations of the nonsmoking and smoking groups are about 17.4 and 18.1 ounces, respectively.

As we did with the mouse weight data, this assessment interactively reviews inference concepts using simulations in R. We will treat the babies dataset as the full population and draw samples from it to simulate individual experiments. We will then ask whether somebody who only received the random samples would be able to draw correct conclusions about the population.

We are interested in testing whether the birth weights of babies born to non-smoking mothers are significantly different from the birth weights of babies born to smoking mothers.

Power Calculations Exercises #1

0/1 punto (calificado)

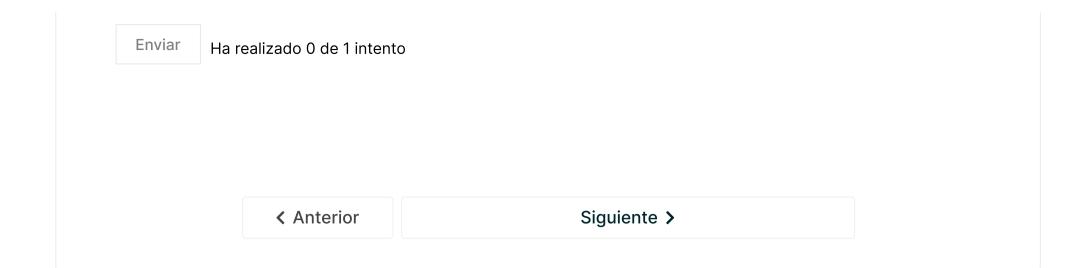
We can explore the trade off of power and Type I error concretely using the babies data. Since we have the full population, we know what the true effect size is (about 8.93) and we can compute the power of the test for true difference between populations.

Set the seed at 1 and take a random sample of N=5 measurements from each of the smoking and nonsmoking datasets. Use the t-test function to find the p-value. (Note that you already performed this calculation in the last assessment.)

The p-value is larger than 0.05 so using the typical cut-off, we would not reject. This is a type II error. Which of the following is *not* a way to decrease this type of error?

O Increase our chance of a type I error.	

×	a higher $lpha$ level.
Enviar	Ha realizado 2 de 2 intentos
Dower (Calculations Exercises #2
-	ible (calificable) ed at 1, then use the replicate() function to repeat the code used in the exercise above
	nes. What proportion of the time do we reject at the 0.05 level?
Enviar	Ha realizado 0 de 5 intentos
Power (Calculations Exercises #3
•	ible (calificable)
	not surprisingly, the power is lower than 10%. Repeat the exercise above for samples sizes of
30, 60, 90	and 120. Which of those four gives you power of about 80%?
Enviar	
LIIVIAI	Ha realizado 0 de 5 intentos
Power (Calculations Exercises #4
l punto pos	ible (calificable)
-	e problem above, but now require an $lpha$ level of 0.01. Which of those four gives you power of
about 80%	5?
Enviar	Ha realizado 0 de 5 intentos
-IIVIGI	
	Calculations Exercises #5
Power (
Power (ible (calificable)
Power (punto pos Consider t	
Power (I punto pos Consider t	ible (calificable) this statement: If you fail to reject the null hypothesis because the p value exceeds 0.05 and the your test is 90%, the chance that your finding is a false negative is 10%.



© Todos los Derechos están Reservados



edX

Acerca de

Afiliados

edX para negocios

Open edX

Carreras

Noticias

Legal

Condiciones de Servicio y Código de Honor

Política de privacidad

Políticas de Accesibilidad

Política de marcas

Mapa del Sitio

Contáctanos

Blog

Contáctenos

Centro de Ayuda

Kit Multimedia















© 2022 edX LLC. All rights reserved.

深圳市恒宇博科技有限公司 <u>粤ICP备17044299号-2</u>