edX

HarvardX PH525.1x
Statistics and R

Ayuda   Mguzmanfredes ⌄

Curso   Progreso   Fechas   Discusión   Notas   Data Analysis for Life Sciences Series   Preguntas Frecuentes

🏠 Curso / Week... / Inference II: Monte Carlo Simulation, Permutation Tests and Association t...

Permutations Exercises

🔖 Marcar esta página

Homework fecha límite Jul 13, 2022 19:00 CEST

Read through the permutation book chapter ☑.

We will use the following dataset to demonstrate the use of permutations:

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

Permutations Exercises #1

1 punto posible (calificable)
We will generate the following random variable based on a sample size of 10 and observe the following difference:

```
N=10
set.seed(1)
nonsmokers <- sample(bwt.nonsmoke , N)
smokers <- sample(bwt.smoke , N)
obs <- mean(smokers) - mean(nonsmokers)
```

The question is whether this observed difference is statistically significant. We do not want to rely on the assumptions needed for the normal or t-distribution approximations to hold, so instead we will use permutations. We will reshuffle the data and recompute the mean. We can create one permuted sample with the following code:

```
dat <- c(smokers,nonsmokers)
shuffle <- sample( dat )
smokersstar <- shuffle[1:N]
nonsmokersstar <- shuffle[(N+1):(2*N)]
mean(smokersstar)-mean(nonsmokersstar)
```

The last value is one observation from the null distribution we will construct. Set the seed at 1, and then repeat the permutation 1,000 times to create a null distribution. What is the permutation derived p-value for our observation?

Enviar   Ha realizado 0 de 5 intentos

Permutations Exercises #2

1 punto posible (calificable)
Repeat the above exercise, but instead of the differences in mean, consider the differences in m   ✏ Ocultar Notas

`obs <- median(smokers) - median(nonsmokers)` . What is the permutation based p-value?

Enviar   Ha realizado 0 de 5 intentos

‹ Anterior   Siguiente ›

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

Ocultar Notas