< Anterior | Siguiente >

T-test Exercises

🔖 Marcar esta página

To check your understanding of the t-test videos, please complete the following exercises.

For these exercises we will load the babies dataset from `babies.txt`. We will use this data to review the concepts behind the 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)
```

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.

T-test Exercises #1

0/1 punto (calificado)

Set the seed at 1 and obtain a samples from the non-smoking mothers (`dat.ns`) of size $N = 25$. Then, without resetting the seed, take a sample of the same size from and smoking mothers (`dat.s`). Compute the t-statistic (call it `tval`).

What is the absolute value of the t-statistic?

2.120904   ✖

2.120904

Ocultar Notas

T-test Exercises #2

0/1 punto (calificado)

Recall that we summarize our data using a t-statistics because we know that in situations where the null hypothesis is true (what we mean when we say "under the null") and the sample size is relatively large, this t-value will have an approximate standard normal distribution. Because we know the distribution of the t-value under the null, we can quantitatively determine how unusual the observed t-value would be if the null hypothesis were true.

The standard procedure is to examine the probability a t-statistic that actually does follow the null hypothesis would have larger absolute value than the absolute value of the t-value we just observed -- this is called a two-sided test.

We have computed these by taking one minus the area under the standard normal curve between `-abs(tval)` and `abs(tval)` . In R, we can do this by using the `pnorm()` function, which computes the area under a normal curve from negative infinity up to the value given as its first argument:

```
pval <- 1-(pnorm(abs(tval))-pnorm(-abs(tval)))
```

What is this estimated p-value?

0.03392985 ✘

0.03392985

Enviar Ha realizado 1 de 5 intentos

By reporting only p-values, many scientific publications provide an incomplete story of their findings. As we have mentioned, with very large sample sizes, scientifically insignificant differences between two groups can lead to small p-values. Confidence intervals are more informative as they include the estimate itself. We will learn how to compute these in the next module.

T-test Exercises #3

0/1 punto (calificado)

What would be the null hypothesis in this study?

○ There are no significant differences between birth weights of babies born to non-smoking mothers and the birth weights of babies born to smoking mothers.

○ There are significant differences between birth weights of babies born to non-smoking mothers and the birth weights of babies born to smoking mothers.

○ There are no significant differences between birth weights of babies born to non-smoking mothers in the sample (N=25) and the birth weights of babies born to smoking mothers in the sample (N=25).

● There are significant differences between birth weights of babies born to non-smoking mothers in the sample (N=25) and the birth weights of babies born to smoking mothers in the sample (N=25).

✘

Enviar Ha realizado 2 de 2 intentos

✎ Ocultar Notas

© 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

✏ Ocultar Notas