Práctica 6

Alejandro Cáceres
UPC - Statistics 2019/2020

Objetivo

- Intervalos de confianza

Intervalos de confianza

Generalmente hacemos un sólo experimento con n mediciones. Por ejemplo el resultado de medir la impedancia de 10 altavoces de 4 Ohmios es:

```
exp1 <- c(3.889533, 4.800898, 4.262898, 4.361508,
5.617505, 3.423186, 3.575229, 6.961411, 3.833203,
3.560329)

barx0 <- mean(exp1)

> barx0
[1] 4.42857
```

Intervalos de confianza

- tenemos n valores para las mediciones, que provienen de la variable aleatoria X.
- tenemos 1 valore para la media de las mediciones, que proviene de **otra** variable aleatoria $\bar{X}$.

Intervalos de confianza

Recuerda:

- hay 2 distribuciones una para X (distribución poblacional, para la que asumimos modelos de probabilidad; binomial, poisson, etc...) y otra para $\bar{X}$ (que intentamos averiguar con desde los modelos: TCL, t, χ^2).
- $\bar{x}$ lo obetnemos de los datos y lo usamos para estimar μ (un parámetro de X definido como $E(X)$ y que nunca observaremos diractamente).
- Si tomamos $\mu \sim \bar{x}_0$ ($\hat{\mu} = \bar{x}_0$) queremos una idea del %95 de valores que puede tomar $\bar{X}$ como una medida de confianza (precisión) de la estimación de μ dada por $\bar{x}_0$.

Intervalos de confianza

Buscamos el valor *xdiff* tal que

$$P(\bar{x}_0 - xdiff \leq \bar{X} \leq \bar{x}_0 + xdiff) = 0.95$$

o transformando en una variable standard

$$P(\frac{-xdiff}{\sigma_{\bar{X}}} \leq \frac{\bar{X} - \bar{x}_0}{\sigma_{\bar{X}}} \leq \frac{xdiff}{\sigma_{\bar{X}}}) = 0.95$$

o

$$P(\frac{-xdiff}{\sigma_X/\sqrt{n}} \leq \frac{\bar{X} - \bar{x}_0}{\sigma_X/\sqrt{n}} \leq \frac{xdiff}{\sigma_X/\sqrt{n}}) = 0.95$$

Intervalos de confianza

$$P(\frac{-xdiff}{\sigma_X/\sqrt{n}} \leq \frac{\bar{X}-\bar{x}_0}{\sigma_X/\sqrt{n}} \leq \frac{xdiff}{\sigma_X/\sqrt{n}}) = 0.95$$

Asumimos 1: que $\bar{X}$ es una distribución normal $N(\mu, \sigma)$

Asumimos 2: Que sabemos (nos dan) σ_X como información adicional. Por ejemplo el fabricante sabe que la desviación estándard de los amplificdores que produce es $\sigma_X = 1.4\Omega$

Entonces: $Z = \frac{\bar{X}-\bar{x}_0}{\sigma_X/\sqrt{n}}$ es una variable aleatoria estandard $N(\mu = 0, \sigma = 1)$:

$$P(z_{0.025} \leq Z \leq z_{0.975}) = 0.95$$

Intervalos de confianza

$$P(z_{0.025} \leq Z \leq z_{0.975}) = 0.95$$

```
> qnorm(0.025)
[1] -1.959964
> qnorm(0.975)
[1] 1.959964
```

o sea:

$$xdiff = 1.95 * \sigma_X / \sqrt{n}$$

Para nuestro experimento, tenemos n=10 y $\sigma_X = 1.4cm$

```
> xdiff<- 1.95*1.4/sqrt(10)
> xdiff
[1] 0.8633018
```

Intervalos de confianza

```
> CI=x0+c(-xdiff,xdiff)
> CI
[1] 3.503469 5.353670
```

En este intervalo tenemos una confianza del 95% en la precisión en la estimación de μ, la verdadera impedancia media de X, dada por $\bar{x}_0 = 4.42857$:

$$\bar{x}_0 = 4.4 \pm 0.8$$

al 95% de confianza. O sea: nos tomamos con cierto escepticismo el .4 en esta estimación, y los siguientes decimales los descartamos.

Intervalos de confianza

Los intervalos de confianza para

- un experimento de n muestras.
- de una distribución normal para X.
- del que sabemos la desviación estándard para X: σ_X.

se pueden calcular con la función
z.test

Intervalos de confianza

```
> install.packages("TeachingDemos") #solo se instala una vez
> library(TeachingDemos) #cada vez que queramos usar el paquete
> z.test(exp1,sd=1.4)


One Sample z-test

data:  exp1
z = 10.003, n = 10.00000, Std. Dev. = 1.40000, Std.
Dev. of the sample mean = 0.44272, p-value <
2.2e-16
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 3.560857 5.296283
sample estimates:
mean of exp1
     4.42857
```

Intervalos de confianza

Calculemos los intervalos de confianza para

- un experimento de n muestras
- una distribución para X normal (o desconocida si $n > 30$)
- donde **no** sabemos su desviación estándard

Intervalos de confianza

Recordemos que queremos:

$$P(\frac{-xdiff}{\sigma_X/\sqrt{n}} \leq \frac{\bar{X} - \bar{x}_0}{\sigma_X/\sqrt{n}} \leq \frac{xdiff}{\sigma_X/\sqrt{n}}) = 0.95$$

Asumimos 1: que $\bar{X}$ es una distribución normal $N(\mu, \sigma)$
Asumimos 2: posemos sustituir σ_x por su estimador $s = \frac{1}{n-1}\Sigma_i(x_i - \bar{x})^2$ que es calculado de los datos.
Entonces: $T = \frac{\bar{X}-\bar{x}_0}{\sigma_X/\sqrt{n}}$ es una variable de una distyribucion t-student con n-1 grados de libertad $t(n-1)$:

Intervalos de confianza

$$P(\frac{-xdiff}{s/\sqrt{n}} \leq \frac{\bar{X} - \bar{x}_0}{s/\sqrt{n}} \leq \frac{xdiff}{s/\sqrt{n}}) = 0.95$$

$$P(t_{0.025,n-1} \leq T \leq t_{0.975,n-1}) = 0.95$$

siendo T una variable aleatoria bajo student-t con n-1 grados de libertad

```
> qt(0.025,9)
[1] -2.262157
> qt(0.975,9)
[1] 2.262157
```

$$xdiff = 2.262157 * s/sqrt(10)$$

Intervalos de confianza

```
x0 <- mean(exp1)
> x0
[1] 4.42857
s <-sd(exp1)
> s
[1] 1.112419
```

Entonces:

```
> xdiff<-2.262157*s/sqrt(10)
> xdiff
[1] 0.7957762
> CI<-x0+c(-xdiff,xdiff)
> CI
[1] 3.632794 5.224346
```

Intervalos de confianza

En este caso se usa la fución **t.test**

```
> t.test(exp1,conf.level = 0.95)

One Sample t-test

data:  exp1
t = 12.589, df = 9, p-value = 5.114e-07
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 3.632794 5.224346
sample estimates:
mean of x
```

Intervalos de confianza

Por último la varianza.
Para un experimento queremos saber el intervalo de confiaza para su varianza muestral, cuando sabemos la varianza de la población σ^2 (1.4^2)

```
> s<-sd(exp1)
> s^2
[1] 1.112419
```

Recordemos que la variable
$X = s^2(n-1)/\sigma^2 \rightarrow \chi^2_{n-1}$

Intervalos de confianza

Buscamos

$$P(\chi^2_{0.025,n-1} \leq S^2(n-1)/\sigma^2 \leq \chi^2_{0.975,n-1}) = 0.95$$

Que en términos de la varianza es

$$P(\frac{s^2(n-1)}{\chi^2_{0.975,n-1}}, S^2, \frac{s^2(n-1)}{\chi^2_{0.025,n-1}}) = 0.95$$

```
> chi1<-qchisq(0.025, df=n-1)
> chi1
[1] 2.700389
> chi2<-qchisq(0.975, df=n-1)
> chi2
[1] 19.02277
```

Intervalos de confianza

```
> n<-10
> CI<-c(s^2*(n-1)/chi2,s^2*(n-1)/chi1)
> CI
[1] 0.5854708 4.1243219
```

Que en R se calcula con **sigma.test**

```
> sigma.test(exp1)

One sample Chi-squared test for variance

data:  exp1
X-squared = 11.137, df = 9, p-value = 0.5328
alternative hypothesis: true variance is not equal to 1
95 percent confidence interval:
 0.5854708 4.1243219
sample estimates:
var of exp1
   1.237475
```