Ejercicios Tema 3 - Intervalos de Confianza. Taller 1

Ricardo Alberich, Juan Gabriel Gomila y Arnau Mir

Curso completo de estadística inferencial con R y Python

Contenidos

1 Estimación por intervalos taller 1

1.1 Ejercicio 1

De una población de barras de hierro se extrae una muestra de 64 barras y se calcula la resistencia a la rotura por tracción se obtiene que $\overline{X} = 1012\ Kg/cm^2$. Se sabe por experiencia que en este tipo de barras $\sigma = 25$ y que la resistencia a la rotura sigue la distribución normal. Calcular un intervalo de confianza para μ al nivel 0.95.

1.1.1 Solución

El intervalo para la media es

```
n=64
xmedia=1012
sigma=25
alpha=1-0.95
IC=c(xmedia-qnorm(1-alpha/2)*sigma/sqrt(n),xmedia+qnorm(1-alpha/2)*sigma/sqrt(n))
IC
```

```
## [1] 1005.875 1018.125
```

Con un nivel de confianza del 95% el intervalo $(1005.8751125, 1018.1248875)$ contiene el verddero valor de la ressitencia a la rotura de estas barras de hierro.

1.2 Ejercicio 2

Para investigar el C.I. medio de una cierta población de estudiantes, se realiza un test a 400 estudiantes. La media y la desviación típica muestrales obtenidas son $\overline{x} = 86$ y $\tilde{s}_X = 10.2$. Calcular un intervalo para μ con un nivel de significación del 98%.

1.2.1 Solución

Como el tamaño de la muestra es grande aproximaremos por una distribución normal aproximando la desviación típica poblacional por la muestral. El intervalo obtenido es

```
n=400
xmedia=86
stilde=10.2
alpha=1-0.98
IC=c(xmedia-qnorm(1-alpha/2)*stilde/sqrt(n),xmedia+qnorm(1-alpha/2)*stilde/sqrt(n))
IC
```

```
## [1] 84.81356 87.18644
```

Con un nivel de confianza del 98% el intervalo $(84.8136, 87.1864)$ contiene el verdadero valor del C.I. medio de los estudiantes, ## Ejercicio 3

Para investigar un nuevo tipo de combustible para cohetes espaciales, se disparan cuatro unidades y se miden las velocidades iniciales. Los resultados obtenidos, expresados en Km/h, son :19600, 20300, 20500, 19800. Calcular un intervalo para la velocidad media μ con un nivel de confianza del 95%, suponiendo que las velocidades son normales.

1.2.2 Solución

En este caso nos dicen que la población sigue una distribución normal y que el tamañó de la muestra $n = 4$ es pequeño, así que utilizremos la t de Student para el cálculo de los cuantiles de intervalo.

```
n=4
muestra=c(19600, 20300, 20500, 19800)
xmedia=mean(muestra)
xmedia
```

```
## [1] 20050
```

```
stilde=sd(muestra)
stilde
```

```
## [1] 420.3173
```

```
alpha=1-0.95
IC=c(xmedia-qt(1-alpha/2,df=n-1)*stilde/sqrt(n),xmedia+qt(1-alpha/2,df=n-1)*stilde/sqrt(n))
IC
```

```
## [1] 19381.18 20718.82
```

Al nivel de confianza del 95% el intervalo $(19381.1813165, 20718.8186835)$ contiene el verdadero valor de la velocidad media de estos cohetes.

1.3 Ejercicio 4

Un fabricante de cronómetros quiere calcular un intervalo de estimación de la desviación típica del tiempo marcado en 100 horas por todos los cronómetros de un cierto modelo. Para ello pone en marcha 10 cronómetros del modelo durante 100 horas y encuentra que $\tilde{s}_X = 50$ segundos. Encontrar un intervalo para el parámetro σ^2 con $\alpha = 0.01$, suponiendo que la población del tiempo marcado por los cronómetros es normal.

1.3.1 Solución

Ekl tamaño de la muestra $n = 10$ no es muy grande pero este caso la distribución poblacional del tiempo es normal .

```
n=10
stilde=50
alpha=0.01 # nivel de confianza del 99%
IC=c((n-1)*stilde^2/qchisq(1-alpha/2,df = n-1),(n-1)*stilde^2/qchisq(alpha/2,df = n-1))
IC
```

```
## [1]   953.8202 12968.8012
```

Al nivel de confianza del 99% el intervalo $(953.8202305, 12968.8012344)$ contiene el verdadero valor de la varianza σ^2 del tiempo.

1.4 Ejercicio 5

Un auditor informático quiere investigar la proporción de rutinas de un programa que presentan una determinada irregularidad. Para ello observa 120 rutinas, resultando que 30 de ellas presentan alguna irregularidad. Con estos datos buscar unos límites de confianza para la proporción p de rutinas de la población que presentan esa irregularidad con probabilidad del 95%.

1.4.1 Solución

En este problema podemos utilizar distintas aproximaciones el intervalo. Siempre es mejor la exacta pero es posible que, para muestras grandes, tengan menor error de cálculo las soluciones aproximadas.

Calculemos varias de ellas; empecemos cargando los datos del problema:

```
n=120
n
```

```
## [1] 120
```

```
x=30# datos bernoulli a 1
x
```

```
## [1] 30
```

```
p_muestral=x/n  # proporción muestral
p_muestral
```

```
## [1] 0.25
```

- Método exacto Cloper Pearson

```
#install.packages("epitools") # descomentar para intalar epitools
epitools::binom.exact(x=x,n=n,conf.level=0.95)
```

```
##    x   n proportion     lower     upper conf.level
## 1 30 120       0.25 0.1754646 0.3372692       0.95
```

- Método de Wilson

```
epitools::binom.wilson(x,n,conf.level=0.95)
```

```
##    x   n proportion     lower     upper conf.level
## 1 30 120       0.25 0.1810982 0.3344114       0.95
```

- Aproximación normal fórmula de Laplace

```
epitools::binom.approx(x,n,conf.level=0.95)
```

```
##    x   n proportion     lower     upper conf.level
## 1 30 120       0.25 0.1725256 0.3274744       0.95
```

En este caso la podemos reproducir el cáculo de forma sencilla con funciones básicas de R

```
alpha=1-0.95
IC=c(p_muestral-qnorm(1-alpha/2)*sqrt(p_muestral*(1-p_muestral)/n),
     p_muestral+qnorm(1-alpha/2)*sqrt(p_muestral*(1-p_muestral)/n))
IC
```

```
## [1] 0.1725256 0.3274744
```