

Ejemplo de viñeta. Determinar el calculo de tamaño de muestra utilizando distintas afijaciones, considerando costes y sin considerar costes

[illegible]

Preparamos los datos

La población está compuesta por 2 estratos generados sintéticamente a partir de dos distribuciones normales.

Primer estrato: pensionistas en CyL

```
N1<-585479
pen<-rnorm(N1, 750, 100)
```

Segundo estrato: asalariados en CyL

```
N2<-932992
ass<-rnorm(N2, 1500, 500)
```

```
datos<-cbind(c(pen, ass), c(rep("pensionista", N1), rep("asalariado", N2)))
N<-N1+N2
```

Ejercicio 1. Consideremos una muestra global de $n=800$ individuos y afijación uniforme

a) Reparto de la muestra

```
Nh<-c(N1, N2)
n<-800
strata.allocation(Nh=Nh, n=n, alloc="unif")
#> [1] 400 400
```

b) Suponiendo una función de coste

```
C<-12000
Cini<-5000
ch<-c(45, 20)
size<-strata.samplesize.cost(Nh=Nh, C=C, cini=Cini, ch=ch, alloc="unif")
paste("Tamaño de muestra", size)
#> [1] "Tamaño de muestra 215.384615384615"
nh.unif<-strata.allocation(Nh=Nh, n=size, alloc="unif")

paste(c("Estrato 1:", "Estrato 2"), nh.unif)
#> [1] "Estrato 1: 107.692307692308" "Estrato 2 107.692307692308"
paste("Coste:", Cini+sum(ch*nh.unif))
#> [1] "Coste: 12000"
```

Como los tamaños de muestra deben ser enteros redondeamos al entero inferior.

```
nh.unif<-floor(nh.unif)
paste(c("Estrato 1:", "Estrato 2"), nh.unif)
#> [1] "Estrato 1: 107" "Estrato 2 107"
paste("Coste:", Cini+sum(ch*nh.unif))
#> [1] "Coste: 11955"
```

Ejercicio 2. Afijación proporcional

a) Reparto de la muestra

```
strata.allocation(Nh=Nh, n=n, alloc="prop")
#> [1] 308.4571 491.5429
```

b) Con función de coste igual a la anterior

```
size<-strata.samplesize.cost(Nh=Nh, C=C, cini=Cini, ch=ch, alloc="prop")
paste("Tamaño de muestra", size)
#> [1] "Tamaño de muestra 236.173037187271"
nh.unif<-floor(strata.allocation(Nh=Nh, n=size, alloc="prop"))
paste(c("Estrato 1:", "Estrato 2"), nh.unif)
#> [1] "Estrato 1: 91" "Estrato 2 145"
paste("Coste:", Cini+sum(ch*nh.unif))
#> [1] "Coste: 11995"
```

Ejercicio 3. Afijación de mínima varianza

La afijación de Neyman depende de las cuasivarianzas de los estratos, por lo que se deben estimar.

a) Reparto de la muestra

Opción 1: usar las varianzas reales como estimadores (solución teórica)

```
var<-c(var(pen), var(ass))
strata.allocation(Nh=Nh, n=n, var=var, alloc="min")
#> [1] 89.15727 710.84273
```

Opción 2: tomar una muestra previa para estimar las cuasivarianzas de los estratos.

```
sample<-strata.sample(data=datos, n=c(20, 20))
var<-c(var(sample[which(sample[,2]=="asalariado"),1]), var(sample[which(sample[,2]=="pensionista"),1]))
strata.allocation(Nh=Nh, n=n, var=var, alloc="min")
#> [1] 612.9519 187.0481
```

Opción 3: estimación más conservadora. Suponemos cuasivarianza máxima en todos los estratos = $\frac{N_h}{N_h-1}p(1-p)$ con $p=0.5$

```
var<-c(Nh/(Nh-1)*0.5*(1-0.5))
strata.allocation(Nh=Nh, n=n, var=var, alloc="min")
#> [1] 308.4572 491.5428
```

Si no se fija la varianza en la función se mostrará un warning y la varianza declarada en la función será la máxima para cada estrato.

```
strata.allocation(Nh=Nh, n=n, alloc="min")
#> Warning in strata.allocation(Nh = Nh, n = n, alloc = "min"):
#> Necessary var argument missing, will be set to worst case scenario value for each strata.
#> [1] 308.4572 491.5428
```

La estimación conservadora coincide con el cálculo de afijación proporcional.

b) Con función de coste igual a la anterior y solución teórica

En la afijación de mínima varianza optimizando con una función de costes es equivalente a utilizar la afijación óptima.

```
size<-strata.samplesize.cost(Nh=Nh, var=var, C=C, cini=Cini, ch=ch, alloc="optim")
paste("Tamaño de muestra", size)
#> [1] "Tamaño de muestra 319.21044160007"
nh.optim<-floor(strata.allocation(Nh=Nh, n=size, var=var, alloc="optim", C=C, cini=Cini, ch=ch))
paste(c("Estrato 1:", "Estrato 2"), nh.optim)
#> [1] "Estrato 1: 24" "Estrato 2 294"
paste("Coste:", Cini+sum(ch*nh.optim))
#> [1] "Coste: 11960"
```