

Entrega Final - Estadística Robusta 2024

Ignacio Spiousas

Ejercicio 4

Probar que tanto s como la **MAD** cumplen estas dos propiedades:

1. **Propiedad de ser invariante por traslaciones:** Sea $a \in \mathbb{R}$ e $y_i = x_i + a$ para todo $1 \leq i \leq n$. Luego, si $\mathbf{y} = (y_1, \dots, y_n)$ entonces:

$$\hat{\sigma}(\mathbf{y}) = \hat{\sigma}(\mathbf{x})$$

2. **Propiedad de ser equivariante por cambios de escala:** Sea $a \in \mathbb{R}$ e $y_i = ax_i$ para todo $1 \leq i \leq n$. Luego, si $\mathbf{y} = (y_1, \dots, y_n)$ entonces:

$$\hat{\sigma}(\mathbf{y}) = |a|\hat{\sigma}(\mathbf{x})$$

Resolución para el s

Empecemos por lo más fácil: s . Dada $\mathbf{x} = (x_1, \dots, x_n)$ una secuencia de \mathbf{n} variables aleatorias, podemos definir a $s(\mathbf{x})$ como:

$$s(\mathbf{x}) = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2},$$

donde $\bar{x} = 1/n \sum_{i=1}^n x_i$.

Invarianza por traslación de s

Recordemos que $y_i = x_i + a$ con $a \in R$ y que $\mathbf{y} = (y_1, \dots, y_n)$. Empecemos calculando \bar{y} que seguramente lo vayamos a necesitar:

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i = \frac{1}{n} \sum_{i=1}^n (x_i + a) = \frac{1}{n} \sum_{i=1}^n x_i + \frac{1}{n} \sum_{i=1}^n a = \bar{x} + \frac{1}{n} an = \bar{x} + a.$$

Y ahora a lo que nos importa, a probar que $s(\mathbf{y}) = s(\mathbf{x})$. Empecemos escribiendo $s(\mathbf{y})$:

$$s(\mathbf{y}) = \sqrt{\frac{1}{n-1}} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}.$$

Y reemplacemos cada y_i por $x_i + a$ y \bar{y} por $\bar{x} + a$ (que sale del cálculo auxiliar):

$$\begin{aligned} s(\mathbf{y}) &= \sqrt{\frac{1}{n-1}} \sqrt{\sum_{i=1}^n [x_i + a - (\bar{x} + a)]^2} = \sqrt{\frac{1}{n-1}} \sqrt{\sum_{i=1}^n (x_i + a - \bar{x} - a)^2} = \\ &= \sqrt{\frac{1}{n-1}} \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} = s(\mathbf{x}). \end{aligned}$$

Que es lo que queríamos demostrar.

Equivarianza por cambio de escala de s

Recordemos que $y_i = ax_i$ con $a \in R$ y que $\mathbf{y} = (y_1, \dots, y_n)$. Empecemos calculando \bar{y} que seguramente lo vayamos a necesitar:

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i = \frac{1}{n} \sum_{i=1}^n ax_i = a \frac{1}{n} \sum_{i=1}^n x_i = a\bar{x}.$$

En este caso queremos probar que $s(\mathbf{y}) = |a|s(\mathbf{x})$. Podemos retomar el $s(\mathbf{y})$ definido anteriormente pero esta vez reemplazando cada y_i por ax_i y \bar{y} por $a\bar{x}$ (que sale del cálculo auxiliar):

$$\begin{aligned} s(\mathbf{y}) &= \sqrt{\frac{1}{n-1}} \sqrt{\sum_{i=1}^n (ax_i - a\bar{x})^2} = \sqrt{\frac{1}{n-1}} \sqrt{\sum_{i=1}^n [a(x_i - \bar{x})]^2} = \sqrt{\frac{1}{n-1}} \sqrt{a^2 \sum_{i=1}^n (x_i - \bar{x})^2} = \\ &= \sqrt{a^2} \sqrt{\frac{1}{n-1}} \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} = |a| \sqrt{\frac{1}{n-1}} \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} = |a|s(\mathbf{x}). \end{aligned}$$

Que es lo que queríamos demostrar.

Resolución para la MAD

Empecemos refrescando las propiedades de la mediana que vimos en el **ejercicio 1** (para más detalle ver la sección Anexo:Ejercicio 1)¹.

$$\begin{aligned}\tilde{\mathbf{y}} &= \tilde{\mathbf{x}} + a & \text{si} & & y_i = x_i + a & & y & & \mathbf{y} = (y_1, \dots, y_n) \\ \tilde{\mathbf{y}} &= a\tilde{\mathbf{x}} & \text{si} & & y_i = ax_i & & y & & \mathbf{y} = (y_1, \dots, y_n),\end{aligned}$$

en ambos casos con $a \in \mathbb{R}$.

Ahora sí, definamos la *MAD*. Dada $\mathbf{x} = (x_1, \dots, x_n)$ una secuencia de \mathbf{n} variables aleatorias, podemos definir a $MAD(\mathbf{x})$ como:

$$MAD(\mathbf{x}) = Med_{1 \leq i \leq n} \{|x_i - \tilde{\mathbf{x}}|\}$$

Invarianza por traslación de la MAD

Recordemos que $y_i = x_i + a$ con $a \in \mathbb{R}$ y que $\mathbf{y} = (y_1, \dots, y_n)$. Queremos probar que $MAD(\mathbf{y}) = MAD(\mathbf{x})$. Empecemos escribiendo la $MAD(\mathbf{y})$:

$$MAD(\mathbf{y}) = Med_{1 \leq i \leq n} \{|y_i - \tilde{\mathbf{y}}|\}$$

Y reemplacemos cada y_i por $x_i + a$ y $\tilde{\mathbf{y}}$ por $\tilde{\mathbf{x}} + a$ (que sale de las propiedades del ejercicio 1.a):

$$\begin{aligned}MAD(\mathbf{y}) &= Med_{1 \leq i \leq n} \{|y_i - \tilde{\mathbf{y}}|\} = Med_{1 \leq i \leq n} \{|x_i + a - (\tilde{\mathbf{x}} + a)|\} = \\ &= Med_{1 \leq i \leq n} \{|x_i + a - \tilde{\mathbf{x}} - a|\} = Med_{1 \leq i \leq n} \{|x_i - \tilde{\mathbf{x}}|\} = MAD(\mathbf{x}).\end{aligned}$$

Que es lo que queríamos demostrar.

¹Donde $\tilde{\mathbf{x}}$ es la mediana de \mathbf{x} .

Equivarianza por cambio de escala de la MAD

Recordemos que $y_i = ax_i$ con $a \in R$ y que $\mathbf{y} = (y_1, \dots, y_n)$. En este caso queremos probar que $MAD(\mathbf{y}) = |a|MAD(\mathbf{x})$. Podemos retomar la $MAD(\mathbf{y})$ definido anteriormente pero esta vez reemplazando cada y_i por ax_i y $\tilde{\mathbf{y}}$ por $a\tilde{\mathbf{x}}$ (que sale de las propiedades del ejercicio 1.b):

$$\begin{aligned} MAD(\mathbf{y}) &= Med_{1 \leq i \leq n} \{|y_i - \tilde{y}|\} = Med_{1 \leq i \leq n} \{|ax_i - a\tilde{x}|\} = \\ &= Med_{1 \leq i \leq n} \{|a||x_i - \tilde{x}|\} = |a|MAD(\mathbf{x}). \end{aligned}$$

Que es lo que queríamos demostrar.

Ejercicio 8

Realizar una simulación en la que se pueda visualizar la eficiencia asintótica de la mediana en comparación con la media cuando los datos son normales. Para ello, construir un gráfico en el que se observe cómo aumentando el tamaño de muestra, la eficiencia de la mediana tiende a $2/\pi \approx 0.637$. A modo comparativo, realizar un gráfico para observar que si bien el cociente de las varianzas tiende a $2/\pi$, el cociente de las medias tiende a 1. ¿A qué se debe esto? ¿Se puede decir u observar algo más?

Resolución

Lo que vamos a hacer es simular varias realizaciones ($nrep$) de una muestra de tamaño n y calcular para cada una de ellas la media y la mediana. Luego, para cada n vamos a calcular la eficiencia como el cociente entre la varianza de las $nrep$ medias y la varianza de las $nrep$ medianas.

Para la simulación vamos a definir algunos parámetros:

- **min_n**: Mínimo tamaño de muestra a considerar.
- **max_n**: Máximo tamaño de muestra a considerar.
- **step**: Pasos para los tamaños de muestra que voy a explorar.
- **nrep**: Cantidad de repeticiones de cada simulación.

```
min_n <- 1
max_n <- 1e3
step <- 20
nrep <- 4000
```

Luego, vamos a generar una matriz x , de la cual luego utilizaremos una sección en particular de acuerdo al n^2 .

```
set.seed(1234)
x <- matrix(rnorm(nrep * max_n, 1, 1), nrep, max_n)
```

Para calcular la eficiencia vamos a echar mano a algunas herramientas del *{tidyverse}*. Para empezar utilizamos la función `expand_grid` que nos genera un tibble con todas las combinaciones de, en este caso, `rep` y `ns`. Donde `rep` varía de 1 a `nrep` y `ns` es un vector de `min_n` a `max_n` con paso `step`.

```
grilla <- expand_grid(ns = seq(min_n, max_n, step),
                     rep = 1:nrep)
slice(grilla, c(1:3, (nrow(grilla)-2):nrow(grilla)))
```

```
# A tibble: 6 x 2
```

	ns	rep
	<dbl>	<int>
1	1	1
2	1	2
3	1	3
4	981	3998
5	981	3999
6	981	4000

Podemos ver las primeras tres y las últimas tres filas del tibble `grilla`, donde puede verse que para cada valor de `ns` hay `nrep` filas con el índice `rep`.

Lo siguiente que hacemos es calcular para cada fila el la media y la mediana para los primeras `ns` columnas de la fila `rep` de `x` (es un poco un trabalenguas, pero juro que tiene sentido). El código nos puede ayudar a entender:

```
medias_medianas <- grilla |>
  rowwise() |>
  mutate(mean = mean (x[rep, 1:ns]),
         median = median(x[rep, 1:ns]))
```

En este código usamos la opción `rowwise()` que nos asegura que las operaciones del verbo `mutate` se están haciendo fila por fila.

²Para evitar que la compilación sea una pesadilla el documento no calcula x ni las medias y medianas sino que lo tengo precalculado y lo cargo para graficarlo. Entonces, las explicaciones de código son más bien para comprender la lógica de trabajo.

Finalmente, sólo nos queda calcular la eficiencia para cada valor de `ns`. Para esto vamos a utilizar dos intrucciones muy útiles de *{dplyr}*: `group_by()` y `summarise()`. La primera nos permite agrupar un tibble de acuerdo a los valores de una columna mientras que el verbo `summarise()` nos permite calcular operaciones sobre estas filas agrupadas. Entonces el código final queda de esta forma:

```
eficiencia_mediana <- medias_medianas |>
  group_by(ns) |>
  summarise(eficiencia = var(mean) / var(median))
```

Una de las cosas que nos permite hacer *{tidyverse}* es encadenar todo el código en una sola línea usando el operador pipe `|>`:

```
eficiencia_mediana <- expand_grid(rep = 1:nrep,
                                ns = seq(min_n, max_n, step)) %>%
  rowwise() %>%
  mutate(mean = mean(x[rep, 1:ns]),
         median = median(x[rep, 1:ns])) %>%
  group_by(ns) %>%
  summarise(eficiencia = var(mean) / var(median))
```

Ahora que tenemos este código, resulta trivial agregar el cociente entre las medias (en lugar de las varianzas):

```
eficiencia_mediana <- expand_grid(rep = 1:nrep,
                                ns = seq(min_n, max_n, step)) %>%
  rowwise() %>%
  mutate(mean = mean(x[rep, 1:ns]),
         median = median(x[rep, 1:ns])) %>%
  group_by(ns) %>%
  summarise(eficiencia = var(mean) / var(median),
         coc_medias = mean(mean) / mean(median))
```

En la figura [Figura 1](#) puede verse la representación gráfica de la eficiencia y el cociente de medias en función del tamaño de la muestra (n). En la misma podemos ver que efectivamente la eficiencia de la mediana converge a $2/\pi = 0.637$ mientras que el cociente entre las medias es 1 para todo n . Esto último se debe a que tanto la media como la mediana son estimadores insesgados de μ entonces, sin importar el n la esperanza de ambos (en este caso simulada como la media de `nrep` repeticiones) es μ .

Como última observación quiero retomar algo que tal vez haya pasado desapercibido. Cuando generamos `x` no tomamos muestras de una $\mathcal{N}(0, 1)$ sino de una $\mathcal{N}(1, 1)$ (hicimos `rnorm(nrep * max_n, 1, 1)`) ¿Por qué? Bueno, como estamos realizando una simulación numérica, tenemos

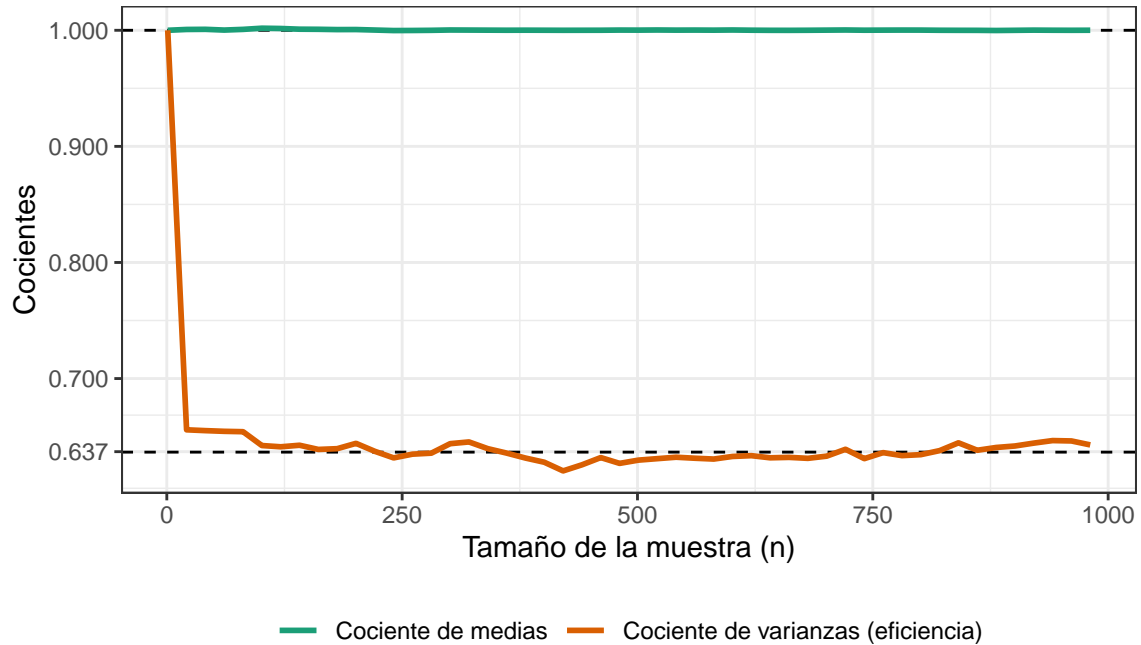


Figura 1: Eficiencia de la mediana y cociente de medias.

que tener cuidado de no hacer cocientes con números muy chicos y, sobre todo cuando n es grande, la media y la mediana son muy cercanas a cero. Esto lo que genera es que aparezcan inestabilidades numéricas. Repitamos la misma simulación que antes pero muestreando de una $\mathcal{N}(0, 1)$ y veamos qué pasa.

En la figura Figura 2 vemos que, efectivamente, el cociente entre las medidas tiene puntos raros debido a inestabilidades numéricas. También se observa que, para muestras más grandes donde es más esperable que la media y la mediana sean más parecidas hay menos puntos inestables. Sin embargo, como puede verse en el insert que magnifica la zona de n s 250 a 1000, el cociente de las medias está lejos de estabilizarse en 1.

Ejercicio 15

- Verificar que el MLE de escala para la distribución t de Student con ν grados de libertad es equivalente a considerar:

$$\rho(t) = \frac{t^2}{t^2 + \nu} \quad y \quad \delta = \frac{1}{\nu + 1}$$

- Probar que la función ρ del item anterior es una ρ -función.

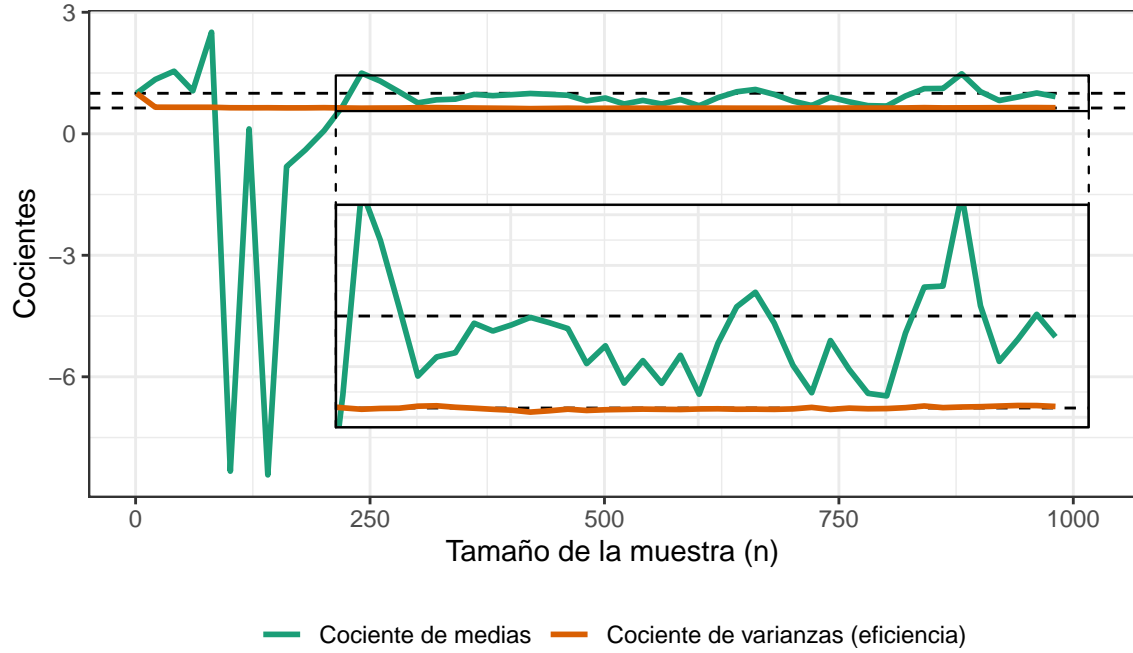


Figura 2: Eficiencia de la mediana y cociente de medias muestrando una $N(0,1)$.

- c. Mostrar que si se usa como $\rho(t) = I(|t| > c)$, con $c > 0$, y $\delta = 1/2$, se tiene que $\hat{\sigma} = \text{Med}(|x|)/c$.

Resolución

Inciso a

Recordemos primero el modelo de escala:

$$x_i = \sigma u_i.$$

Lo que queremos es minimizar la verosimilitud de este modelo que tiene función de densidad:

$$f_x = \frac{1}{\sigma} f_0 \left(\frac{x_i}{\sigma} \right),$$

donde $f_0()$ es una t de Student de ν grados de libertad.

La función de verosimilitud la podemos escribir como:

$$L(\sigma) = \prod_{i=1}^n f_x = \prod_{i=1}^n \frac{1}{\sigma} f_0\left(\frac{x_i}{\sigma}\right) = \frac{1}{\sigma^n} \prod_{i=1}^n f_0\left(\frac{x_i}{\sigma}\right),$$

que, aplicando logaritmo me queda:

$$l(\sigma) = \log[l(\sigma)] = -n \log(\sigma) + \sum_{i=1}^n \log\left(f_0\left(\frac{x_i}{\sigma}\right)\right).$$

Ahora para minimizar lo que tengo que hacer es derivar e igualar a 0:

$$\frac{dl(\sigma)}{d\sigma} = \frac{-n}{\sigma} + \sum_{i=1}^n \frac{f'_0\left(\frac{x_i}{\sigma}\right)}{f_0\left(\frac{x_i}{\sigma}\right)} \left(\frac{-x_i}{\sigma^2}\right) = 0.$$

Sacando el $1/\sigma$ de factor común y operando un poco llegamos a que:

$$\frac{1}{n} \sum_{i=1}^n \frac{f'_0\left(\frac{x_i}{\sigma}\right)}{f_0\left(\frac{x_i}{\sigma}\right)} \left(\frac{-x_i}{\sigma}\right) = 1,$$

donde el σ que cumple esta ecuación es efectivamente el $\hat{\sigma}$. Reemplazando $t = x_i/\sigma$ podemos reescribir la igualdad anterior como:

$$\frac{1}{n} \sum_{i=1}^n -\frac{f'_0(t)}{f_0(t)} t = 1.$$

Ahora escribamos explícitamente $f_0(t)$:

$$f_0(t) = H(\nu) \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}$$

donde $H(\nu)$ es una función de los grados de libertad y no de t . Escribamos también su derivada:

$$f'_0(t) = \frac{df_0(t)}{dt} = H(\nu) \left(-\frac{\nu+1}{2}\right) \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}-1} \frac{2t}{\nu},$$

que, luego de operar algebraicamente queda como:

$$f'_0(t) = -H(\nu) \frac{\nu+1}{\nu} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+3}{2}} t.$$

Ahora que tenemos estas expresiones podemos tratar de simplificar un poco $f'_0(t)/f_0(t)$:

$$\frac{f'_0(t)}{f_0(t)} = \frac{-H(\nu) \frac{\nu+1}{\nu} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+3}{2}} t}{H(\nu) \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}} = -\frac{(\nu+1)t}{\nu} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+3}{2} + \frac{\nu+1}{2}},$$

donde, haciendo un pequeño cálculo auxiliar podemos ver que:

$$-\frac{\nu+3}{2} + \frac{\nu+1}{2} = \frac{-\nu-3+\nu+1}{2} = \frac{-2}{2} = -1,$$

Entonces:

$$\frac{f'_0(t)}{f_0(t)} = -\frac{(\nu+1)t}{\nu} \left(1 + \frac{t^2}{\nu}\right)^{-1} = -\frac{(\nu+1)}{\nu} \left(\frac{\nu t}{\nu + t^2}\right) = -(\nu+1) \left(\frac{t}{\nu + t^2}\right),$$

entonces el sumando de la suma nos queda:

$$-\frac{f'_0(t)}{f_0(t)} t = (\nu+1) \left(\frac{t^2}{\nu + t^2}\right).$$

Entonces el estimador MLE $\hat{\sigma}$ será el que cumpla:

$$\frac{1}{n} \sum_{i=1}^n (\nu+1) \left(\frac{t^2}{\nu + t^2}\right) = \frac{\nu+1}{n} \sum_{i=1}^n \left(\frac{t^2}{\nu + t^2}\right) = 1,$$

recordando que $t = x_i/\sigma$. Entonces, lo siguiente que vamos a hacer es reescribirlo de forma tal de poder identificar $\rho()$ y δ :

$$\frac{1}{n} \sum_{i=1}^n \left(\frac{t^2}{\nu + t^2}\right) = \frac{1}{\nu+1},$$

o sea:

$$\frac{1}{n} \sum_{i=1}^n \rho\left(\frac{x_i}{\hat{\sigma}}\right) = \delta,$$

donde puede verse claramente que:

$$\rho(t) = \frac{t^2}{\nu + t^2},$$

y,

$$\delta = \frac{1}{\nu + 1},$$

tal como se quería demostrar.

Inciso b

Para ser $t^2/(\nu + t^2)$ un ρ -función, esta debe cumplir:

1. Ser una función no decreciente en $|t|$ (en particular, es una función par).
2. Que se cumpla que $\rho(0) = 0$.
3. Ser creciente (monótona) para $t > 0$ tal que $\rho(x) < \lim_{t \rightarrow +\infty} \rho(t) = \rho(\infty)$.
4. Que si ρ es acotada, se supone $\rho(\infty) = 1$.

Lo primero que vamos a verificar es que el $\rho(t)$ propuesto sea par:

$$\rho(-t) = \frac{(-t)^2}{\nu + (-t)^2} = \frac{t^2}{\nu + t^2} = \rho(t),$$

es decir, es una función par. Eso lo podemos ver también gráficamente. En la figura [Figura 3](#) podemos ver a $\rho(t)$ en función de t para $\nu = 1$.

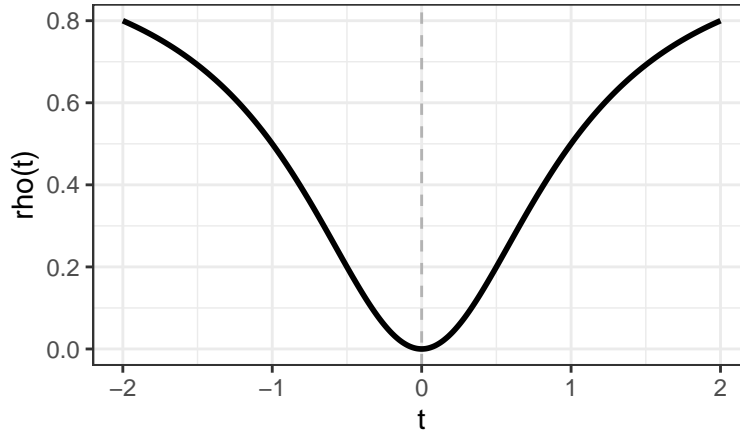


Figura 3: Función $\rho(t)$ en función de t para 10 grados de libertad.

Ahora una fácil $\rho(0) = 0^2/(\nu + 0^2) = 0/\nu = 0$. Algo que también puede verse en la figura [Figura 3](#).

Sigamos con la monotonía para $x > 0$. Dados $0 < t_a < t_b$ queremos demostrar que $\rho(t_a) < \rho(t_b)$. Primero evaluemos $\rho(t)$ en ambos puntos:

$$\rho(t_a) = \frac{t_a^2}{\nu + t_a^2} \quad y \quad \rho(t_b) = \frac{t_b^2}{\nu + t_b^2},$$

que podemos reescribir como:

$$\rho(t_a) = \frac{1}{\nu/t_a^2 + 1} \quad y \quad \rho(t_b) = \frac{1}{\nu/t_b^2 + 1}.$$

Como t_a y t_b son mayores que cero demostrar que $\rho(t_a) < \rho(t_b)$ es equivalente a demostrar que $1/\rho(t_a) > 1/\rho(t_b)$. Escribamos $1/\rho(t_a)$ y $1/\rho(t_b)$:

$$\frac{1}{\rho(t_a)} = \nu/t_a^2 + 1 \quad y \quad \frac{1}{\rho(t_b)} = \nu/t_b^2 + 1.$$

De nuevo, demostrar que $1/\rho(t_a) > 1/\rho(t_b)$ es equivalente a demostrar que $1/\rho(t_a) - 1 > 1/\rho(t_b) - 1$. Operemos de nuevo:

$$\frac{1}{\rho(t_a)} - 1 = \nu/t_a^2 \quad y \quad \frac{1}{\rho(t_b)} - 1 = \nu/t_b^2.$$

Como ν es una constante positiva y $t_a < t_b$ entonces $\nu/t_a > \nu/t_b$ y $1/\rho(t_a) - 1 > 1/\rho(t_b) - 1$, que es lo que queríamos demostrar. Esto también lo podemos ver graficamente. En la figura Figura 4 se observa que, para valores de $t > 0$, la $\rho(t)$ propuesta es monótona creciente.

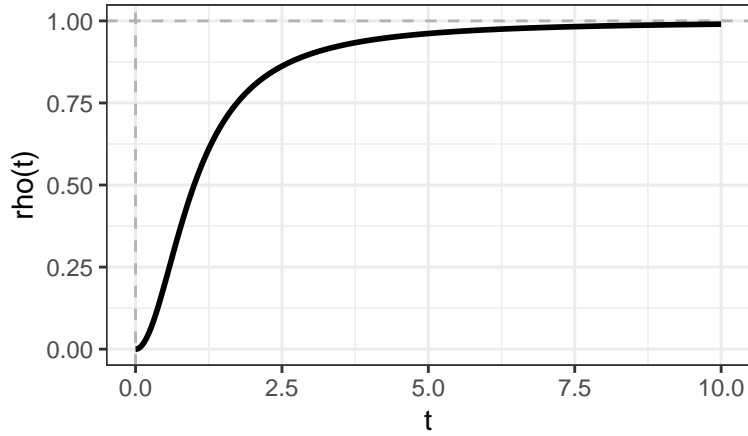


Figura 4: Función $\rho(t)$ en función de t para 10 grados de libertad solo para valores de $t > 0$.

Lo último que nos queda por probar es que $\rho(t)$ es acotada ($\rho(\infty) = 1$), algo que se ve fuertemente sugerido en la figura Figura 4. Para esto vamos a resolver el límite cuando $t \rightarrow \infty$ de $\rho(t)$:

$$\lim_{t \rightarrow \infty} \rho(t) = \lim_{t \rightarrow \infty} \frac{1}{\nu/t^2 + 1} = 1,$$

ya que ν/t^2 tiende a 0 cuando $t \rightarrow \infty$.

Finalmente, hemos demostrado que la $\rho(t)$ propuesta $t^2/(\nu + t^2)$ es una ρ -función.

Inciso c

Se nos pide mostrar que si se usa como $\rho(t) = I(|t| > c)$, con $c > 0$, y $\delta = 1/2$, se tiene que $\hat{\sigma} = \text{Med}(|x|)/c$. O sea, en este caso el estimador $\hat{\sigma}$ sería el que cumple la siguiente ecuación:

$$\frac{1}{n} \sum_{i=1}^n \rho\left(\frac{x}{\hat{\sigma}}\right) = \delta \quad \text{o sea} \quad \frac{1}{n} \sum_{i=1}^n I(|t| > c) = \frac{1}{2}.$$

Como por LGN la media tiene a la esperanza, y considerando a $I(|t| > c)$ como una variable aleatoria, podemos decir que:

$$\sum_{i=1}^n I(|t| > c) \xrightarrow{\mathcal{P}} \mathbb{E}[I(|t| > c)].$$

Entonces, como $\mathbb{E}[I(|t| > c)] = 1/2$ podemos decir que $P(|t| > c) = 1/2$. A partir de esto podemos despejar $\hat{\sigma}$:

$$P(|t| > c) = P\left(\left|\frac{x}{\hat{\sigma}}\right| > c\right) = P(|x| > c\hat{\sigma}) = \frac{1}{2},$$

ya que $\hat{\sigma}$ es siempre mayor que cero. Que, por la definición de mediana podemos reescribir como:

$$c\hat{\sigma} = \text{Med}(|x|)$$

O sea, que $\hat{\sigma} = \text{Med}(|x|)/c$ tal como queríamos demostrar.

Anexo: Ejercicio 1

Vamos a probar que la mediana es un estimador equivariante por transformaciones afines.

Primero definamos la mediana de \mathbf{x} ($\mathbf{x} = x_1, \dots, x_n$):

$$\tilde{\mathbf{x}} = Med_{1 \leq i \leq n} \{x_i\} = \begin{cases} x_{(\frac{n+1}{2})} & \text{si } n \text{ es impar} \\ \frac{x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}}{2} & \text{si } n \text{ es par} \end{cases}$$

Ejercicio 1.a

Probar que la mediana es un estimador equivariante por traslaciones.

Para evaluar una traslación solidaria, definamos una nueva variable y_i tal que $y_i = x_i + a$. Veamos que pasa con los estadísticos de orden de \mathbf{x} e \mathbf{y} :

$$\begin{aligned} x_{(1)} &\leq x_{(2)} \leq \dots \leq x_{(n)} \\ y_{(1)} &\leq y_{(2)} \leq \dots \leq y_{(n)} \end{aligned}$$

Es decir, los estadísticos de orden de \mathbf{x} e \mathbf{y} son iguales. Esto significa que si ordenamos de mayor a menor todas las x y todas las y , al sumarle a a la i -ésima \mathbf{x} obtenemos como resultado al i -ésima \mathbf{y} .

n impar

Ahora calculemos la mediana de \mathbf{y} para el caso en el que n es impar.

$$\tilde{\mathbf{y}} = y_{(\frac{n+1}{2})} = x_{(\frac{n+1}{2})} + a = \tilde{\mathbf{x}} + a$$

n par

Y para el caso en el que n es par.

$$\tilde{\mathbf{y}} = \frac{y_{(\frac{n}{2})} + y_{(\frac{n}{2}+1)}}{2} = \frac{x_{(\frac{n}{2})} + a + x_{(\frac{n}{2}+1)} + a}{2} = \frac{x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}}{2} + \frac{2a}{2} = \tilde{\mathbf{x}} + a$$

Finalmente, demostramos que, tanto cuando n es par como impar, la mediana es un estimador equivariante por traslaciones.

Ejercicio 1.b

Probar que la mediana es un estimador equivariante por cambios de escala.

Para evaluar el cambio de escala, definimos una nueva variable y_i tal que $y_i = ax_i$. En este caso la transformación de estadísticos de orden no es tan directa por lo que tendremos que considerar los casos en los que $a \geq 0$ y $a < 0$ separadamente.

$$a \geq 0$$

En este caso, los estadísticos de orden se mantienen iguales al caso de la traslación:

$$\begin{aligned}x_{(1)} &\leq x_{(2)} \leq \dots \leq x_{(n)} \\ y_{(2)} &\leq y_{(2)} \leq \dots \leq y_{(n)}\end{aligned}$$

$a \geq 0$ y n **impar**

Ahora calculemos la mediana de y para el caso en el que n es impar.

$$\tilde{y} = y_{(\frac{n+1}{2})} = ax_{(\frac{n+1}{2})} = a\tilde{x}$$

$a \geq 0$ y n **par**

Y para el caso en el que n es par.

$$\tilde{y} = \frac{y_{(\frac{n}{2})} + y_{(\frac{n}{2}+1)}}{2} = \frac{ax_{(\frac{n}{2})} + ax_{(\frac{n}{2}+1)}}{2} = a \frac{x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}}{2} = a\tilde{x}$$

$$a < 0$$

En este caso, los estadísticos de orden se invierten ya que, por ejemplo, el valor de \mathbf{x} más alto (x_n) pasa a ser el valor de \mathbf{y} más bajo (y_1):

$$\begin{aligned}x_{(1)} &\leq x_{(2)} \leq \dots \leq x_{(n-1)} \leq x_{(n)} \\ ax_{(1)} &\geq ax_{(2)} \geq \dots \geq ax_{(n-1)} \geq ax_{(n)} \\ y_{(1)} &\geq y_{(2)} \geq \dots \geq y_{(n-1)} \geq y_{(n)}\end{aligned}$$

Entonces, el estadístico de orden k de la nueva variable \mathbf{y} puede definirse como $y_{(k)} = ax_{(n-k+1)}$. Con esto en mente vamos a continuar con las demostraciones para n impar y par.

$a < 0$ y n **impar**

Ahora calculemos la mediana de y para el caso en el que n es impar.

$$\tilde{y} = y_{(\frac{n+1}{2})} = ax_{(\frac{n+1}{2})} = a\tilde{x}$$

O sea, $k = \frac{n+1}{2}$, y:

$$y_{(\frac{n+1}{2})} = y_{(k)} = ax_{(n-k+1)} = ax_{(n-\frac{n+1}{2}+1)} = ax_{(n-\frac{n}{2}-\frac{1}{2}+1)} = ax_{(\frac{n}{2}+\frac{1}{2})} = ax_{(\frac{n+1}{2})}$$

Entonces:

$$\tilde{y} = y_{(\frac{n+1}{2})} = ax_{(\frac{n+1}{2})} = a\tilde{x}$$

$a < 0$ y n **par**

Veamos como se modifican los estadísticos de orden en este caso. Primero, consideremos el caso en el que $k = \frac{n}{2}$:

$$y_{(\frac{n}{2})} = y_{(k)} = ax_{(n-k+1)} = ax_{(n-\frac{n}{2}+1)} = ax_{(\frac{n}{2}+1)}$$

Y segundo, consideremos el caso en el que $k = \frac{n}{2} + 1$:

$$y_{(\frac{n}{2}+1)} = y_{(k)} = ax_{(n-k+1)} = ax_{(n-(\frac{n}{2}+1)+1)} = ax_{(n-\frac{n}{2}-1+1)} = ax_{(\frac{n}{2})}$$

Entonces, veamos como se modifica la mediana de x por un cambio de escala cuando $a < 0$ y n es par:

$$\tilde{y} = \frac{y_{(\frac{n}{2})} + y_{(\frac{n}{2}+1)}}{2} = \frac{ax_{(\frac{n}{2}+1)} + ax_{(\frac{n}{2})}}{2} = a \frac{x_{(\frac{n}{2}+1)} + x_{(\frac{n}{2})}}{2} = a\tilde{x}$$

Finalmente, demostramos que, tanto cuando n es par como impar, la mediana es un estimador equivariante por cambios de escala.

Conclusión

Hemos demostrado que la mediana es un estimador equivariante por transformaciones afines, es decir, que es equivariante por traslaciones y cambios de escala.