En cierto experimento oligorradiactivo se registra el ángulo X (en grados) que cada partícula emitida sigue respecto a la dirección referencial. Se obtiene la muestra de tamaño 30 siguiente:

```
## c(43.26, 37.78, 43.56, 34.45, 44.95, 43.4, 38.34, 40.94, 45.03,
## 36.63, 40.92, 39.27, 37.46, 39.98, 35.05, 42.03, 41.51, 42.43,
## 44.78, 39.51, 39.35, 42.89, 49.09, 43.62, 45.6, 30.01, 40.87,
## 38.45, 41.01, 40.74)
```

La densidad de X es

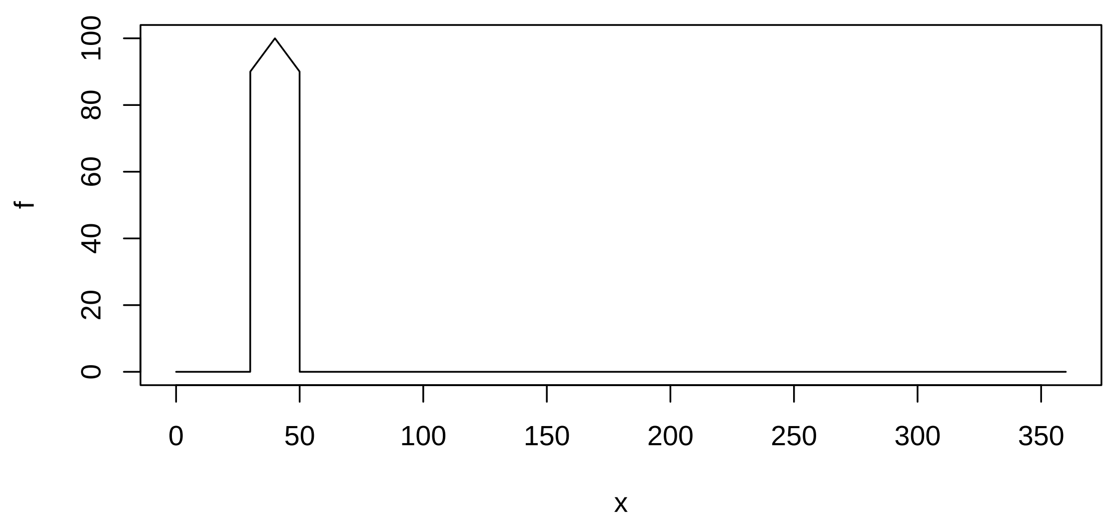
$$f(x) = a \cdot (100 - |x - \theta|) \cdot I(\theta - 10 < x < \theta + 10)$$

El parámetro θ toma valores entre 20 y 60.

1. (0'8 puntos) **Calcula la constante a.**

   En primer lugar representamos gráficamente la densidad, teniendo en cuenta que θ es un parámetro de ubicación y que a es una constante de proporcionalidad:

```r
a <- 1  # por ejemplo
θ <- 40 # por ejemplo
f <- function (x) a * (100-abs(x-θ)) * (θ-10<x) * (x<θ+10)
plot (f, 0, 360, n=10001)
```



Tiene forma de triángulo isósceles sobre rectángulo, ambos con base $(\theta+10)-(\theta-10)=20$. La altura del rectángulo es $f(\theta-10)=a\cdot\big(100-|(\theta-10)-\theta|\big)=f(\theta+10)=a\cdot\big(100-|(\theta+10)-\theta|\big)=90a$. La altura del triángulo es $a\cdot\big(100-|\theta-\theta|\big)-90a=100a-90a=10a$. El área bajo la densidad es pues $20\cdot 90a+(20\cdot 10a)/2=1900a$, luego $1=\int f=1900a \implies a=1/1900$. Podemos comprobarlo numéricamente, manteniendo $a=1$ y, por ejemplo, para $\theta=40$:

```
integrate(f, 0, 360)

## 1900 with absolute error < 0.013
```

o bien comprobándolo para muchas θ:

```
summary (sapply (seq (20, 60, .01),
                 function (θ)
                     integrate (function (x) (100-abs(x-θ)),
                                θ-10, θ+10) $ value))

##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1900    1900    1900    1900    1900    1900
```

En vez de geométricamente, puede resolverse analíticamente como sigue:

$$
\begin{aligned}
1=\int f &= \int_{\theta-10}^{\theta+10} a\cdot(100-|x-\theta|)\cdot dx \\
\left\langle \begin{array}{c} t=x-\theta \\ dt=dx \end{array} \right\rangle &= \int_{-10}^{10} a\cdot(100-|t|)\cdot dt \\
\langle \text{simetría} \rangle &= 2\cdot\int_0^{10} a\cdot(100-t)\cdot dt \\
&= 2\cdot a\cdot\Big[100\cdot t-t^2/2\Big]_{t=0}^{10} \\
&= 2\cdot a\cdot 950=1900\cdot a \\
&\Rightarrow a=1/1900
\end{aligned}
$$

2. (0’8 puntos) **Comprueba si $(X_{(1)},\bar{X},X_{(n)})$ es un estadístico suficiente para θ.**

   La verosimilitud es

$$
\begin{aligned}
\mathcal{L} &= \prod_{i=1}^{n} a\cdot(100-|x_i-\theta|)\cdot I(\theta-10<x_i<\theta+10) \\
&= a^n\cdot\prod_i\Big(100-|x_i-\theta|\Big)\cdot I(\theta-10<x_{(1)})\cdot I(x_{(n)}<\theta+10)
\end{aligned}
$$

No parece fácil expresarla en función también de $\bar{x}$, por lo que buscaremos un contraejemplo. Sean las muestras $\vec{x} = (39, 40, 40, 41)$ y $\vec{x} = (39, 39, 41, 41)$. Por un lado, $(x_{(1)}, \bar{x}, x_{(n)}) = (39, 40, 41) = (y_{(1)}, \bar{y}, y_{(n)})$. Sin embargo,

$$
\begin{aligned}
\frac{\mathcal{L}(\vec{x},\theta)}{\mathcal{L}(\vec{y},\theta)} &= \frac{a^n \prod_i \Big(100 - |x_i - \theta|\Big) \cdot I(\theta - 10 < x_{(1)}) \cdot I(x_{(n)} < \theta + 10)}{a^n \prod_i \Big(100 - |y_i - \theta|\Big) \cdot I(\theta - 10 < y_{(1)}) \cdot I(y_{(n)} < \theta + 10)} \\
&= \frac{\prod_i \Big(100 - |x_i - \theta|\Big) \cdot I(\theta - 10 < 39) \cdot I(41 < \theta + 10)}{\prod_i \Big(100 - |y_i - \theta|\Big) \cdot I(\theta - 10 < 39) \cdot I(41 < \theta + 10)} \\
&= \frac{\prod_i \Big(100 - |x_i - \theta|\Big) \cdot I(31 < \theta < 49)}{\prod_i \Big(100 - |y_i - \theta|\Big) \cdot I(31 < \theta < 49)} \\
&= \left\langle \frac{I(31 < \theta < 49)}{I(31 < \theta < 49)} \text{ independiente de } \theta \right\rangle = \frac{\prod_i \Big(100 - |x_i - \theta|\Big)}{\prod_i \Big(100 - |y_i - \theta|\Big)} \\
&= \frac{\Big(100 - |39 - \theta|\Big) \cdot \Big(100 - |40 - \theta|\Big) \cdot \Big(100 - |40 - \theta|\Big) \cdot \Big(100 - |41 - \theta|\Big)}{\Big(100 - |39 - \theta|\Big) \cdot \Big(100 - |39 - \theta|\Big) \cdot \Big(100 - |41 - \theta|\Big) \cdot \Big(100 - |41 - \theta|\Big)} \\
&= \frac{\Big(100 - |40 - \theta|\Big)^2}{\Big(100 - |39 - \theta|\Big) \cdot \Big(100 - |41 - \theta|\Big)}
\end{aligned}
$$

depende de θ pues, por ejemplo,

$$\frac{\mathcal{L}(\vec{x},40)}{\mathcal{L}(\vec{y},40)} = \frac{10.000}{9.801} = 1\text{'}0203... \neq 1\text{'}0001... = \frac{9.801}{9.800} = \frac{\mathcal{L}(\vec{x},41)}{\mathcal{L}(\vec{y},41)}$$

Otro enfoque sería coger muestras como $\vec{x} = (\theta - 5, \theta - \frac{\theta}{10}, \theta + \frac{\theta}{10}, \theta + 5)$ y $\vec{y} = (\theta - 5, \theta - 5, \theta + 5, \theta + 5)$, con lo que el cociente depende claramente de θ,

$$\frac{\mathcal{L}(\vec{x},\theta)}{\mathcal{L}(\vec{y},\theta)} = \frac{95 \cdot (100 - \frac{\theta}{10})^2 \cdot 95}{95^4}$$

teniendo en cuenta que $2 < \frac{\theta}{10} < 6$ y se trata de muestras válidas.

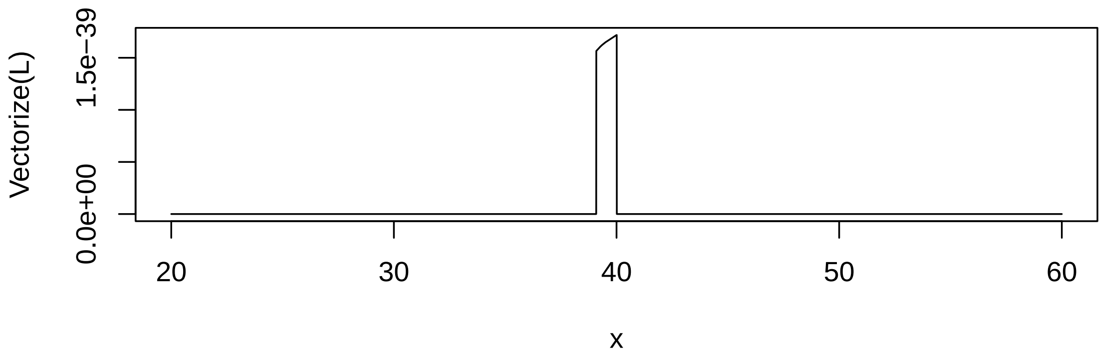
3. (0'8 puntos) **Halla la estimación máximo-verosímil de θ.**

La verosimilitud es

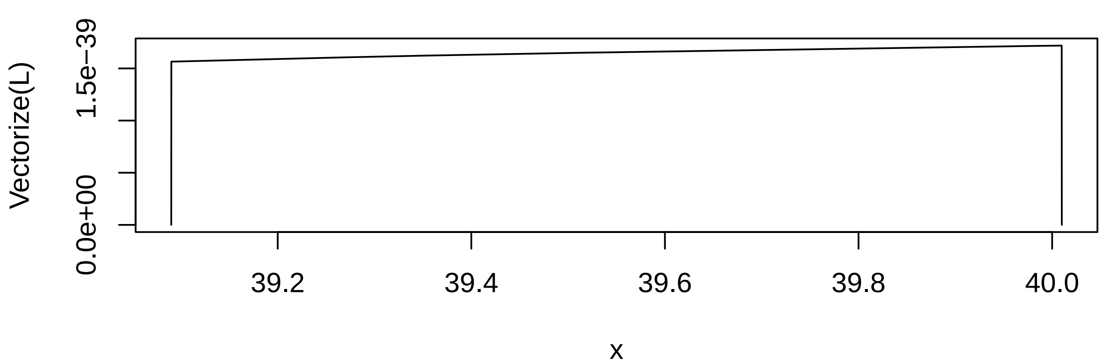
$$
\begin{aligned}
\mathcal{L}(\theta) &= a^n \cdot \prod_i \Big(100 - |x_i - \theta|\Big) \cdot I(\theta - 10 < x_{(1)}) \cdot I(x_{(n)} < \theta + 10) \\
&= a^n \cdot \prod_i \Big(100 - |\theta - x_i|\Big) \cdot I(x_{(n)} - 10 < \theta < x_{(1)} + 10)
\end{aligned}
$$

Como no es una función derivable y la obtención analítica del máximo parece difícil, usaremos maximización numérica:

```r
a <- 1/1900
L <- function (θ)
    a^n * prod(100-abs(θ-X)) * (max(X)-10<θ) * (θ<min(X)+10)
## representación gráfica para comprobar su comportamiento:
plot (Vectorize(L), 20, 60, n=10001)
```




```r
(emv <- optimize (L, c(20,60), maximum=TRUE)$maximum)

## [1] 59.99995

## estimación cerca del extremo, incorrecta; refinamos:
θmin <- max(X)-10
θmax <- min(X)+10
plot (Vectorize(L), θmin, θmax, n=10001)
```



```r
(emv <- optimize (L, c(θmin,θmax), maximum=TRUE))

## $maximum
## [1] 40.00994
##
## $objective
## [1] 1.718717e-39
```

Por tanto, la estimación máximo-verosímil de θ para la muestra del enunciado es $40'0099$ con un error menor que $1'23 \times 10^{-4}$:

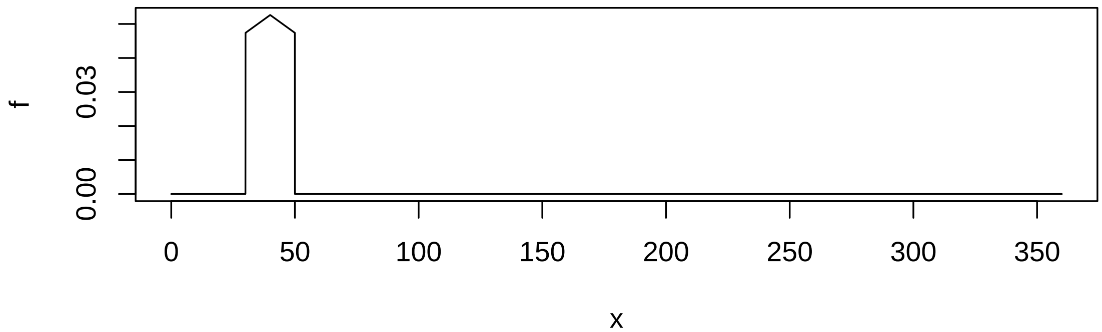
```r
## véase ?optimize
sqrt(.Machine$double.eps) * emv$maximum +
    .Machine$double.eps^0.25

## [1] 0.0001226665
```

En el caso de esta muestra concreta, se puede considerar que la verosimilitud es creciente (con otra muestra podría no serlo) por lo visto en su representación gráfica, lo que llevaría a que el máximo se alcanzaría en el mayor valor posible de θ, es decir, $x_{(1)} + 10 = 30'01 + 10 = 40'01$, el cual es un valor dentro del margen de precisión numérica comentado antes, $40'0099429 \pm 1'2266651 \times 10^{-4} = [40'0098202, 40'0100655]$.

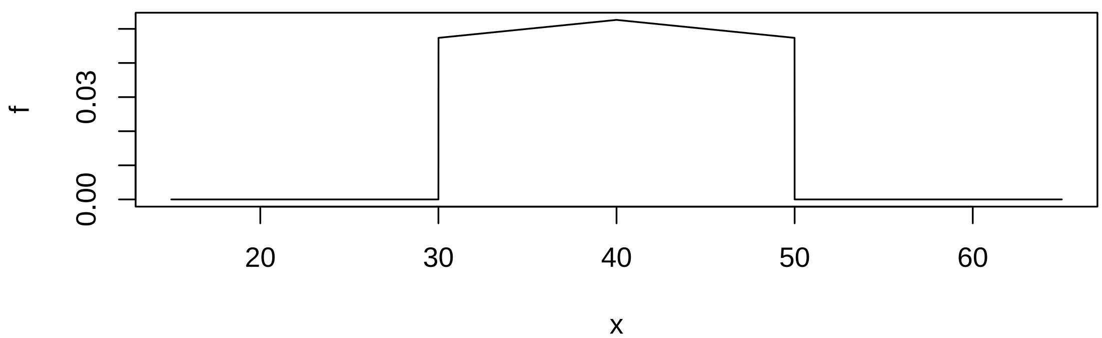
4. (0'8 puntos) **Obtén el estimador por el método de los momentos de θ y calcula su valor para la muestra.**

   Como se vio antes, la densidad es simétrica respecto a θ:

```r
a <- 1/1900
θ <- 40 # por ejemplo
f <- function (x) a * (100-abs(x-θ)) * (θ-10<x) * (x<θ+10)
plot (f, 0, 360, n=10001)
```




```r
plot (f, 15, 65, n=10001)
```



Las ilustraciones sugieren la simetría, que puede demostrarse considerando, dado $h > 0$, los puntos $x_1 = \theta - h$ y $x_2 = \theta + h$

$$
\begin{aligned}
f(x_1) &= a \cdot (100 - |x_1 - \theta|) \cdot I(\theta - 10 < x_1 < \theta + 10) \\
&= a \cdot (100 - |\theta - h - \theta|) \cdot I(\theta - 10 < \theta - h < \theta + 10) \\
&= a \cdot (100 - |-h|) \cdot I(-10 < -h < +10) \\
&= a \cdot (100 - h) \cdot I(-10 < h < +10)
\end{aligned}
$$

$$
\begin{aligned}
f(x_2) &= a \cdot (100 - |x_2 - \theta|) \cdot I(\theta - 10 < x_2 < \theta + 10) \\
&= a \cdot (100 - |\theta + h - \theta|) \cdot I(\theta - 10 < \theta + h < \theta + 10) \\
&= a \cdot (100 - |h|) \cdot I(-10 < h < +10) \\
&= a \cdot (100 - h) \cdot I(-10 < h < +10) \\
&= f(x_1)
\end{aligned}
$$

Por tanto,

$$
\begin{aligned}
f(\theta - h) = f(\theta + h) &\implies f \text{ es simétrica respecto a } \theta \\
&\implies E(X) = \theta \\
&\implies \hat{\theta}_{\mathrm{MM}} = \bar{X}
\end{aligned}
$$

Si no se quiere emplear la simetría,

$$
\begin{aligned}
E(X) &= \int_{-\infty}^{\infty} x \cdot f(x) \cdot dx \\
&= \int_{\theta-10}^{\theta+10} x \cdot a \cdot (100 - |x-\theta|) \cdot dx \\
&= \int_{\theta-10}^{\theta} x \cdot a \cdot (100 + x - \theta) \cdot dx + \int_{\theta}^{\theta+10} x \cdot a \cdot (100 - x + \theta) \cdot dx \\
&= a \cdot 100 \cdot \int_{\theta-10}^{\theta} x \cdot dx + a \cdot \int_{\theta-10}^{\theta} x^2 \cdot dx - a \cdot \theta \cdot \int_{\theta-10}^{\theta} x \cdot dx \\
&+ a \cdot 100 \cdot \int_{\theta}^{\theta+10} x \cdot dx - a \cdot \int_{\theta}^{\theta+10} x^2 \cdot dx + a \cdot \theta \cdot \int_{\theta}^{\theta+10} x \cdot dx \\
&= a \cdot 100 \cdot \int_{\theta-10}^{\theta+10} x \cdot dx \\
&+ a \cdot \left(\int_{\theta-10}^{\theta} x^2 \cdot dx - \int_{\theta}^{\theta+10} x^2 \cdot dx \right) \\
&- a \cdot \theta \cdot \left(\int_{\theta-10}^{\theta} x \cdot dx - \int_{\theta}^{\theta+10} x \cdot dx \right) \\
&= a \cdot 100 \cdot \left[\frac{x^2}{2} \right]_{x=\theta-10}^{\theta+10} \\
&+ a \cdot \left(\left[\frac{x^3}{3} \right]_{\theta-10}^{\theta} - \left[\frac{x^3}{3} \right]_{\theta}^{\theta+10} \right) \\
&- a \cdot \theta \cdot \left(\left[\frac{x^2}{2} \right]_{x=\theta-10}^{\theta} - \left[\frac{x^2}{2} \right]_{x=\theta}^{\theta+10} \right) \\
&= a \cdot 100 \cdot \left[\frac{(\theta+10)^2}{2} - \frac{(\theta-10)^2}{2} \right] \\
&+ a \cdot \left(\left[\frac{\theta^3}{3} - \frac{(\theta-10)^3}{3} \right] - \left[\frac{(\theta+10)^3}{3} - \frac{\theta^3}{3} \right] \right) \\
&- a \cdot \theta \cdot \left(\left[\frac{\theta^2}{2} - \frac{(\theta-10)^2}{2} \right] - \left[\frac{(\theta+10)^2}{2} - \frac{\theta^2}{2} \right] \right) \\
&= a \cdot 2000 \cdot \theta \\
&+ a \cdot -200 \cdot \theta \\
&- a \cdot \theta \cdot -100 \\
&= a \cdot 1900 \cdot \theta = \theta \implies \hat{\theta}_{\mathrm{MM}} = \bar{X}
\end{aligned}
$$

Para la muestra del enunciado, la estimación por el método de los momentos es

```r
mean(X)
```

```
## [1] 40.76367
```