Ejercicios de Probabilidad y Simulación

Tema 2: Variables aleatorias y distribuciones

Juan Cantero Jimenez

3 de noviembre, 2021

1. *Sea X una variable aleatoria absolutamente continua con función de densidad:*

$$f(x) = \begin{cases} 0 \ si\ x < 0 \\ 2x \ si\ 0 \leq x < 1/2 \\ 6(x-1) \ si\ 1/2 \leq x \leq 1 \\ 0 \ si\ x > 1 \end{cases}$$

a Dibuja $f(x)$ y comprueba que realmente es una función de densidad de probabilidad

b Calcula la función de distribución de X y represéntala gráficamente.

c Calcula $P(\frac{3}{4} \leq X \leq 1)$ y $P(\frac{1}{3} \leq X \leq \frac{2}{3})$

a Para garantizar que $f(x)$ sea una función de densidad, se debe de integrar esta función, y comprobar que el area bajo la curva sea 1. Para comprobar que efectivamente esto se cumple se define la función en R y se integra entre 0 y 1 gracias a la función *integrate*().

```
dcustom <- function(z){
  fdensi <- function(x){
    if (x < 0){return(0)}
    if (0 <= x & x < 1/2){return(x*2)}
    if (1/2 <= x & x <= 1){return(6*(1-x))}
    if (x > 1){return(0)}
  }
  if (length(z) == 1){
    fdensi(z)
  }else{
    sapply(z,fdensi)

  }
}

integrate(dcustom,0,1)

## 1 with absolute error < 1.1e-14
```

b Para obtener la función de distribución de la variable X se debe de integrar esta, para esto se puede usar la función *integrate*() mencionada anteriormente o aprovecharse de que el área bajo la curva de la función de densidad se corresponde con dos triángulos rectángulos. Aprovechando esto la función de distribución se puede definir en R:

```
pcustom <-function(x){

fundistri_numeric_ <- function(x){
```

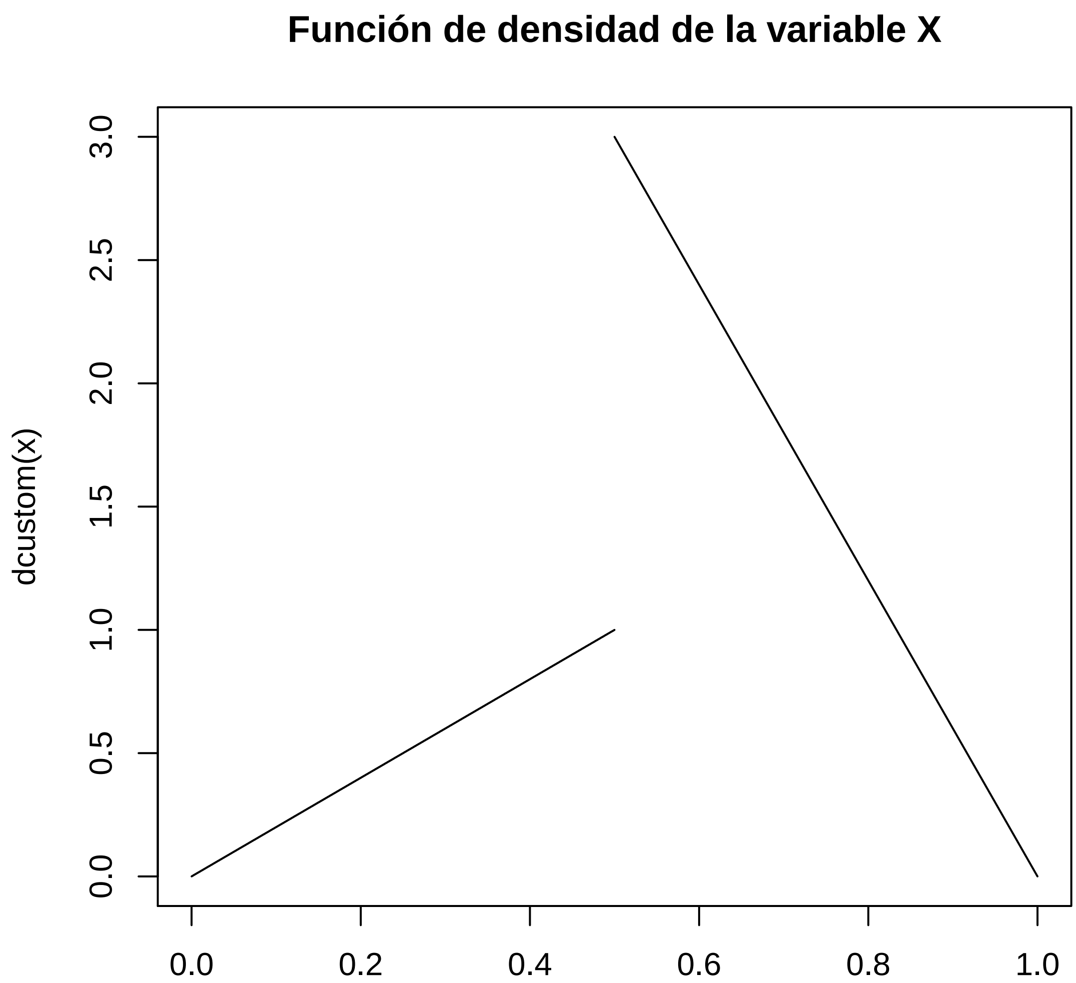

Figure 1: Representación gráfica de la función de densidad de la variable X

```
  if (x < 1/2){
    return((dcustom(x)*x)/2)
  }else{
    # [0,1/2)
    fi <- 1-((dcustom(1/2)*1/2)/2)
    # [1/2,x]
    fi2 <- ( ((dcustom(1/2)-dcustom(x))*(x-1/2))/2 + (dcustom(x)*(x-1/2) ))
    return(fi+fi2)
  }
}
  if(length(x) == 1){
    fundistri_numeric_(x)
  }else{
    sapply(x,fundistri_numeric_)
  }
```

```
}
```

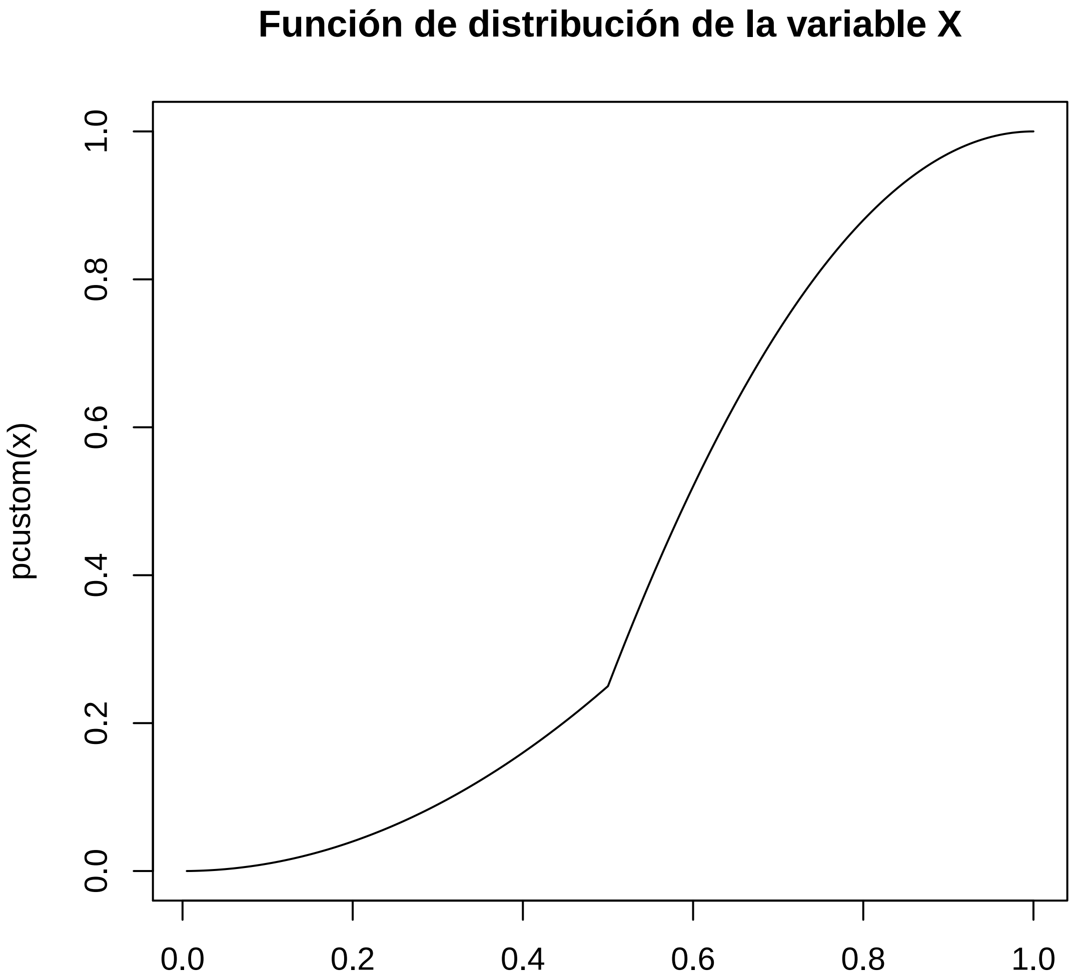

Figure 2: Representación gráfica de la función de distribución de la variable X

c Las probabilidades pedidas pueden ser halladas haciendo uso de la función de distribución definida anteriormente en R como $pcustom()$.

```
## P(3/4 <= X <= 1)
pcustom(1)-pcustom(3/4)

## [1] 0.1875

## P(1/3 <= X <= 2/3)
pcustom(2/3)-pcustom(1/3)

## [1] 0.5555556
```

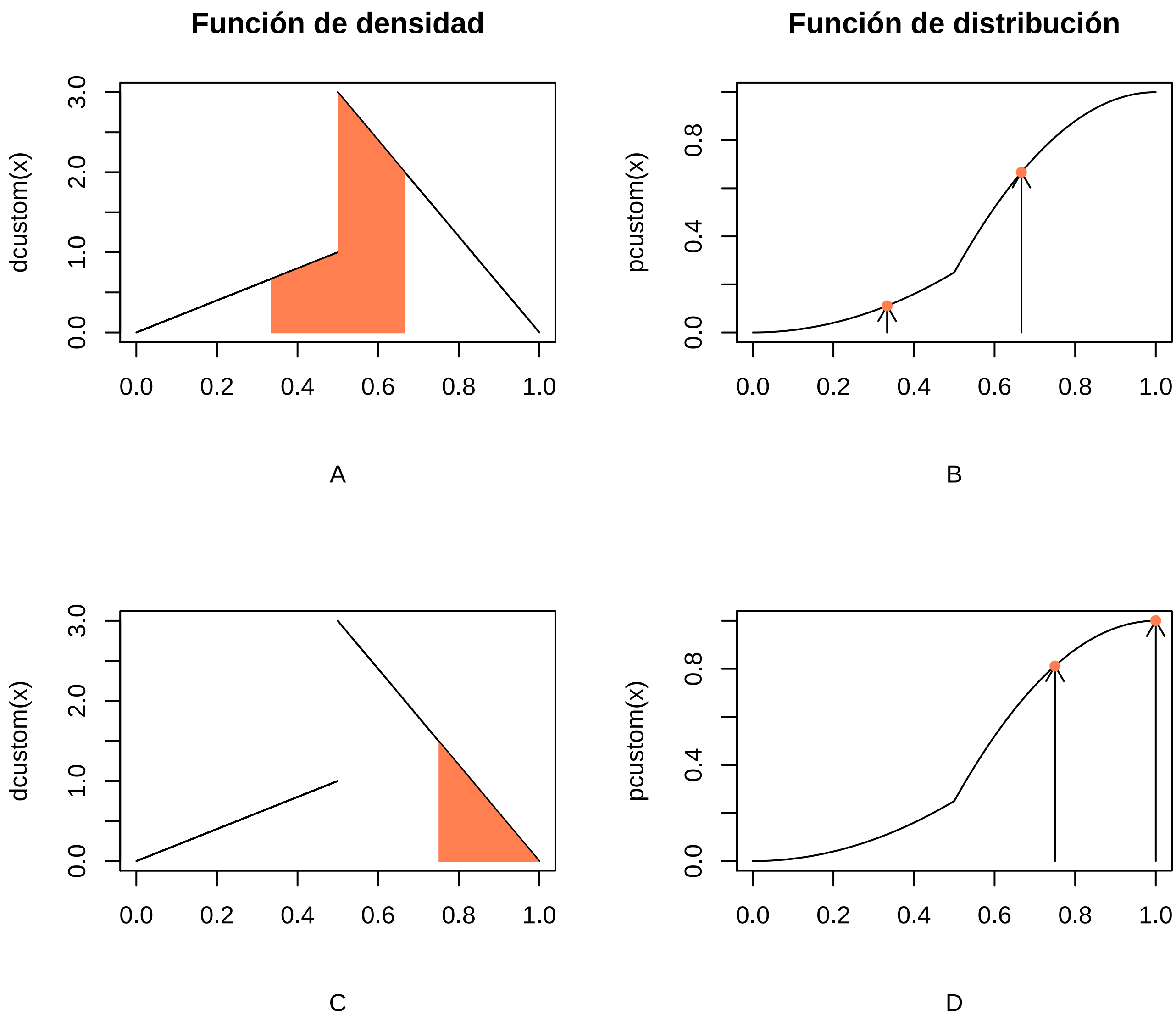

Figure 3: Las figuras A y B representan la probabilidad pedida de $P(\frac{3}{4} \leq X \leq 1)$ en la función de densidad y de distribución respectivamente. Las figuras C y D representan la probabilidad pedida de $P(\frac{1}{3} \leq X \leq \frac{2}{3})$ en la función de densidad y de distribución respectivamente

2. *En un estudio realizado en la ciudad Magdalena de Kino (México) para evaluar el Índice de Masa Corporal (IMC) en menores en edad escolar (de 8-11 años) se tomaron medidas a una muestra aleatoria de 474 menores, el 49.8 % de los cuales eran niñas. Según el estudio publicado por Chiquete et al (2014), el índice de masa corporal de esta población sigue una distribución aproximadamente normal de media 18.7 y desviación típica de 4.1. Se considera que un menor tiene sobre- peso si su IMC está por encima de 21.8 y tiene una desnutrición leve por debajo de 14.1. En el mismo estudio también se midió el peso (en kg) de los menores, cuya distribución es aproximadamente normal de media 35.1 kg y desviación típica 12.3 kg.*

- a *¿Cuál es el porcentaje de la población de menores que presenta sobrepeso? ¿cuál es el porcentaje de los que tendrían desnutrición leve?*
- b *¿Cuál es la proporción de menores en la población que tendría. normopeso?*
- c *Calcula los percentiles 35 y 65 para la distribución del Índice de Masa Corporal (IMC) en menores en edad escolar.*

d *Calcula el primer y el tercer cuartil del Índice de Masa Corporal (IMC) en menores en edad escolar. ¿Cuánto vale el rango intercuartılico?*

e *Visitamos una clase de primaria de esta ciudad, en la que hay 21 menores. Calcula la probabilidad de que como máximo tres menores de esa clase tengan un peso que exceda los 38 kg.*

Se pregunta por probabilidades asociadas a dos variables continuas que describen el Indice de masa corporal, IMC, y peso, p, de una población de menores en la ciudad de Magdalena de Kino de la cual se tomo una muestra aleatoria de tamaño 474. Ambas variables siguen una distribución normal de media 18.7 y desviación típica 4.1 para la variable IMC y media 35.1 kg y desviación tipica 12.3 kg para la variable p.Esto es equivalente a decir:

$$IMC \sim N(\mu = 18.7, \sigma = 4.1)$$

$$p \sim N(\mu = 35.1kg, \sigma = 12.3kg)$$

Para obtener las probabilidades pedidas se usará R. Primero se asignan los datos proporcionados en el ejercicio a variables definidas en el entorno de R.

```
n <- 474
mu_IMC <- 18.7
sigma_IMC <- 4.1
sobrepeso_IMC <- 21.8
desnutricion_IMC <- 14.1
mu_peso <- 35.1
sigma_peso <- 12.3
```

a $P(porcentaje\ menores\ con\ sobrepeso) = P(IMC \geq 21.8)$

```
1-pnorm(sobrepeso_IMC, mu_IMC, sigma_IMC)

## [1] 0.2247954
```

$P(porcentaje\ menores\ con\ desnutrición) = P(IMC \leq 14.1)$

```
pnorm(desnutricion_IMC, mu_IMC,sigma_IMC)

## [1] 0.1309416
```

b $P(porcentaje\ de\ menores\ con\ normopeso) = P(14.1 \leq IMC \leq 21.8)$

```
pnorm(sobrepeso_IMC,mu_IMC,sigma_IMC) - pnorm(desnutricion_IMC,mu_IMC,sigma_IMC)

## [1] 0.6442631
```

c P_{35}

```
qnorm(0.35,mu_IMC,sigma_IMC)

## [1] 17.12019
```

P_{65}

```r
qnorm(0.65,mu_IMC,sigma_IMC)
```

```
## [1] 20.27981
```

d Q_1

```r
qnorm(0.25,mu_IMC,sigma_IMC)
```

```
## [1] 15.93459
```

Q_3

```r
qnorm(0.75,mu_IMC,sigma_IMC)
```

```
## [1] 21.46541
```

$Q_3 - Q_1$

```r
qnorm(0.75,mu_IMC,sigma_IMC)-qnorm(0.25,mu_IMC,sigma_IMC)
```

```
## [1] 5.530816
```

e Primero se debe de calcular la probabilidad de que un individuo de la población pese mas de 38 kg, esto es equivalente a decir $P(p \geq 38\ kg)$

```r
p <- 1-pnorm(38, mu_peso, sigma_peso)
print(p)
```

```
## [1] 0.4068047
```

La probabilidad de que como máximo tres menores pesen mas de 38 kg en una clase de 21 individuos es una variable discreta, X, que puede ser modelizado por una distribución binomial

$$X \sim Bi(p = 0.406,\ n = 21)$$

$$P(X \leq 3)$$

```r
pbinom(3,21,p)
```

```
## [1] 0.009376732
```