PRACTICA 5: VARIABLES ALEATORIAS

.

Grado en biología sanitaria. Universidad de Alcalá, 2023/24.

1. Considera la variable binomial $X \sim B(7, 1/4)$. Para calcular probabilidades de X existen sendas funciones en R

   - La función `dbinom` permite calcular una probabilidad puntual, $P(X = 3)$ de calcula con

     ```
     dbinom(3, size = 7, prob = 1/4)
     ```

     ```
     [1] 0.1730347
     ```

     Observa que hay que especificar el número total de experimentos ($n = 7$) y la probabilidad de éxito ($prob = 1/4$) de la binomial con la que trabajas.

   - La función `pbinom` permite calcular una probabilidad acumulada, $P(X \leq 3)$ de calcula con

     ```
     pbinom(3, size = 7, prob = 1/4)
     ```

     ```
     [1] 0.9294434
     ```

     Observa que el resultado es el mismo que el de calcular la probabilidad de que X valga 0, 1, 2, 3 (cada una con dbinom) y sumarlas

     ```
     sum(dbinom(0:3, size = 7, prob = 1/4))
     ```

     ```
     [1] 0.9294434
     ```

   Se recomienda aprender a resolver el ejercicio tanto usando probabilidades puntuales como con probabilidades acumuladas.

   Si quieres más información, puedes leer la sección 1 del tutorial05 (visita https://postdata-statistics.com/courses/introestadistica/).

   Calcula las probabilidades y valores que se indican a continuación.

   (a) $P(X = 4)$.
   (b) $P(X \leq 4)$.
   (c) $P(X < 3)$.
   (d) $P(X \geq 2)$.
   (e) $P(X > 3)$.
   (f) $P(2 \leq X \leq 4)$ (o, lo que es lo mismo, $P((X \geq 2) \cap (X \leq 4))$).
   (g) $P(2 < X < 4)$.
   (h) $P(2 \leq X < 4)$.
   (i) $P(2 < X \leq 4)$.
   (j) $P(X < 2 | X \leq 5)$

2. En los siguientes ejercicios sea X una variable aleatoria de tipo $N(5, 3)$.

   En el caso de una variable aleatoria continua las probabilidades puntuales valen 0, por lo que hay que hacer los cálculos directamente con la función de distribución de probabilidad.

   - La función `pnorm` permite calcular probabilidades acumuladas: $P(X \leq 3)$ de calcula con

     ```
     pnorm(3, mean = 5, sd = 3)
     ```

     ```
     [1] 0.2524925
     ```

     Fíjate en que hay que especificar la media ($mean = 5$) y la desviación típica ($sd = 3$) de la variable normal con la que trabajas.

   Si quieres más información, puedes leer la sección 1 del tutorial05 (visita https://postdata-statistics.com/courses/introestadistica/).

   Calcula las probabilidades y valores que se indican a continuación.

   (a) Calcula $P(X > 6)$, $P(X > 7)$ y, finalmente, $P(X > 10)$. ¿Qué observas?

   (b) Calcula $P(X < 4)$, $P(X < 3)$ y, finalmente, $P(X < 0)$. ¿Ves alguna relación con los valores del anterior apartado? Para entenderlo, puedes hacer un dibujo de la normal con la que estás trabanjando

   (c) $P(4.5 < x < 5.5)$, $P(2 < X < 8)$ y, finalmente, $P(0 < X < 10)$.

   (d) Los valores k_1 y k_2 tales que $P(X < k_1) = 0.90$ y $P(X < k_2) = 0.95$.

   (e) Los valores k_1 y k_2 tales que $P(X > k_1) = 0.1$ y $P(X > k_2) = 0.05$. ¿Ves alguna relación con los valores del anterior apartado?

   (f) $P(X > 2|X \leq 6)$

   Ahora, con la misma variable X, responde a estas preguntas sin usar el ordenador:

   (g) El valor $P(X < 7)$ ¿es mayor o menor que $\frac{1}{2}$?

   (h) El valor $P(X > 8)$ ¿es mayor o menor que $\frac{1}{2}$?

   (i) ¿Cuál de estos dos valores es más grande, $P(X > 4)$ o $P(X < 5.5)$?

   (j) El valor k tal que $P(X > k) = 0.6$ ¿es mayor o menor que 5?

   (k) El valor k tal que $P(X < k) = 0.1$ ¿es mayor o menor que 5? Soluciones en la página 4.

Soluciones.

1. Ejercicio 1, pág. 1

   Fijamos los valores de n y p en este ejercicio:

   ```
   n = 7
   p = 1/4
   ```

   y empezamos con los cálculos.

   (a) Es directo (resuelto en el enunciado):

```r
dbinom(4, size = n, prob = p)
```

```
## [1] 0.05767822
```

(b) Resuelto en el enunciado

```r
     pbinom(4, size = n, prob = p)
```

```
## [1] 0.9871216
```

(c) $P(X < 3) = P(X \leq 2)$, luego:

```r
pbinom(2, size = n, prob = p)
```

```
## [1] 0.7564087
```

(d) $P(X \geq 2) = 1 - P(X < 2) = 1 - P(X \leq 1)$. ¡Cuidado en el segundo paso!

```r
1- pbinom(1, size = n, prob = p)
```

```
## [1] 0.5550537
```

De otra manera, usando `dbinom` y sumando:

```r
sum(dbinom(2:n, size = n, prob = p))
```

```
## [1] 0.5550537
```

(e) $P(X > 3) = 1 - P(X \leq 3)$

```r
1- pbinom(3, size = n, prob = p)
```

```
## [1] 0.07055664
```

(f) Podemos usar que $P(2 \leq X \leq 4) = P(X \leq 4) - P(x \leq 1)$

```r
pbinom(4, size = n, prob = p) - pbinom(1, size = n, prob = p)
```

```
## [1] 0.5421753
```

Compruébalo sumando valores de `dbinom`.

(g) Ahora usamos $P(2 < X < 4) = P(X \leq 3) - P(x \leq 2) = P(X = 3)$

```r
pbinom(3, size = n, prob = p) - pbinom(2, size = n, prob = p)
```

```
## [1] 0.1730347
```

```r
dbinom(3, size = n, prob = p)
```

```
## [1] 0.1730347
```

El resultado es el mismo, claro.

(h) $P(2 \leq X < 4) = P(X \leq 3) - P(x \leq 1)$

```r
pbinom(3, size = n, prob = p) - pbinom(1, size = n, prob = p)

## [1] 0.4844971
```

(i) $P(2 < X \leq 4) = P(X \leq 4) - P(x \leq 2)$

```r
n = 7
pbinom(4, size = n, prob = p) - pbinom(2, size = n, prob = p)

## [1] 0.2307129
```

(j) $P(X < 2|X \leq 5) = \frac{P((X<2)\cap(X\leq 5))}{p(x\leq 5)} = \frac{P(X<2)}{p(x\leq 5)}$

```r
      n = 7
      p = 1/4
pbinom(1, size = n, prob = p)/pbinom(5, size = n, prob = p)

## [1] 0.4455446
```

2. Ejercicio 2, pág. 2

   Todas las respuestas son aproximadas. Definir los parámetros

```r
mu = 5
sigma = 3
```

(a) $P(X > 6) = 1 - P(X < 6)$

```r
1-pnorm(6, mean = mu, sd =sigma)

## [1] 0.3694413
```

También sepuede hacer el cálculo "directo": observa el uso del argumento `lower.tail` (su valor por defecto es `FALSE`)

```r
1 - pnorm(7, mean = mu, sd =sigma)

## [1] 0.2524925

pnorm(7, mean = mu, sd =sigma, lower.tail = FALSE)

## [1] 0.2524925
```

```r
pnorm(10, mean = mu, sd =sigma, lower.tail = FALSE)

## [1] 0.04779035
```

A medida que nos desplazamos hacia la derecha los valores van siendo más y más pequeños.

(b) En este caso:

```r
pnorm(4, mean = mu, sd =sigma)

## [1] 0.3694413

pnorm(3, mean = mu, sd =sigma)

## [1] 0.2524925

pnorm(0, mean = mu, sd =sigma)

## [1] 0.04779035
```

Las respuestas son las mismas del anterior apartado, por la simetría de la curva normal respecto de la media, que está en $\mu = 5$. Por ejemplo, los valores 3 y 7 están a la misma distancia, a izquierda y derecha de μ, respectivamente, y por eso

$$P(X < 3) = P(X > 7) = 0.2525$$

(c) El cálculo es

```r
pnorm(5.5, mean = mu, sd = sigma)-pnorm(4.5, mean = mu, sd = sigma)

## [1] 0.1323677
```

Observa que también se puede calcular así

```r
(pnorm(8, mean = mu, sd = sigma)-0.5)*2

## [1] 0.6826895

(pnorm(10, mean = mu, sd = sigma)-0.5)*2

## [1] 0.9044193
```

(d) Se trata de los percentiles 90 y 95, y la sintaxis vertorial permite calcularlos de una vez

```r
qnorm(c(0.9, 0.95), mean = mu, sd = sigma)

## [1] 8.844655 9.934561
```

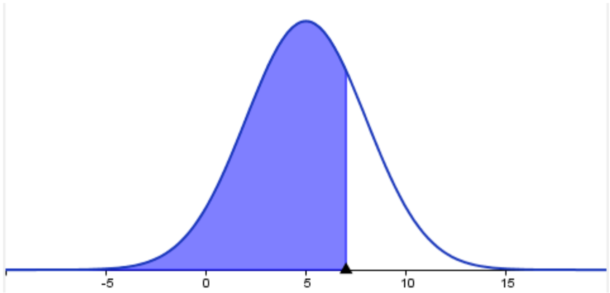
(e) Los mismos valores de k_1 y k_2, de nuevo por la simetría de la curva normal.

(f) $P(X > 2|X \leq 6) = \frac{P((X>2)\cap(X\leq 6))}{p(x\leq 6)}$

```r
numerador2 = pnorm(6, mean = mu, sd = sigma)-
             pnorm(2, mean = mu, sd = sigma)
denominador2 = pnorm(6, mean = mu, sd = sigma)
numerador2/denominador2

## [1] 0.7483894
```

(g) Es mayor que 1/2. Se tiene $P(X < 7) = 0.7475$. Este apartado y el siguiente se resuelven teniendo en cuenta que el área de cada una de las dos mitades de la curva es $\frac{1}{2}$, observando si el punto que hemos tomado está a la derecha o la izquierda de μ, y si la probabilidad que calculamos incluye todos los valores mayores o menores. Por ejemplo, para la primera pregunta tenemos que pensar en un dibujo como este:



con el que resulta evidente que la respuesta es mayor que $\frac{1}{2}$.

(h) Es menor que 1/2. Se tiene $P(X > 8) = 0.1587$.

(i) El valor $P(X > 4)$ es más grande. Se tiene $P(X > 4) = 0.6306$, mientras que $P(X < 5.5) = 0.5662$. En este caso la respuesta es fácil de ver porque el valor 4 está más lejos de $\mu = 5$ que 5.5.

(j) Las ideas para este apartado y el siguiente son las mismas, pero ahora tenemos que pensar en los valores del eje x, en vez de pensar en las probabilidades (áreas) que definen esos valores. El valor tiene que ser menor que 5 (de otra manera, la probabilidad sería menor que $\frac{1}{2}$). Se obtiene $k = 4.24$

(k) El valor tiene que ser menor que 5. Se obtiene $k = 1.156$.