HISTOGRAMAS Y FUNCIÓN EMPÍRICA

domingo, 10 de abril de 2022 17:04

2.11 Para estimar el precio (en dólares) del kilo de asado se examinaron los precios en las pizarras de 24 carnicerías elegidas al azar. Se obtuvieron los siguientes resultados:

8.59	8.77	8.29	7.50	9.18	8.53	10.03	8.97	8.47	9.98	9.00	9.44
8.02	10.35	7.15	9.00	9.15	9.11	7.63	9.66	10.20	9.01	8.73	10.54.

$\} \to x_i$

(**a**) Graficar la función de distribución empírica basada en esa muestra y estimar, usándola, la probabilidad de que el precio de un kilo de asado supere los 9.5 dólares.

(**b**) Usando los intervalos con extremos 7.1, 7.85, 8.35, 9.65, 10.15, 10.90 hallar la función histograma basada en la muestra observada y estimar, usándola, la probabilidad de que el precio de un kilo de asado supere los 9.5 dólares.

a)

FUNCIÓN EMPÍRICA: $\hat{F}(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(x_i \le x)$

Donde $x_1, x_2, \ldots, x_n$ se asumen realizaciónes de las variables aleatorias $X_1, X_2, \ldots, X_n$, todas con distribución $F_X(x)$ e independientes

MUESTRAS ORDENADAS:

```
7.15  7.50  7.63  8.02  8.29  8.47  8.53  8.59  8.73  8.77  8.97  9.00  9.00  9.01
9.11  9.15  9.18  9.44  9.66  9.98 10.03 10.20 10.35 10.54
```

n = 24 (cant. total de muestras) $\Rightarrow F_{X(x)} = \frac{1}{24}\sum_{i=1}^{24} 1\,(x_i \le x)$

$$F_{X(x)} = \tfrac{1}{24}1(7.15 \le x) + \tfrac{1}{24}1(7.50 \le x) + \tfrac{1}{24}1(7.63 \le x) + \tfrac{1}{24}1(8.02 \le x) + \tfrac{1}{24}1(8.29 \le x) + \tfrac{1}{24}1(8.47 \le x) + \tfrac{1}{24}1(8.53 \le x) + \tfrac{1}{24}1(8.59 \le x) + \tfrac{1}{24}1(8.73 \le x) + \tfrac{1}{24}1(8.77 \le x) + \tfrac{1}{24}1(8.97 \le x) + \tfrac{1}{24}1(9.00 \le x) + \tfrac{1}{24}1(9.00 \le x) + \tfrac{1}{24}1(9.01 \le x) + \tfrac{1}{24}1(9.11 \le x) + \tfrac{1}{24}1(9.15 \le x) + \tfrac{1}{24}1(9.18 \le x) + \tfrac{1}{24}1(9.44 \le x) + \tfrac{1}{24}1(9.66 \le x) + \tfrac{1}{24}1(9.98 \le x) + \tfrac{1}{24}1(10.03 \le x) + \tfrac{1}{24}1(10.20 \le x) + \tfrac{1}{24}1(10.35 \le x) + \tfrac{1}{24}1(10.54 \le x)$$

$$F_{X(x)} = \tfrac{1}{24}1(7.15 \le x) + \tfrac{1}{24}1(7.50 \le x) + \tfrac{1}{24}1(7.63 \le x) + \tfrac{1}{24}1(8.02 \le x) + \tfrac{1}{24}1(8.29 \le x) + \tfrac{1}{24}1(8.47 \le x) + \tfrac{1}{24}1(8.53 \le x) + \tfrac{1}{24}1(8.59 \le x) + \tfrac{1}{24}1(8.73 \le x) + \tfrac{1}{24}1(8.77 \le x) + \tfrac{1}{24}1(8.97 \le x) + \tfrac{2}{24}1(9.00 \le x) + \tfrac{1}{24}1(9.01 \le x) + \tfrac{1}{24}1(9.11 \le x) + \tfrac{1}{24}1(9.15 \le x) + \tfrac{1}{24}1(9.18 \le x) + \tfrac{1}{24}1(9.44 \le x) + \tfrac{1}{24}1(9.66 \le x) + \tfrac{1}{24}1(9.98 \le x) + \tfrac{1}{24}1(10.03 \le x) + \tfrac{1}{24}1(10.20 \le x) + \tfrac{1}{24}1(10.35 \le x) + \tfrac{1}{24}1(10.54 \le x)$$

$$F_{X(x)} = \tfrac{1}{24}1(7.15 \le x < 7.50) + \tfrac{2}{24}1(7.50 \le x < 7.63) + \tfrac{3}{24}1(7.63 \le x < 8.02) + \tfrac{4}{24}1(8.02 \le x < 8.29) + \tfrac{5}{24}1(8.29 \le x < 8.47) + \tfrac{6}{24}1(8.47 \le x < 8.53) + \tfrac{7}{24}1(8.53 \le x) + \tfrac{8}{24}1(8.59 \le x < 8.73) + \tfrac{9}{24}1(8.73 \le x < 8.77) + \tfrac{10}{24}1(8.77 \le x < 8.97) + \tfrac{11}{24}1(8.97 \le x < 9.00) + \tfrac{13}{24}1(9.00 \le x < 9.01) + \tfrac{14}{24}1(9.01 \le x < 9.11) + \tfrac{15}{24}1(9.11 \le x < 9.15) + \tfrac{16}{24}1(9.15 \le x < 9.18) + \tfrac{17}{24}1(9.18 \le x < 9.44) + \tfrac{18}{24}1(9.44 \le x < 9.66) + \tfrac{19}{24}1(9.66 \le x < 9.98) + \tfrac{20}{24}1(9.98 \le x < 10.03) + \tfrac{21}{24}1(10.03 \le x < 10.20) + \tfrac{22}{24}1(10.20 \le x < 10.35) + \tfrac{23}{24}1(10.35 \le x < 10.54) + \tfrac{24}{24}1(10.54 \le x)$$

DEBE CUMPLIR:

1. $\hat{F}_X(x) \in [0,1], \forall x \in \mathbb{R}$
2. $\hat{F}_X(x)$ es monótona no decreciente
3. $\hat{F}_X(x)$ es continua a derecha
4. $lim_{x\to-\infty}\hat{F}_X(x) = 0$ y $lim_{x\to\infty}\hat{F}_X(x) = 1$

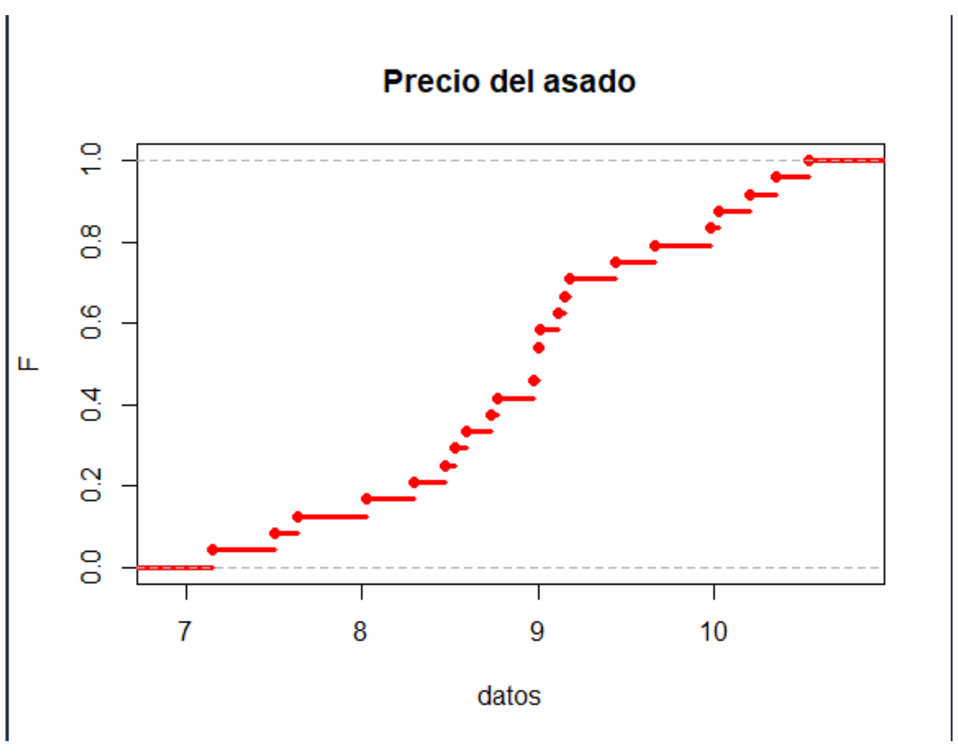
CÓDIGO EN R:

```
muestras<-c(8.59, 8.77, 8.29, 7.50, 9.18, 8.53, 10.03, 8.97, 8.47, 9.98, 9.00, 9.44,
            8.02, 10.35, 7.15, 9.00, 9.15, 9.11, 7.63, 9.66, 10.20, 9.01, 8.73, 10.54)
ordenado<-sort(muestras, decreasing = FALSE)
acumulaciones=ecdf(muestras) #la función ecdf es para
#obtener la función de distribución empírica de una muestra cualquiera

plot(main = "Precio del asado", acumulaciones,col="red",lwd=3,xlab="datos",ylab="F")
```

$P(X > 9.5) = 1 - P(X \le 9.5) = 1 - F_X(9.5) = 1 - 0.75 = 0.25$

```
> acumulaciones(9.5)
[1] 0.75
```



b)

PASOS:

- Se selecciona un origen x_0 y se divide la recta real en intervalos de longitud h

$$B_j = [x_0 + (j-1)h, x_0 + jh], j \in \mathbb{N}$$

No es necesario que todos los intervalos tengan la misma longitud, pero es recomendable que así sea. Esto facilita la lectura.

- Se cuenta cuantas observaciones caen en cada intervalo armando una tabla de frecuencias. Denotamos a la cantidad de observaciones que caen en el intervalo j como n_j
- Para cada intervalo, se divide la frecuencia absoluta por la cantidad total de la muestra n (para convertirlas en frecuencias relativas, análogo a como se hace con las probabilidades) y por la longitud h (para asegurarse que el area debajo del histograma sea igual a 1):

$$f_j = \frac{n_j}{nh}$$

- Se grafica el histograma realizando una barra vertical sobre cada intervalo con altura f_j y ancho h

```
7.15  7.50  7.63  8.02  8.29  8.47  8.53  8.59  8.73  8.77  8.97  9.00  9.00  9.01
9.11  9.15  9.18  9.44  9.66  9.98 10.03 10.20 10.35 10.54
```

intervalos con extremos 7.1, 7.85, 8.35, 9.65, 10.15, 10.90

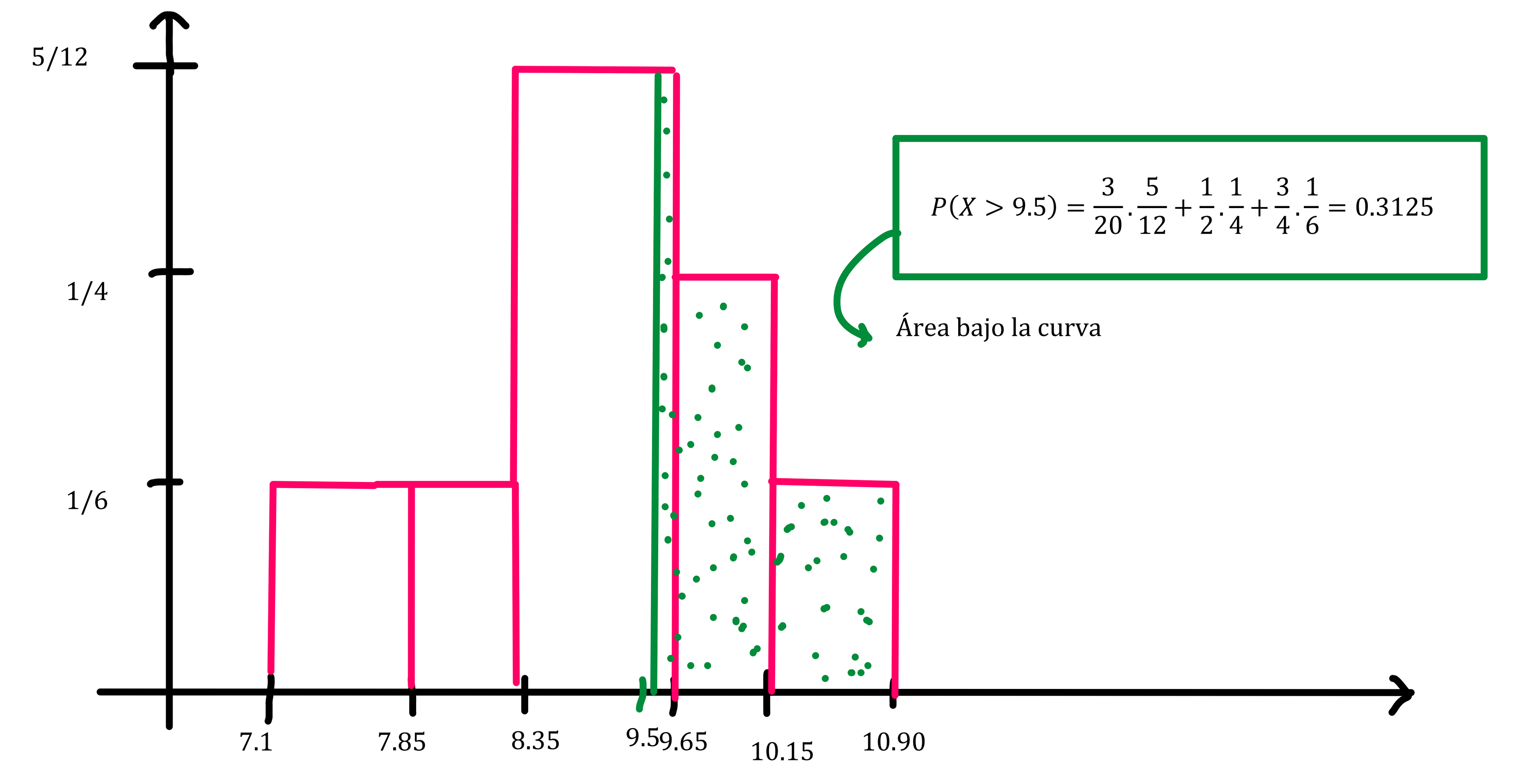
B_j	n_j	h	$f_j = \frac{n_j}{nh}$
[7.1, 7.85)	3	0.75	$\frac{1}{6}$
[7.85, 8.35)	2	0.5	$\frac{1}{6}$
[8.35, 9.65)	13	1.3	$\frac{5}{12}$
[9.65, 10.15)	3	0.5	$\frac{1}{4}$
[10.15,10.90)	3	0.75	$\frac{1}{6}$

recordar: h es la longitud del intervalo

n=24

FUNCIÓN HISTOGRAMA: $\hat{f}_h(x) = \frac{1}{nh}\sum_{i=1}^{n}\sum_{j} \mathbf{1}(x_i \in B_j)\mathbf{1}(x \in B_j)$ $= f_j$. la indicadora del intervalo

Entonces: $f_x(x) = \frac{1}{6}1(7.1 \le x < 7.85) + \frac{1}{6}1(7.85 \le x < 8.35) + \frac{5}{12}1(8.35 \le x < 9.65) + \frac{1}{4}1(9.65 \le x < 10.15) + \frac{1}{6}1(10.15 \le x < 10.90)$



$$P(X > 9.5) = \frac{3}{20}\cdot\frac{5}{12} + \frac{1}{2}\cdot\frac{1}{4} + \frac{3}{4}\cdot\frac{1}{6} = 0.3125$$