

## Autoevaluación

# OCW 2019: Curso práctico para el análisis e inferencia estadística con Mathematica

## Ejercicios propuestos resueltos (1 de 2)

#### Equipo docente del curso

Arrospide Zabala, Eneko Martín Yagüe, Luis Unzueta Inchaurbe, Aitziber Soto Merino, Juan Carlos Durana Apaolaza, Gaizka Bikandi Irazabal, Iñaki

Departamento de Matemática Aplicada Escuela de Ingeniería de Bilbao, Edificio II-I









### **EJERCICIOS DEL BLOQUE I. ESTADÍSTICA DESCRIPTIVA**

#### Ejercicio nº1

#### Enunciado

Un niño apunta en un papel el último número de la matricula de los primeros 60 coches que pasan por delante de su casa

- a) Genere de manera aleatoria la serie estadística de los números apuntados por el niño; utilice las funciones RandomInteger y SeedRandom[34] como semilla
- b) Resuma el conjunto de datos mediante una tabla de frecuencias
- c) Represente el gráfico de barras y el gráfico de barras acumuladas
- d) Calcule los siguientes estadísticos de centralización: media, moda y mediana
- e) Calcule los siguientes estadísticos de dispersión: varianza y desviación típica
- f) Obtenga los siguientes estadísticos de posición: cuartiles y deciles
- g) Calcule los estadísticos de forma: asimetría y curtosis

#### Resolución

Se define la variable aleatoria X="último número de las matrículas observadas por el niño"

- a) Genere de manera aleatoria la serie estadística de los números apuntados por el niño; utilice las funciones RandomInteger y SeedRandom[34] como semilla
  - Semilla para la generación de números aleatorios

#### SeedRandom[34]

• Generación de la serie de números aleatorios

```
ultimo = RandomInteger [{0, 9}, 60]
{3, 5, 5, 4, 5, 0, 7, 4, 1, 2, 6, 1, 6, 5, 6, 2, 0, 8, 9, 0, 8, 8, 1, 5, 8, 0, 1, 2, 2, 6, 9, 0, 9, 9, 6, 8, 1, 4, 1, 8, 0, 3, 3, 4, 9, 1, 6, 1, 1, 2, 1, 4, 3, 1, 3, 4, 0, 6, 7, 9}
```

- b) Resuma el conjunto de datos mediante una tabla de frecuencias
  - Ordenación de la serie y recuento de frecuencias

```
datos = Sort[Tally[ultimo]];
```

■ Tamaño de la serie

#### n = Length[ultimo];

Conjunto de valores de la variable

```
xi = datos[[All, 1]];
```

■ Frecuencias: absoluta (ni) y absoluta acumulada (Ni)

```
ni = datos[[All, 2]]; Ni = Accumulate[ni];
```

■ Frecuencias: relativa (fi) y relativa acumulada (Fi)

```
fi = N[ni/n]; Fi = N[Ni/n];
```





■ Tabla de frecuencias

TableForm[{xi, ni, fi, Ni, Fi},

TableHeadings → {{"Consumos", "Frec. absoluta", "Frec. relativa", "Frec. absoluta acumulada", "Frec. relativa acumulada"}, None}, TableDirections -> Row, TableAlignments → Center]

Consumos Fre	. absoluta	Frec.	relativa	Frec.	absoluta	acumulada	Frec.	relativa	acumulada
--------------	------------	-------	----------	-------	----------	-----------	-------	----------	-----------

0	7	0.116667	7	0.116667
1	11	0.183333	18	0.3
2	5	0.0833333	23	0.383333
3	5	0.0833333	28	0.466667
4	6	0.1	34	0.566667
5	5	0.0833333	39	0.65
6	7	0.116667	46	0.766667
7	2	0.0333333	48	0.8
8	6	0.1	54	0.9
9	6	0.1	60	1.

• c) Represente el gráfico de barras y el gráfico de barras acumuladas

```
datos2 = Map[{#, Count[ultimo, #]} &, Range[0, 9]]
{{0, 7}, {1, 11}, {2, 5}, {3, 5}, {4, 6}, {5, 5}, {6, 7}, {7, 2}, {8, 6}, {9, 6}}
```

Representación del gráfico de barras

```
valores = datos2[[All, 1]];

ni = datos2[[All, 2]];

BarChart[ni, ChartLabels \rightarrow valores, AxesLabel \rightarrow {"xi", "Ni"},

PlotLabel \rightarrow HoldForm[Gráfico de Barras], LabelingFunction \rightarrow Above, BarSpacing \rightarrow 0.2]
```



Representación del gráfico de barras acumuladas

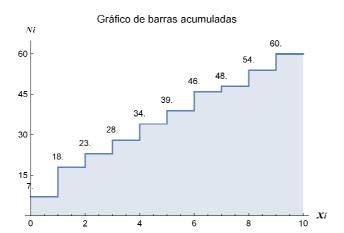
```
Ni = Accumulate[ni];

ListStepPlot[Transpose[{valores, Ni}], AxesLabel \rightarrow {"xi", "Ni"},

PlotLabel \rightarrow "Gráfico de barras acumuladas", LabelingFunction \rightarrow Above,

Ticks \rightarrow {Automatic, {15, 30, 45, 60}}, Filling \rightarrow Axis, PlotRange \rightarrow {0, 65}]
```





- d) Calcule los siguientes estadísticos de centralización: media, moda y mediana
  - Cálculo de la media

media = Mean[ultimo] // N

4.05

■ Obtención de la moda

moda = Commonest[ultimo][[1]]

1

■ Cálculo de la mediana

mediana = Median[ultimo]

4

Solución: 
$$\begin{cases} E[X] = 4.05 \\ Mo[X] = 1 \\ Me[X] = 4 \end{cases}$$

- e) Calcule los siguientes estadísticos de dispersión: varianza y desviación típica
  - Cálculo de la varianza

varianza = (n - 1) \*Variance[ultimo]/n/N

8.88083

■ Cálculo de la desviación típica

destip = Sqrt[varianza] // N

2.98007

Solución: 
$$\begin{cases} Var[X] = 8.88083 \\ SD[X] = 2.98007 \end{cases}$$

- f) Obtenga los siguientes estadísticos de posición: cuartiles, deciles
  - Obtención de los cuartiles

{Q1, Q2, Q3} = Quartiles[ultimo]

**{1, 4, 6**}





- Solución:  $Q_1 = 1, Q_2 = 4, Q_3 = 6$
- Obtención de los deciles

```
q = Range[10, 90, 10];
```

If[IntegerQ[q\*Length[ultimo]/100],
Quantile[ultimo, q/100, {{1/2, 0}, {0, 1}}],
Quantile[ultimo, q/100, {{0, 0}, {1, 0}}]]

 $\{0, 1, 1, 3, 4, 5, 6, 7, 8\}$ 

- Solución:  $P_{10} = 0$ ,  $P_{20} = P_{30} = 1$ ,  $P_{40} = 3$ ,  $P_{50} = 4$ ,  $P_{60} = 5$ ,  $P_{70} = 6$ ,  $P_{80} = 7$ ,  $P_{90} = 8$
- g) Calcule los estadísticos de forma: asimetría y curtosis
  - Obtención del coeficiente de asimetría

```
\gamma_1 = Skewness[ultimo] // N
```

0.238716

- $\blacksquare$  Solución:  $\boxed{\gamma_1 = 0.238716 > 0}$  , la distribución presenta asimetría a la derecha
- Cálculo del coeficiente de curtosis

 $\beta_2 = \text{Kurtosis}[\text{ultimo}] // \text{N}$ 

1.74322

• Solución:  $\beta_2 = 1.74322 < 3$  , por lo que la distribución es platicúrtica

#### Ejercicio nº2

#### Enunciado

En la unidad de maternidad de un hospital se ha analizado la altura de los 50 bebés nacidos durante la última semana.

- a) Genere de manera aleatoria la serie estadística de las alturas de los recién nacidos; utilice la función **SeedRandom[234]** como semilla y la función **RandomInteger** sabiendo que la altura de los recién nacidos está en el rango [45-54]
- b) Obtenga la tabla de frecuencias de la altura de los recién nacidos la última semana con los datos agrupados en intervalos de amplitud 1 cm
- c) Represente gráficamente la información de la tabla de frecuencias del apartado anterior mediante un histograma
- d) Trace un diagrama de caja para la altura de los recién nacidos

#### Resolución

#### Remove["Global`\*"]

- a) Genere de manera aleatoria la serie estadística de las alturas de los recién nacidos; utilice la función SeedRandom[234] como semilla y la función RandomInteger sabiendo que la altura de los recién nacidos está en el rango [45-54]
  - Semilla para la generación de números aleatorios

#### SeedRandom [234]





■ Generación de la serie de números aleatorios y asignación a la variable altura

```
altura = RandomReal [ {45, 54}, 50]
```

```
{45.2849, 47.4157, 47.8184, 51.5263, 46.349, 49.4089, 45.7426, 45.2398, 51.4706, 52.3347, 53.587, 51.9123, 46.4359, 53.2679, 51.5947, 47.9266, 49.0077, 50.4675, 52.0102, 52.1175, 52.4566, 50.1611, 52.4249, 53.6433, 49.8115, 52.6422, 51.0607, 50.7744, 47.4218, 45.5032, 50.6432, 51.16, 50.7061, 50.381, 48.2616, 52.3365, 52.3834, 48.9695, 52.3286, 50.0411, 48.5621, 48.7544, 50.6263, 50.3835, 50.0622, 52.6472, 48.4651, 53.3191, 48.7105, 50.7708}
```

- lacktriangle b) Obtenga la tabla de frecuencias de la altura de los recién nacidos la última semana con los datos agrupados en intervalos de amplitud  $1\ cm$ 
  - Obtención de los intervalos de clase

marca = Map[Mean, intervalos, {1}] // N

```
{45.5, 46.5, 47.5, 48.5, 49.5, 50.5, 51.5, 52.5, 53.5}
```

■ Cálculo de las frecuencias absolutas y frecuencias absolutas acumuladas de cada intervalo

```
ni = BinCounts [altura, {45, 54, 1}]
{4, 2, 4, 6, 3, 11, 6, 10, 4}
```

#### Ni = Accumulate[ni];

• Cálculo de las frecuencias relativas y frecuencias relativas acumuladas de cada intervalo

```
fi = N[ni / Length[altura]];
```

#### Fi = N[Ni / Length[altura]];

■ Tabla de frecuencias

```
tabla = TableForm[{interv, marca, ni, fi, Ni, Fi},
   TableHeadings → {{"Intervalos clase", "Marcas clase", "Fr. absoluta",
        "Fr. relativa", "Fr. absoluta acumulada", "Fr. relativa acumulada"}, None},
   TableDirections → Row, TableAlignments → Center]
```

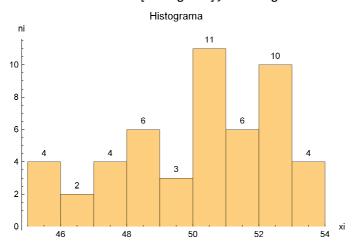
Intervalos clase Marcas clase Fr. absoluta Fr. relativa Fr. absoluta acumulada Fr. relativa acumulada

[45,46)	45.5	4	0.08	4	0.08
[46,47)	46.5	2	0.04	6	0.12
[47,48)	47.5	4	0.08	10	0.2
[48,49)	48.5	6	0.12	16	0.32
[49,50)	49.5	3	0.06	19	0.38
[50,51)	50.5	11	0.22	30	0.6
[51,52)	51.5	6	0.12	36	0.72
[52,53)	52.5	10	0.2	46	0.92
[53,54]	53.5	4	0.08	50	1.



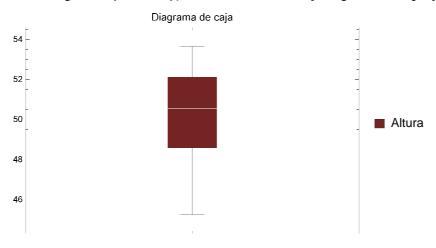
• c) Represente gráficamente la información de la tabla de frecuencias del apartado anterior mediante un histograma

Histogram[altura, {45, 54, 1}, AxesLabel → {HoldForm[xi], HoldForm[ni]},
PlotLabel → HoldForm[Histograma], LabelingFunction → Above]



- Como todos los intervalos de clase tienen la misma amplitud, en el histograma se representan la frecuencias absolutas
- d) Trace un diagrama de caja para la altura de los recién nacidos

BoxWhiskerChart[altura, {"Outliers", {"Outliers", " $\bullet$ "}, {"FarOutliers", " $\circ$ "}}, ChartLegends  $\rightarrow$  {"Altura"}, PlotLabel  $\rightarrow$  HoldForm["Diagrama de caja"], ChartStyle  $\rightarrow$  05]







#### **EJERCICIOS DEL BLOQUE II. PROBABILIDAD**

#### Ejercicio nº3

Obtenga el espacio muestral de los siguientes experimentos:

- a) lanzar 3 dados equilibrados y obtener la suma de las caras
- b) escoger dos objetos de un conjunto de tres (con o sin repetición)

#### Resolución

#### Remove["Global`\*"]

a) lanzar 3 dados equilibrados y obtener la suma de las caras

$$b = Table[i + j + k, \{i, 6\}, \{j, 6\}, \{k, 6\}];$$

DeleteDuplicates [Sort[Flatten[b]]]

- Solución:  $\Omega_1 = \{3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18\}$
- b) escoger dos objetos de un conjunto de tres (con o sin repetición)
  - se definen tres objetos, por ejemplo, un triangulo rojo, un rectángulo azul y un circulo verde

#### objetos =

{Graphics[{Red, Triangle[]}], Graphics[{Blue, Rectangle[]}], Graphics[{Green, Disk[]}]}



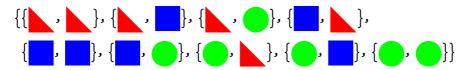
Se escogen dos sin repetición

#### Subsets[objetos, {2}]



Se escogen dos con repetición

#### Tuples[objetos, {2}]



#### Ejercicio nº4

#### Enunciado

Un niño tiene 3 cubos rojos, 3 cubos azules y 3 cubos verdes (todos ellos de diferentes tonalidades) guardados en una caja. Si el niño elige de manera aleatoria 3 cubos, ¿cuál es la probabilidad de que el primero sea rojo, el segundo azul y el tercero verde?





#### Resolución

#### Remove["Global`\*"]

Definición de los nueve cubos

azules = 
$$\{ \bigcap, \bigcap, \bigcap \}$$
; verdes =  $\{ \bigcap, \bigcap, \bigcap \}$ ; rojos =  $\{ \bigcap, \bigcap, \bigcap \}$ ;

■ Elementos contenidos en la caja

caja = Flatten[{azules, verdes, rojos}]



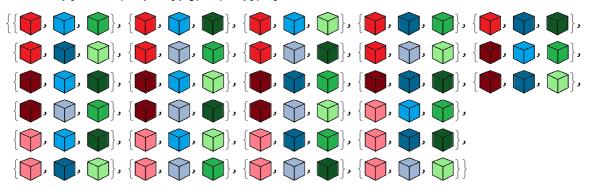
■ Número total de elementos (m) y número de elementos seleccionados (r)

#### m = Length [caja]; r = 3;

■ Se elige un cubo de cada color

 Elección de un cubo rojo en primer lugar, un cubo azul en segundo lugar y un cubo verde en tercer lugar

casos = Map[Flatten, Tuples[{rj, az, vr}], 1]



Número de casos favorables

#### favorables = Length [casos]

27

Número de casos totales (variaciones sin repetición):  $V_{m,r} = \frac{m!}{(m-r)!}$ 

posibles = Factorial[m] / Factorial[m - r]

504

Aplicando la regla de Laplace

prob = favorables / posibles // N

0.0535714

■ La probabilidad pedida es:  $P\left( \bigcap, \bigcap, \bigcap \right) = \frac{27}{504} = 0.0536$ 





#### Ejercicio nº5

#### Enunciado

Sea el experimento consistente en lanzar 10 veces un dado equilibrado.

- a) Genere la serie estadística: utilice la función RandomChoice con SeedRandom[234] como semilla
- b) Repita el experimento lanzando el dado 6000 veces; en esta ocasión, genere la serie estadística utilizando la función RandomChoice con SeedRandom[1234] como semilla
- c) Represente gráficamente la información de las tablas de frecuencias de los apartados anteriores mediante diagramas de barras

#### Resolución

#### Remove["Global`\*"]

- a) Genere la serie estadística: utilice la función RandomChoice con SeedRandom[234] como semilla
  - Semilla para la generación de números aleatorios

#### SeedRandom [234]

• Generación de la serie de números aleatorios

numeros = 
$$\{1, 2, 3, 4, 5, 6\}$$
;

dado10 = RandomChoice[numeros, 10]

■ Solución:  $serie = \{1, 3, 1, 2, 6, 2, 1, 1, 2, 2\}$ 

#### t10 = Tally [Sort [dado10]]

$$\{\{1,4\},\{2,4\},\{3,1\},\{6,1\}\}$$

- b) Repita el experimento lanzando el dado 6000 veces; en esta ocasión, genere la serie estadística utilizando la función RandomChoice con SeedRandom[1234] como semilla
  - Semilla para la generación de números aleatorios

#### SeedRandom[1234]

Generación de la serie de números aleatorios

dado6000 = RandomChoice [caras, 6000];

#### t6000 = Tally [Sort [dado6000]]

$$\{\{1, 967\}, \{2, 972\}, \{3, 1066\}, \{4, 1013\}, \{5, 1018\}, \{6, 964\}\}$$

■ Solución:

caras del dado	1	2	3	4	5	6
frecuencias	967	972	1066	1013	1018	964



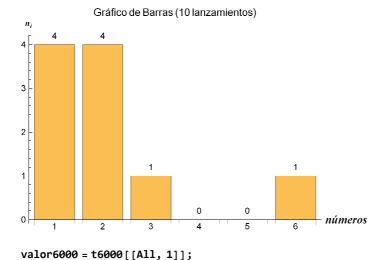


 c) Represente gráficamente la información de las tablas de frecuencias de los apartados anteriores mediante diagramas de barras

```
Grid[{{"10 veces", "6000 veces"},
    {TableForm[{Transpose[t10][[1]], Transpose[t10][[2]], N[Transpose[t10][[2]] / 10]},
    TableHeadings → {{"Nº", "Frecuencia absoluta", "Frecuencia relativa"}},
    TableDirections -> Row, TableAlignments → Center],
    TableForm[{Transpose[t6000][[1]], Transpose[t6000][[2]], N[Transpose[t6000][[2]] / 6000]},
    TableHeadings → {{"Nº", "Frecuencia absoluta", "Frecuencia relativa"}},
    TableDirections -> Row, TableAlignments → Center]}}]
```

10 veces 6000 veces Nº Frecuencia absoluta Frecuencia relativa Nº Frecuencia absoluta Frecuencia relativa 967 0.161167 0.4 2 972 0.162 4 0.177667 2 0.4 3 1066 4 1013 0.168833 3 1 0.1 5 6 1 0.1 1018 0.169667 6 964 0.160667

```
ta10 = Map[{#, Count[dado10, #]} &, Range[1, 6]]  \{\{1, 4\}, \{2, 4\}, \{3, 1\}, \{4, 0\}, \{5, 0\}, \{6, 1\}\}  valor10 = ta10[[All, 1]];  ni10 = ta10[[All, 2]];  BarChart[ni10, ChartLabels \rightarrow valor10, AxesLabel \rightarrow {"números", "n<sub>i</sub>"}, PlotLabel \rightarrow HoldForm[Gráfico de Barras (10 lanzamientos)], LabelingFunction \rightarrow Above, BarSpacing \rightarrow 0.2]
```



```
ni6000 = t6000[[All, 2]];

BarChart[ni6000, ChartLabels \rightarrow valor6000, AxesLabel \rightarrow {"números", "n_i"}, PlotLabel \rightarrow HoldForm[Gráfico de Barras (6000 lanzamientos)], LabelingFunction \rightarrow Above, BarSpacing \rightarrow 0.2]
```







