Trabajo Práctico 1

Especialización en Inteligencia Artificial

Probabilidad y Estadistica

Abril Noguera

Universidad de Buenos Aires

Laboratorio de Sistemas Embebidos

Especialización en Inteligencia Artificial

Probabilidad y Estadística para la Inteligencia Artificial

Docente: Camilo Argoty

Nombre:	Abril Noguera	Código:	a2017
Fecha:	30-03-25		NOGUERA ABRIL

PRIMER TRABAJO PRÁCTICO

1. (3 puntos) De 10 monedas hay 2 monedas falsas, que tienen probabilidad 0,5 de mostrar cara al ser lanzadas. Si se toma una moneda al azar, se lanza 16 veces, y en todas ellas se obtiene cara, ¿qué es más probable, que la moneda elegida sea justa o que esté cargada? Dar las probabilidades tanto de que la moneda elegida sea falsa, como de que sea justa.

Planteamiento del Problema

Total de monedas: 10

- Monedas justas (J): 8 (probabilidad de cara = 0.5).
- Monedas falsas (F): 2 (probabilidad de cara = 0.5).

Experimento: Se elige una moneda al azar y se lanza 16 veces, obteniendo 16 caras seguidas.

Pregunta: ¿Es más probable que la moneda sea justa (J) o falsa (F)?

Note: Las monedas cargadas en realidad son también monedas justas al tener probabilidad de 0.5

```
=== Parametros ===
Total de Monedas: 10
Monedas Falsas: 2
Probabilidad de Cara en Moneda Falsa: 0.5
Lanzamientos Observados: 16

=== PROBABILIDADES A PRIORI ===
P(Moneda Justa): 0.8000  (8/10)
P(Moneda Falsa): 0.2000  (2/10)

=== PROBABILIDAD DE OBSERVAR LOS LANZAMIENTOS ===
P(16 caras | Justa): 0.5^16 = 0.000015
P(16 caras | Falsa): 0.5^16 = 0.000015

=== PROBABILIDAD TOTAL (EVIDENCIA) ===
P(16 caras) = [P(16 caras | Justa)*P(Justa)] + [P(16 caras | Falsa)*P(Fals
a)]
                     = (0.000015 * 0.8000) + (0.000015 * 0.2000)
                     = 0.000015

=== PROBABILIDADES POSTERIORES (TEOREMA DE BAYES) ===
P(Justa | 16 caras) = [P(16 caras | Justa)*P(Justa)] / P(16 caras)
                           = (0.000015 * 0.8000) / 0.000015
                           = 0.800000

P(Falsa | 16 caras) = [P(16 caras | Falsa)*P(Falsa)] / P(16 caras)
                           = (0.000015 * 0.2000) / 0.000015
                           = 0.200000


=== Árbol de Decisión ===
Elección de Moneda
├── Moneda Justa (P=0.8000)
│   └── 16 caras seguidas (P=0.0000)
└── Moneda Falsa (P=0.2000)
    └── 16 caras seguidas (P=0.0000)

=== Resultados ===
Probabilidad de ser justa: 0.8000
Probabilidad de ser falsa: 0.2000

Conclusión: Es más probable que la moneda sea JUSTA.
```

Al tener la **misma probabilidad de cara (0.5) en monedas justas y falsas**, el resultado de 16 caras seguidas no proporciona información útil para distinguirlas. Por lo tanto, las probabilidades posteriores se mantienen iguales a las iniciales (prior):

- La probabilidad de que sea justa sigue siendo 0.8.
- La probabilidad de que sea falsa sigue siendo 0.2.

En este caso, no hay actualización de creencias porque las monedas son indistinguibles en su comportamiento.

¿Qué pasa si la moneda falsa tiene P(cara) > 0.5?

Se hace la misma prueba destinando una **probabilidad de 0.7 de obtener cara en monedas falsas**.

```
=== Parametros ===
Total de Monedas: 10
Monedas Falsas: 2
Probabilidad de Cara en Moneda Falsa: 0.7
Lanzamientos Observados: 16

=== PROBABILIDADES A PRIORI ===
P(Moneda Justa): 0.8000  (8/10)
P(Moneda Falsa): 0.2000  (2/10)

=== PROBABILIDAD DE OBSERVAR LOS LANZAMIENTOS ===
P(16 caras | Justa): 0.5^16 = 0.000015
P(16 caras | Falsa): 0.7^16 = 0.003323

=== PROBABILIDAD TOTAL (EVIDENCIA) ===
P(16 caras) = [P(16 caras | Justa)*P(Justa)] + [P(16 caras | Falsa)*P(Fals
a)]
                     = (0.000015 * 0.8000) + (0.003323 * 0.2000)
                     = 0.000677

=== PROBABILIDADES POSTERIORES (TEOREMA DE BAYES) ===
P(Justa | 16 caras) = [P(16 caras | Justa)*P(Justa)] / P(16 caras)
                           = (0.000015 * 0.8000) / 0.000677
                           = 0.018035

P(Falsa | 16 caras) = [P(16 caras | Falsa)*P(Falsa)] / P(16 caras)
                           = (0.003323 * 0.2000) / 0.000677
                           = 0.981965


=== Árbol de Decisión ===
Elección de Moneda
├── Moneda Justa (P=0.8000)
│   └── 16 caras seguidas (P=0.0000)
└── Moneda Falsa (P=0.2000)
    └── 16 caras seguidas (P=0.0033)

=== Resultados ===
Probabilidad de ser justa: 0.0180
Probabilidad de ser falsa: 0.9820

Conclusión: Es más probable que la moneda sea FALSA.
```

La probabilidad de que la moneda sea falsa es mucho mayor. Aunque solo el 20% de las monedas esten cargadas, el resultado de 16 caras seguidas es tan inverosímil con una moneda justa que la conclusión se inclina decisivamente hacia la moneda falsa.

2. (3 puntos) Sean X e Y dos v.a. continuas con densidad conjunta:

$$f_{X,Y} = \begin{cases} Ky & 25x^2 \leq y \leq 4x \\ 0 & \text{en otro caso} \end{cases}$$

Encontrar:

a) Determine el valor de K

b) Encuentre la densidad marginal $f_Y(y)$ de Y

c) Encuentre la densidad condicional $f_{X|Y}(x|y)$ de X dado Y

a) Determine el valor de K

$$f_{X,Y} = \begin{cases} Ky & 25x^2 \leqslant y \leqslant 4x \\ 0 & \text{en otro caso} \end{cases}$$

a) determinar el valor de K.

densidad conjunta : $\iint_{\mathbb{R}^2} f_{X,Y}(x,y)\,dx\,dy = 1$

I) identifico limites de integración

$25x^2 = 4x \rightarrow x(25x-4)=0$ ⟨ $x=0$, $x=\frac{4}{25}$ $\quad x \in [0, \frac{4}{25}]$

II) calculo la integral

$$\int_0^{\frac{4}{25}} \int_{25x^2}^{4x} Ky\,dy\,dx = 1$$

en y $\rightarrow$ $\displaystyle\int_{25x^2}^{4x} Ky\,dy = K\left[\frac{y^2}{2}\right]_{25x^2}^{4x} = K\left(8x^2 - \frac{625x^4}{2}\right)$

en x $\rightarrow$ $\displaystyle K\int_0^{\frac{4}{25}} 8x^2 - \frac{625x^4}{2}\,dx = K\left[\frac{8x^3}{3} - \frac{625x^5}{10}\right]_0^{\frac{4}{25}} =$

$$= K\left(\frac{8\left(\frac{4}{25}\right)^3}{3} - \frac{625\left(\frac{4}{25}\right)^5}{10}\right) = \frac{1024}{234375}K$$

$\overset{=1}{\rightarrow}$ $\displaystyle\frac{1024}{234375}K = 1 \rightarrow K = \frac{234375}{1024} \simeq 228{,}8818$

```
a) Valor de K: 228.881835937500
```

Encuentre la densidad marginal fY(y) de Y

b) densidad marginal $f_Y(y)$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx$$

i) determino límites de x para y fijo: $25x^2 \leq y \leq 4x$

- $x \geq \frac{y}{4}$
- $x \leq \sqrt{\frac{y}{25}}$
- $y \geq 0$
- $\frac{y}{4} \leq \sqrt{\frac{y}{25}} \rightarrow y \leq \frac{16}{25}$

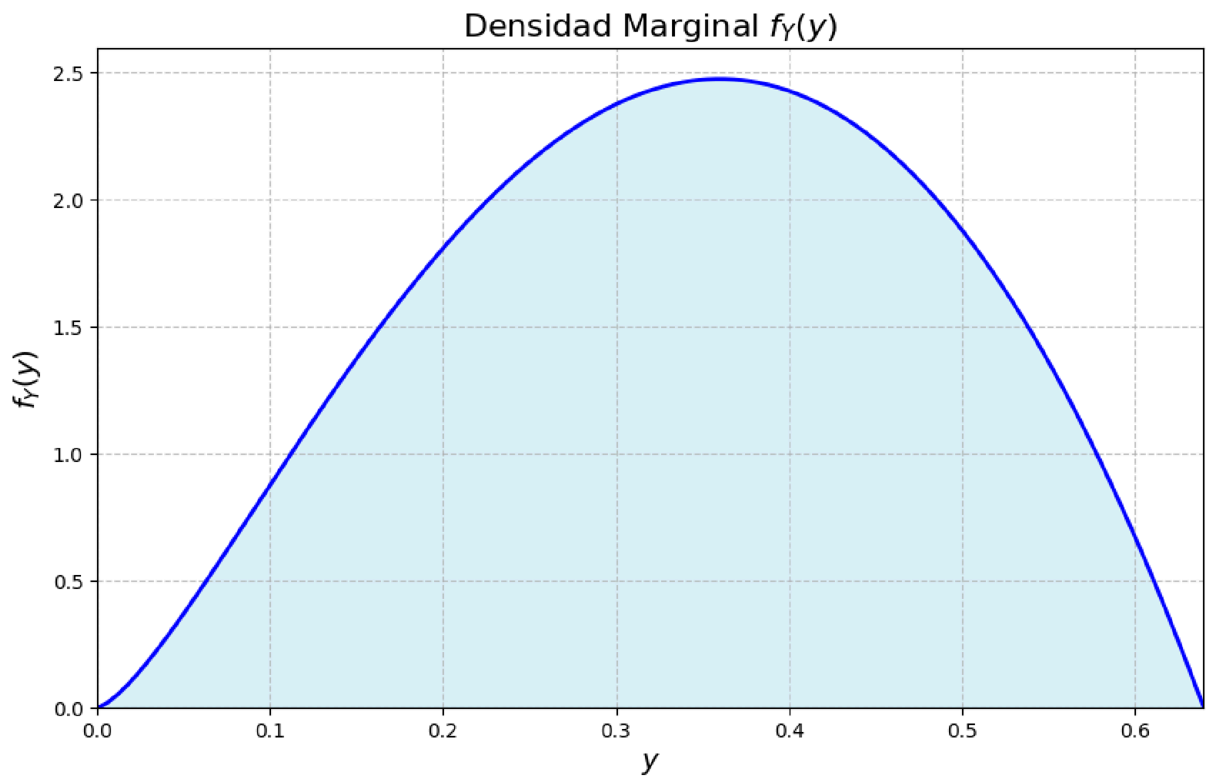
ii) calculo la integral para $y \in [0, \frac{16}{25}]$

$$f_Y(y) = \int_{\frac{y}{4}}^{\sqrt{\frac{y}{25}}} \frac{234375}{1024} y\,dx = \frac{234375}{1024} y \left(\sqrt{\frac{y}{25}} - \frac{y}{4}\right) =$$

$$= \frac{46875}{1024} y^{\frac{3}{2}} - \frac{234375}{4096} y^2 \simeq 45.7764\, y^{\frac{3}{2}} - 57.2205\, y^2$$

$$f_Y(y) = \begin{cases} \frac{46875}{1024} y^{\frac{3}{2}} - \frac{234375}{4096} y^2 & 0 \leq y \leq \frac{16}{25} \\ 0 & \text{en otro caso} \end{cases}$$

```
b) Densidad marginal f_Y(y):
                3/2                    2
45.7763671875·y    - 57.220458984375·y
```



Encuentre la densidad condicional fX|Y (x|y) de X dado Y

c) densidad condicional $f_{X|Y}(x|y)$

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

$$f_{X|Y}(x|Y) = \frac{\cancel{\frac{234375}{1024}}\,\cancel{y}}{\cancel{\frac{234375}{1024}}\,\cancel{y}\left(\sqrt{\frac{y}{25}} - \frac{y}{4}\right)} = \frac{1}{\sqrt{\frac{y}{25}} - \frac{y}{4}} = \frac{20}{4\sqrt{y} - 5y}$$

$$f_{X|Y}(x|Y) = \begin{cases} \dfrac{20}{4\sqrt{y} - 5y} & \dfrac{y}{4} \leq x \leq \sqrt{\dfrac{y}{25}} \\ 0 & \text{en otro caso} \end{cases}$$

```
c) Densidad condicional f_{X|Y}(x|y):

             228.8818359375⋅y
───────────────────────────────────────────────
                  3/2                         2
45.7763671875⋅y     - 57.220458984375⋅y
```

3. (4 puntos) Don Francisco es un pequeño comerciante de barrio que posee un supermercado, con el que sostiene su familia. Uno de sus hijos, Matías, quien recién inicia a cursar la Especialización en Inteligencia Artificial del LSE de la UBA, le propone hacer un análisis de las ventas durante el año anterior, con el fin de hacer pronósticos para el año siguiente, lo que a don Francisco le parece buena idea.

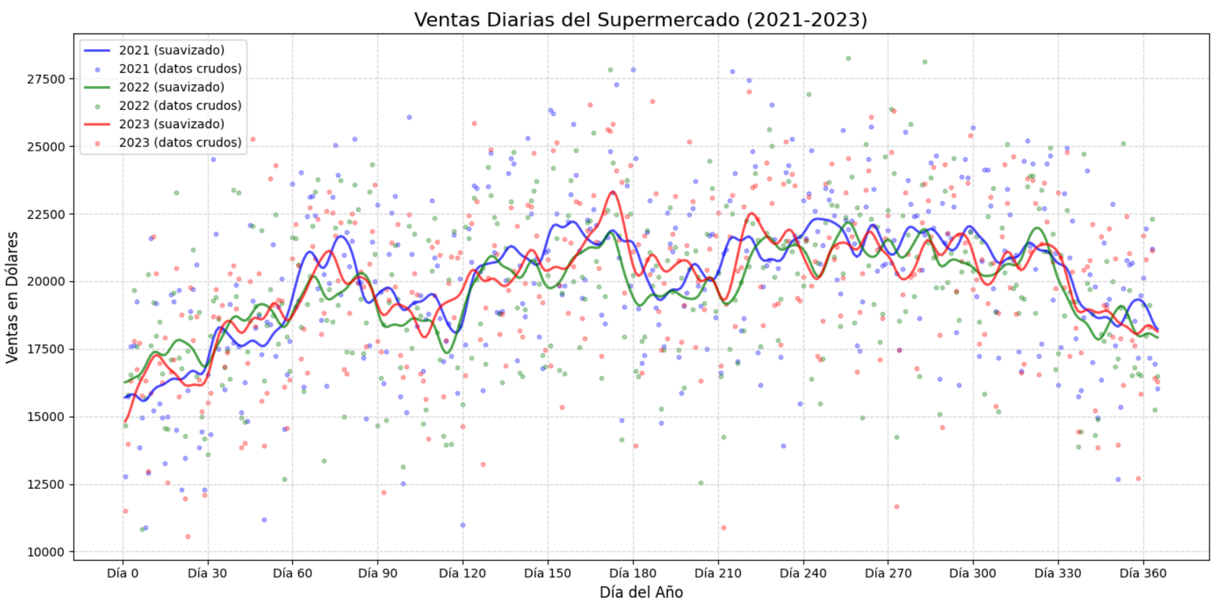
   Don Francisco le entrega a Matías el cuaderno donde tiene registrado el valor total de sus ventas en cada día del año. Con esta información, Matías construye una tabla en la cual la primera columna corresponde a la fecha y la segunda corresponde al monto de las ventas, en dólares para evitarse dolores de cabeza con la inflación. Matías no se siente muy seguro de la tarea a realizar, así que les pide ayuda a ustedes para abordar el problema.

   A partir del archivo de datos correspondiente a su grupo, determine una función empírica de distribución y una aproximación a la función de densidad de las ventas del supermercado de Don Francisco para cada año de registro (2021, 2022 y 2023):

Out[52]:

	Fecha	Ventas
0	2021-01-01	12766.701233
1	2021-01-02	15744.349509
2	2021-01-03	17603.437166
3	2021-01-04	17629.888008
4	2021-01-05	19241.714589

Ventas Diarias del Supermecado a lo largo de los Años



Función Empírica de Distribución (CDF)

Para una variable aleatoria X (ventas diarias), la CDF se define como: FX(x) = P(X<=x)

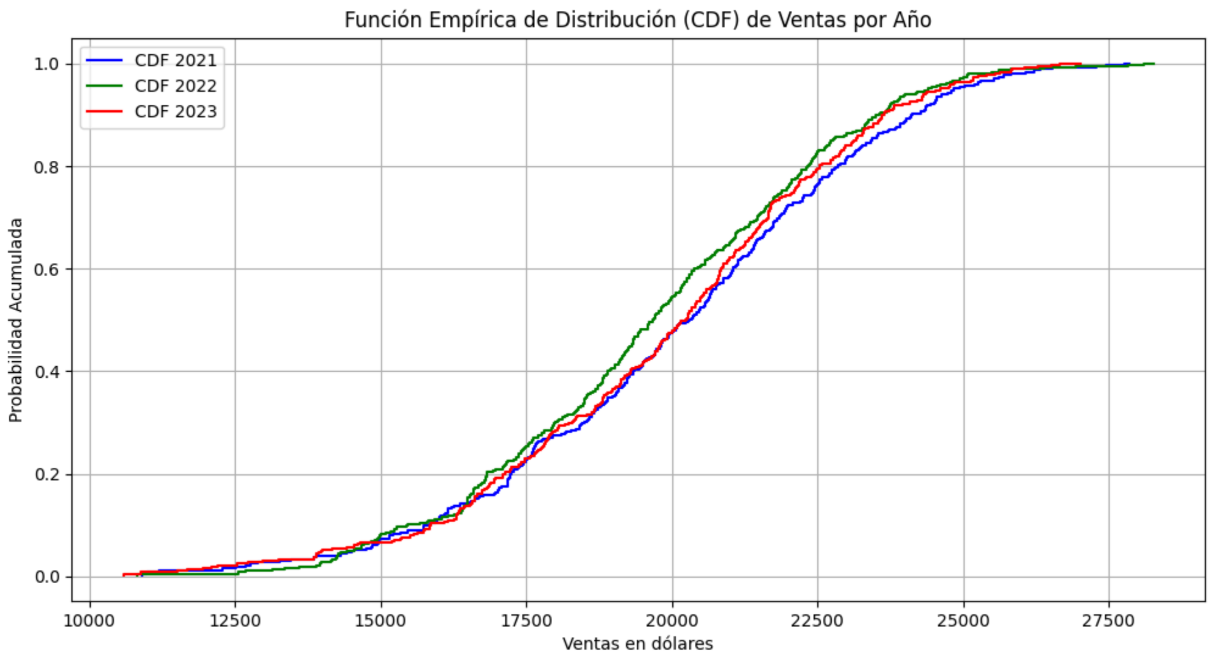
La probabilidad acumulada de que las ventas diarias sean menores o iguales a un valor específico.

```
Definiendo x como $20000

Año : 2021
- La CDF empírica (probabilidad de ventas ≤ $20000) es: 0.47 (47%)

Año : 2022
- La CDF empírica (probabilidad de ventas ≤ $20000) es: 0.54 (54%)

Año : 2023
- La CDF empírica (probabilidad de ventas ≤ $20000) es: 0.47 (47%)
```



Out[105…

Año	count	mean	std	min	25%	
2021	365.0	20082.548050	3229.895938	10901.174491	17641.152381	20342.14
2022	365.0	19697.257088	3024.879804	10815.563915	17491.851787	19694.17
2023	365.0	19906.809559	3114.299003	10584.177191	17799.217434	20247.67

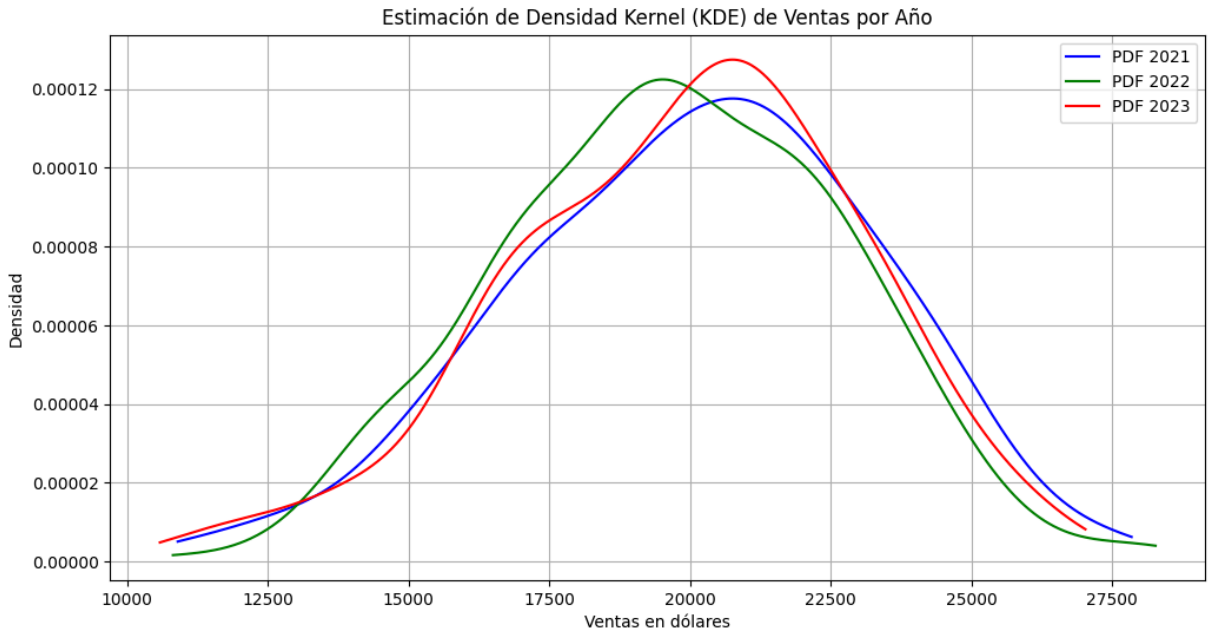
Conclusiones CDF

- 2022 domina en probabilidades acumuladas para montos bajos/medios
- 2023 tiene una curva más empinada entre USD15,000 y USD 22,000 → Menor dispersión (ventas más agrupadas cerca de la moda).
- En 2022, el 54% de las ventas era de valores menores a USD 20,000. Mientras que en 2021 y 2023, el 47% de las ventas era de valores menores a USD 20,000.

Estimación de Densidad Kernel (KDE)

Para una variable aleatoria continua X (ej: ventas diarias), la función de densidad fX(x) describe cómo se distribuye la probabilidad en cada punto x.

El KDE estima esta densidad a partir de datos muestrales, sin asumir una forma paramétrica conocida (como la normal o exponencial).



Conclusiones KDE

- El KDE de 2023 tiene un pico más pronunciado en ~$20,500 USD, lo que indica que este año tuvo más días con ventas cercanas a ese valor.
- En 2023, el pico de densidad (moda) estuvo en ~ $20,500USD.Mientrasqueelañoanteriorestuvoen$ 19,500 USD. Las ventas más frecuentes aumentaron ~$1,000 USD en 2023, indicando un crecimiento en las ventas típicas.
- La curva de 2023 es más estrecha que la de 2021 y 2022, lo que significa menos variabilidad en las ventas diarias.

Link al Notebook

Se puede encontrar el trabajo completo en el siguiente link: Repositorio GitHub

This notebook was converted with convert.ploomber.io