Métodos Estadísticos y Simulación

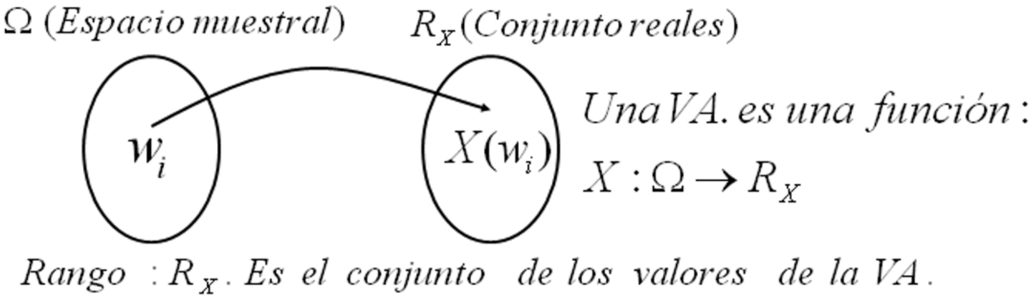
Unidad 3: Variable aleatoria

Mg. J. Eduardo Gamboa U.

2025-10-05

Variable aleatoria

Una variable aleatoria es cualquier función que tiene como dominio a los elementos que constituyen el espacio muestral (Ω) de un experimento aleatorio, y como rango a un subconjunto de los números reales.



Variable aleatoria discreta

Una variable aleatoria discreta es aquella que puede tomar un número finito o numerable de valores. Sus valores posibles conforman su rango R_x y a cada uno se le puede asignar una probabilidad exacta mediante una función $f(x) = P(X = x)$, llamada función de probabilidad.

Ejemplo

Una planta empacadora revisa lotes de 8 productos. Se define la variable aleatoria X como el número de productos defectuosos registrados.

Rango de la variable: $R_x = \{0, 1, 2, 3, 4, 5, 6, 7, 8\}$

Ejemplo

Un área de atención al cliente registra cuántos reclamos se reciben por día. Aunque puede variar, se observan típicamente entre 0 y 5 reclamos. Se define la v.a. X: número de reclamos que recibe por día

Rango de la variable: $R_x = \{0, 1, 2, 3, 4, 5\}$

Ejemplo

En una parcela agrícola se registra cuántas plantas con plagas hay por cada 100 metros cuadrados. Este conteo puede ser 0, 1, 2, etc. La variable aleatoria X es el número de plantas con plagas por cada $100m^2$.

Rango de la variable: $R_x = \{0, 1, 2, ...\}$

Ejercicio

Una empresa controla la calidad de su línea de producción registrando el número de productos defectuosos por cada lote de 500 unidades. La empresa sigue la política de que si en un lote aparecen 5 defectuosos, la producción se detiene de inmediato para inspección. Definir la variable aleatoria y su rango.

Función de probabilidad de una variable aleatoria discreta

Para una variable aleatoria discreta X, se denota como $f(x)$ su función de probabilidad, tal que satisface las siguientes condiciones:

- $f(x) = P(X = x) = 0, \quad x \notin R_x$
- $f(x) = P(X = x) > 0, \quad x \in R_x$
- $\sum_{x_i \in R_x} f(x_i) = \sum_{x_i \in R_x} P(X = x_i) = 1$

Ejemplo

En una planta de empaquetado de frutas, se selecciona una caja con 3 frutas al azar para control de calidad. Cada fruta puede estar en buen estado (B) o defectuosa (D).

Queremos estudiar el comportamiento de la variable aleatoria X: número de frutas en buen estado dentro de la caja inspeccionada

Espacio muestral: $\Omega = \{BBB, BBD, BDB, DBB, BDD, DBD, DDB, DDD\}$

Hallando el rango de X

$X(DDD) = 0$

$X(BDD) = X(DBD) = X(DDB) = 1$

$X(BBD) = X(BDB) = X(BBD) = 2$

$X(BBB) = 3$

Entonces: $R_x = \{0, 1, 2, 3\}$

Hallando las probabilidades:

$$f(0) = P(X = 0) = 1/8 = 0.125$$

$$f(1) = P(X = 1) = 3/8 = 0.375$$

$$f(2) = P(X = 2) = 3/8 = 0.375$$

$$f(3) = P(X = 3) = 1/8 = 0.125$$

Definiendo la función de probabilidades:

x	0	1	2	3
$f(x) = P(X = x)$	0.125	0.375	0.375	0.125

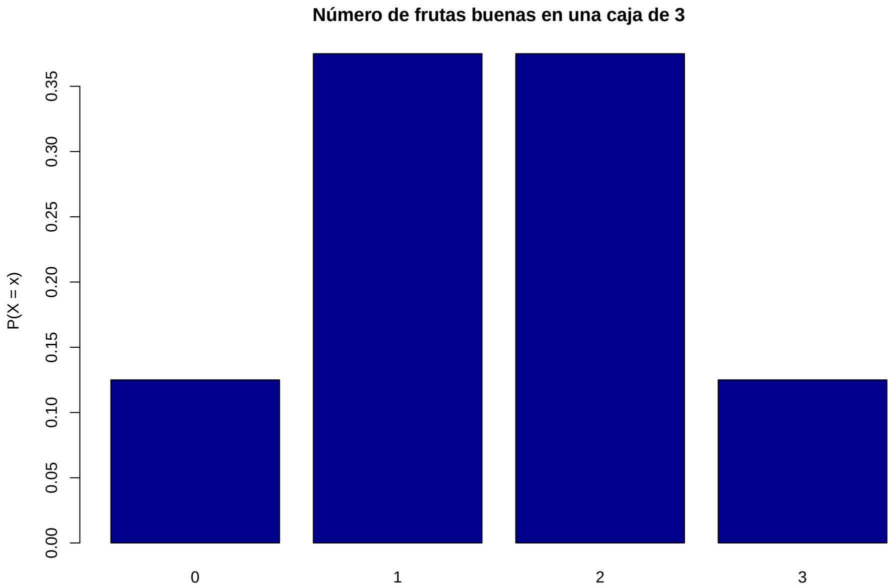
Se cumple:

$\sum_{x_i \in R_x} f(x_i) = f(0) + f(1) + f(2) + f(3)$

$\sum_{x_i \in R_x} f(x_i) = 0.125 + 0.375 + 0.375 + 0.125 = 1$

Otra presentación:

$$f(x) = P(X = x) = \begin{cases} 0.125, & \text{si } x \in \{0, 3\} \\ 0.375, & \text{si } x \in \{1, 2\} \end{cases}$$


Número de frutas buenas en una caja de 3
P(X = x)
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0
1
2
3

Ejemplo

El número de visitas por día que realiza un técnico agropecuario para verificar el manejo de campo de los agricultores de espárragos, es una v.a. con la siguiente f.d.p.:

$$f(x) = P(X = x) = \begin{cases} c, & \text{si } x \in \{1, 2\} \\ \frac{1}{8}, & \text{si } x \in \{3, 4\} \\ 2c, & \text{si } \in \{5, 6\} \\ 0, & \text{Otro caso.} \end{cases}$$

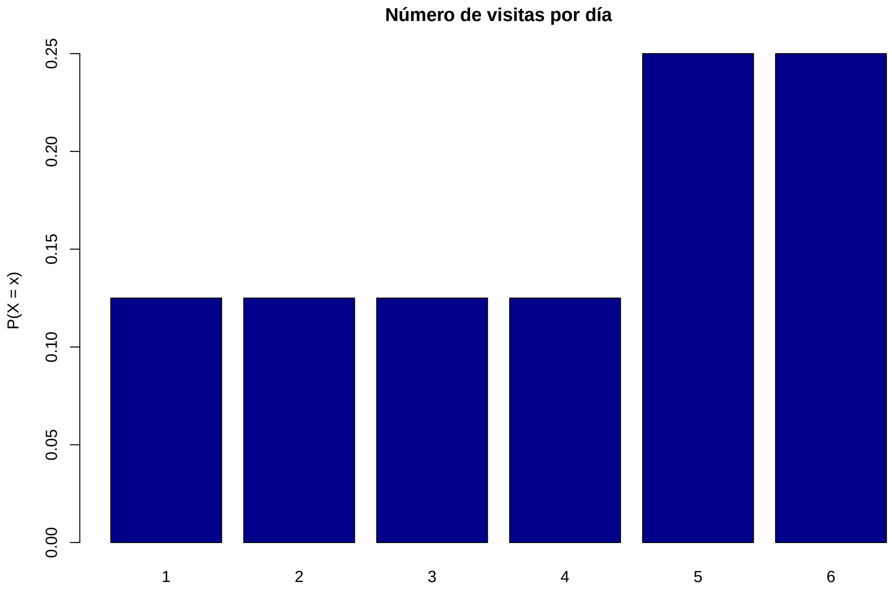
Sea la V.A.D. X=Número de visitas por día del técnico agropecuario.

$R_x = \{1, 2, 3, 4, 5, 6\}$, por propiedad: $\sum_{x_i \in R_x} f(x_i) = 1$

$\sum_{x_i \in R_x} f(x_i) = f(1) + f(2) + f(3) + f(4) + f(5) + f(6)$

$\sum_{x_i \in R_x} f(x_i) = c + c + 1/8 + 1/8 + 2c + 2c = 1 \quad \rightarrow \quad c = \frac{1}{8}$

$$f(x) = P(X = x) = \begin{cases} \frac{1}{8}, & \text{si } x \in \{1, 2, 3, 4\} \\ \frac{2}{8}, & \text{si } x \in \{5, 6\} \\ 0, & \text{Otro caso.} \end{cases}$$


Número de visitas por día
P(X = x)
0.00
0.05
0.10
0.15
0.20
0.25
1
2
3
4
5
6

Ejercicio

Un supervisor de atención al cliente evalúa aleatoriamente 2 llamadas realizadas por los operadores de un call center, para medir la calidad del servicio.

Cada llamada se califica con una puntuación de 1 (mala), 2 (regular) o 3 (buena), según una rúbrica de desempeño.

La selección de llamadas es con reposición porque se toman de una base grande y se permite que una misma calificación se repita.

Sea X la variable aleatoria, definida como la suma de las dos calificaciones. Hallar la función de probabilidad de la variable aleatoria X.

En un distrito de Lima el número de hijos por familia es una variable aleatoria con la siguiente función de probabilidad:

$$P(X = x) = \begin{cases} 0.5k, & \text{si } x \in \{0, 1\} \\ k, & \text{si } x \in \{2, 3\} \\ 2.0k, & \text{si } x = 4 \\ 0, & \text{otra manera.} \end{cases}$$

a. Halle el valor de k para que $f(x)$ sea una función de probabilidad.

b. Si se escoge al azar una familia, ¿cuál es la probabilidad de que tenga por lo menos dos hijos?

c. Si se escoge al azar una familia con al menos un hijo, ¿cuál es la probabilidad de que tenga 3 como máximo?

Ejercicio

Un dispositivo electrónico está compuesto por tres elementos independientes: A, B y C. Cada uno puede funcionar correctamente o fallar durante una prueba. La probabilidad de que cada elemento falle es 0.1

a. Halle la función de probabilidad de la variable aleatoria X: número de elementos que fallan en una prueba

b. ¿Cuál es la probabilidad de que falle al menos un elemento en una prueba?

Medidas estadísticas de una variable aleatoria discreta

- La media o valor esperado de una variable aleatoria discreta X, con función de probabilidad $f(x)$, está dada por:

$$\mu_X = E(X) = \sum_{x \in R_x} x f(x)$$

- La variancia de una v.a.d. X, con f.d.p. $f(x)$, está dada por:

$$\sigma_X^2 = Var(X) = E[(X - \mu_X)^2] = E(X^2) - (E(X))^2 = E(X^2) - \mu_X^2$$

donde $E(X^2) = \sum_{x \in R_x} x^2 f(x)$

- La desviación estándar para una v.a. X:

$$\sigma = \sqrt{\sigma^2}$$

- El coeficiente de variación de una v.a. X:

$$CV(X) = \frac{\sigma}{\mu} \times 100\%$$

Ejemplo

Una agencia de colocación de empleo a determinado que el número de solicitudes de empleo que debe enviar a una empresa hasta ser aceptado el postulante, es una variable aleatoria con la siguiente función de probabilidad:

x	1	2	3	4	5	6
$f(x)$	1/8	1/8	1/8	1/8	2/8	2/8

a. Hallar el número esperado del número de solicitudes de empleo

$$\mu_X = E(X) = 1 \times \frac{1}{8} + 2 \times \frac{1}{8} + 3 \times \frac{1}{8} + 4 \times \frac{1}{8} + 5 \times \frac{2}{8} + 6 \times \frac{2}{8} = 4$$

b. Hallar la variancia del número de solicitudes de empleo

$$E(X^2) = 1^2 \times \frac{1}{8} + 2^2 \times \frac{1}{8} + 3^2 \times \frac{1}{8} + 4^2 \times \frac{1}{8} + 5^2 \times \frac{2}{8} + 6^2 \times \frac{2}{8} = 19$$

$$\sigma_X^2 = Var(X) = E(X^2) - \mu_X^2 = 19 - 4^2 = 3$$

c. Hallar el coeficiente de variación

$$CV(X) = \frac{\sigma}{\mu} \times 100\% = \frac{\sqrt{3}}{4} \times 100\% = 43.3\%$$

d. Si la agencia cobra al postulante por cada solicitud que envía para solicitar empleo 80 soles más 10 por cargas administrativas. Hallar el cobro esperado y el coeficiente de variación del número de solicitudes de empleo.

Sea: Y = Cobro por solicitud enviada

$$Y = 80X + 10$$

$$\mu_Y = E(Y) = 80E(X) + 10 = 80 \times 4 + 10 = 330.0$$

$$\sigma_Y = 80\sigma_X = 80\sqrt{3} = 138.56$$

$$CV_Y = (138.56/330.0) \times 100\% = 42.0\%$$

e. Hallar la probabilidad que el número de solicitudes de empleo enviadas sea a los más 5, si se sabe que ya se enviaron más de 2

Recordando la función de probabilidad de la variable aleatoria:

x	1	2	3	4	5	6
$f(x)$	1/8	1/8	1/8	1/8	2/8	2/8

$$P(X \leq 5|X > 2) = \frac{P(2 < X \leq 5)}{P(X > 2)} = \frac{P(3 \leq X \leq 5)}{1 - P(X \leq 2)}$$

$$P(X \leq 5|X > 2) = \frac{1/8 + 1/8 + 2/8}{1 - 1/8 - 1/8} = 0.667$$

Ejercicio

Sea la v.a. X: número de reclamos por día en una compañía de telefonía móvil, cuya función de probabilidad es:

x	8	12	18	24	32
$P(X = x)$	0.32	0.25	0.20	0.15	0.08

a. Hallar la probabilidad de que en un día se tenga más de 18 reclamos.

b. Calcule e interprete la media, la desviación estándar y el coeficiente de variación.

c. Si por cada reclamo la empresa tiene un costo de \$3.5, al cual se añade un costo fijo de \$0.50 por gastos administrativos. Halle el cv del costo por reclamo.

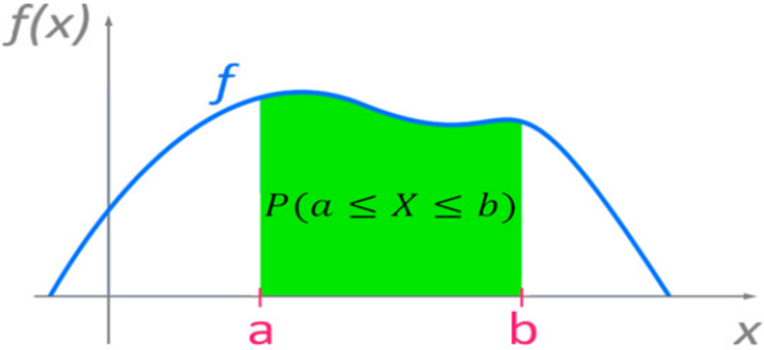
d. Si se sabe que un día se registraron más de 15 reclamos, ¿cuál es el valor esperado del número de reclamos? ¿Y cuál es su desviación estándar?

Variable aleatoria continua

Función de densidad

- Describe la distribución de una variable aleatoria continua.
- No da probabilidades directas, sino que su integral en un intervalo representa la probabilidad.
- Una función $f(x)$ es una densidad si cumple:

1. $f(x) \geq 0, \quad \forall x \in \mathbb{R}$
2. $\int_{-\infty}^{\infty} f(x)dx = 1$
3. $P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b) = \int_a^b f(x)dx$


f(x)
f
$P(a \leq X \leq b)$
a
b
x

Ejemplo

Suponga que el tiempo, en horas, que necesita un técnico para reparar cierta avería de un artefacto eléctrico es una variable aleatoria que tiene la siguiente función de densidad:

$$f(x) = \begin{cases} kx, & 1 \leq x \leq 3 \\ 0, & \text{para otros valores de } x \end{cases}$$

a. Halle el valor de k para que $f(x)$ sea una función de densidad

   Por propiedad, $\int_{-\infty}^{\infty} f(x)dx = 1$. Luego:

$$\int_1^3 kxdx = k\int_1^3 xdx = k\frac{x^2}{2}|_1^3 = k \times \frac{1}{2}(3^2 - 1^2) = 1 \rightarrow k = \frac{1}{4}$$

```r
fx0 = function(x){x}
integrate(fx0, 1,3)$value
```

```
[1] 4
```

Entonces la función de densidad es:

$$f(x) = \begin{cases} \frac{x}{4}, & 1 \leq x \leq 3 \\ 0, & \text{para otros valores de } x \end{cases}$$

Verificando:

```r
fx = function(x){x/4}
integrate(fx, 1,3)$value
```

```
[1] 1
```

b. ¿Qué porcentaje de reparaciones tiene un tiempo menor a 2.4 horas?

$$P(X < 2.4) = \int_{1}^{2.4} \frac{x}{4} dx = \frac{x^2}{8}|_1^{2.4} = \frac{1}{8}(2.4^2 - 1^2) = 0.595 = 59.5\%$$

```
integrate(fx, 1.0, 2.4)$value
```

```
[1] 0.595
```

c. Calcule la probabilidad de que un técnico demore por lo menos 1.8 horas pero menos de 2 horas en reparar un artefacto eléctrico.

$$P(1.8 \leq X < 2) = \int_{1.8}^{2} \frac{x}{4} dx = \frac{x^2}{8}|_{1.8}^{2} = \frac{1}{8}(2^2 - 1.8^2) = 0.095$$

```
integrate(fx, 1.8, 2.0)$value
```

```
[1] 0.095
```

d. ¿Cuál es el tiempo máximo que necesita un técnico para estar dentro del 18% más rápido en reparar un artefacto eléctrico?

Se tiene $P(X \leq k) = 0.18$

$$P(X \leq k) = \int_1^k \frac{x}{4} dx = \frac{x^2}{8}|_1^k = \frac{1}{8}(k^- 1^2) = 0.18$$

Despejando: $k = 1.56$ horas

```
area <- function(k) {integrate(fx, 1, k)$value - 0.18}
uniroot(area, lower = 1, upper = 3)$root
```

```
[1] 1.562041
```

Ejercicio Una empresa fabrica un tipo de artículo cuya producción está sujeta a pequeñas variaciones de peso. Según estudios de control de calidad, el peso (en kg) de los artículos sigue la siguiente distribución:

$$f(x) = \begin{cases} \frac{2}{3}x, & 1 \leq x \leq 2 \\ 0, & \text{para otros valores de } x \end{cases}$$

a. Verifique que f(x) es una función de densidad
b. ¿Cuál es la probabilidad de que un artículo pese entre 1.5 kg y 2 kg?
c. ¿Qué porcentaje de artículos pesa menos de 1.8 kg?
d. Dado que un artículo pesa al menos 1.2 kg, ¿cuál es la probabilidad de que pese menos de 1.6 kg?
e. ¿Cuál debe ser el peso mínimo para que un artículo esté considerado dentro del 12.5% de los que más pesan?
f. ¿Cuál debe ser el peso máximo para que un artículo esté considerado dentro del 14.5% de los que menos pesan?

Medidas estadísticas de una V.A. continua.

Media

La media, valor esperado o esperanza matemática de una v.a. X, que tiene como función de densidad $f(x)$ está dada por:

$$\mu_X = E(X) = \int_{-\infty}^{\infty} x f(x) dx$$

Variancia

La variancia de una v.a X está dada por el valor esperado de la función $h(x) = (x - \mu_x)^2$, es decir:

$$\sigma_X^2 = Var(X) = E[(X - \mu_X)^2] = E(X^2) - (E(X))^2 = E(X^2) - \mu_X^2$$

donde $E(X^2) = \int_{-\infty}^{\infty} x^2 f(x) dx$

La v.a. continua X definida como la proporción de accidentes fatales por mes que ocurren en determinada ciudad tiene como función de densidad:

$$f(x) = \begin{cases} 2x, & 0 \leq x \leq 1 \\ 0, & \text{para otros valores de } x \end{cases}$$

a. Calcule la probabilidad de que la proporción de accidentes por mes sea menor a 0.4

$$P(X < 0.4) = P(0 \leq X < 0.4) = \int_0^{0.4} 2x dx = 2\frac{x^2}{2}|_0^{0.4} = (0.4^2 - 0^2) = 0.16$$

```
fx = function(x){2*x}
integrate(fx, 0, 0.4)$value
```

```
[1] 0.16
```

b. Encuentre la proporción media de accidentes fatales por mes en esa ciudad

$$\mu_X = E(X) = \int_{-\infty}^{\infty} x \times 2x dx = \int_0^1 2x^2 dx = 2\frac{x^3}{3}|_0^1 = \frac{2}{3}$$

```
fEx = function(x){x*2*x}
(Ex = integrate(fEx, 0, 1)$value)
```

```
[1] 0.6666667
```

c. Halle la variancia y el coeficiente de variación de la proporción de accidentes fatales por mes en esa ciudad.

$$E(X^2) = \int_{-\infty}^{\infty} x^2 \times 2x dx = \int_0^1 2x^3 dx = 2\frac{x^4}{4}|_0^1 = \frac{1}{2}$$

```
fEx2 = function(x){x**2*2*x}
(Ex2 = integrate(fEx2, 0, 1)$value)
```

```
[1] 0.5
```

$$\sigma_X^2 = E(X^2) - (E(X))^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{18}$$

```
(VX = Ex2 - Ex**2)
```

```
[1] 0.05555556
```

$$CV(X) = \frac{\sigma_X}{\mu_X} \times 100\% = \frac{\sqrt{1/18}}{2/3} \times 100\% = 35.4\%$$

Ejercicio

En un centro de inspección vehicular, se ha determinado que el tiempo (en minutos) que demora un automóvil en pasar la revisión técnica sigue una distribución continua, cuya función de densidad es:

$$f(x) = \begin{cases} \frac{x}{600}, & 20 \leq x \leq 40 \\ 0, & \text{para otros valores de } x \end{cases}$$

a. ¿Cuál es la probabilidad de que un automóvil termine su revisión en menos de 32 minutos?

b. Si un automóvil ya lleva 30 minutos en revisión, ¿cuál es la probabilidad de que termine en menos de 5 minutos adicionales?

c. ¿Qué porcentaje de vehículos termina la revisión en un tiempo entre 25 y 35 minutos?

d. Halle la media y el coeficiente de variación del tiempo de revisión técnica.

e. Con la implementación de dos casetas adicionales, se estima que el tiempo promedio disminuirá en un 12.5%. ¿Cuál será el nuevo tiempo promedio de revisión?

f. Determine la mediana del tiempo de revisión.