Tema 4. Probabilidad y variables aleatorias

En este tema:

- Probabilidad:
  - Experimentos aleatorios, espacio muestral, sucesos.
  - Interpretaciones de la probabilidad.
  - Propiedades de la probabilidad.
  - Probabilidad condicionada y teorema de Bayes.
- Variables aleatorias:
  - Concepto de variable aleatoria.
  - Variables aleatorias discretas.
  - Variables aleatorias continuas.
  - Esperanza, varianza y desviación típica.
- Modelos de variables aleatorias

  ...

Tema 4. Probabilidad y variables aleatorias

En este tema:

- Probabilidad

  ...
- Variables aleatorias

  ...
- Modelos de variables aleatorias:
  - Distribución Bernoulli
  - Distribución Binomial
  - Distribución de Poisson
  - Distribución uniforme
  - Distribución exponencial
  - Distribución normal
  - Distribuciones asociadas a la normal

Conceptos básicos

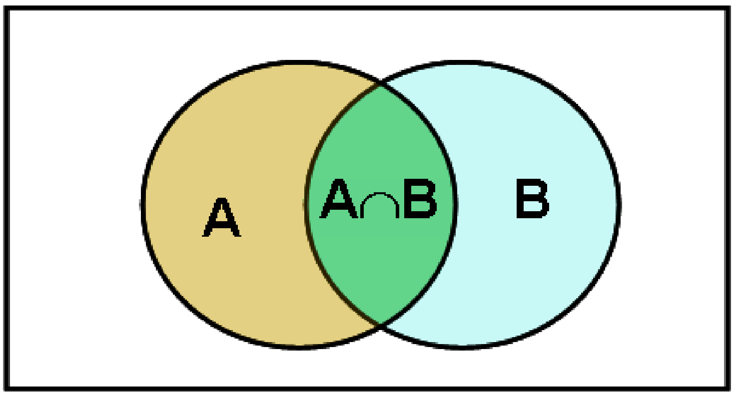
- Experimento aleatorio: proceso de observar un fenómeno del que se conocen de antemano todos sus posibles resultados, pero a partir de las condiciones iniciales no puede predecirse exactamente cuál de estos resultados se producirá.
- Espacio muestral: es el conjunto de todos los posibles resultados de un experimento aleatorio. Se denota por $\Omega = \{e_1, e_2, \ldots, e_n, \ldots\}$ y cada uno de sus elementos se denomina suceso elemental o punto muestral.
- Un espacio muestral (correspondiente a un determinado experimento aleatorio) tiene asociada una colección $\mathcal{F}$ no varía de subconjuntos de Ω. Los elementos de $\mathcal{F}$ se denominan sucesos y se denotan por las letras $A, B, C, \ldots$.

Ejemplo:

El espacio muestral correspondiente al experimento aleatorio *puntuación obtenida al lanzar un dado* es $\Omega = \{1, 2, 3, 4, 5, 6\}$. Podemos considerar los sucesos $A =$"obtener una puntuación par" y $B =$"obtener una puntuación superior a 3". Entonces, $A = \{2, 4, 6\}$ y $B = \{4, 5, 6\}$.

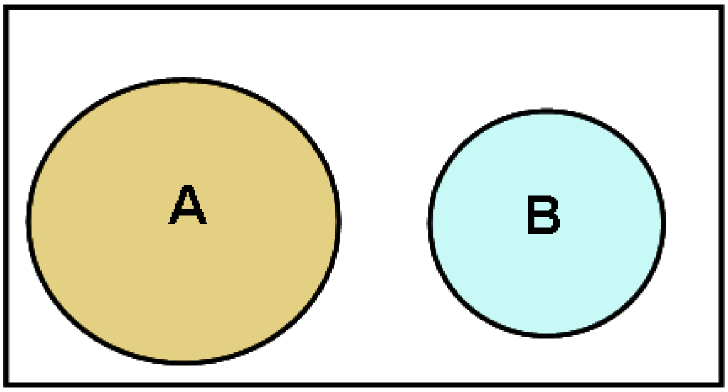
Sucesos: conceptos básicos

Intersección de sucesos: Si A y B son dos sucesos del espacio muestral Ω, entonces la intersección, $A \cap B$, es el conjunto de todos los sucesos de Ω que están en A y en B.



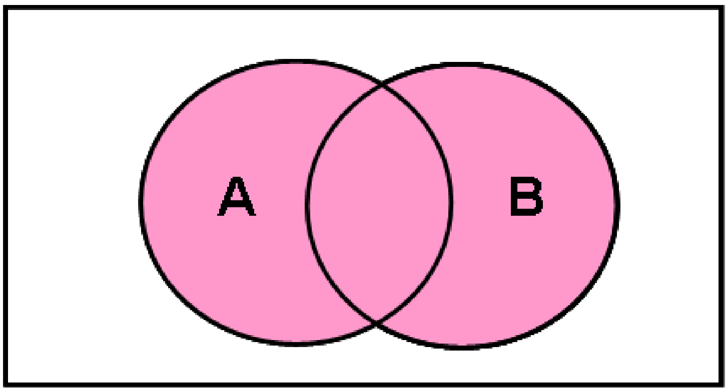
Sucesos: conceptos básicos

A y B son sucesos incompatibles si no tienen ningún suceso elemental en común i.e., el conjunto $A \cap B$ es vacío



Sucesos: conceptos básicos

Unión de sucesos: Si A y B son dos sucesos de un espacio muestral Ω, entonces la unión, $A \cup B$, es el conjunto de todos los sucesos de Ω que pertenecen a cualquiera de los dos, A ó B.



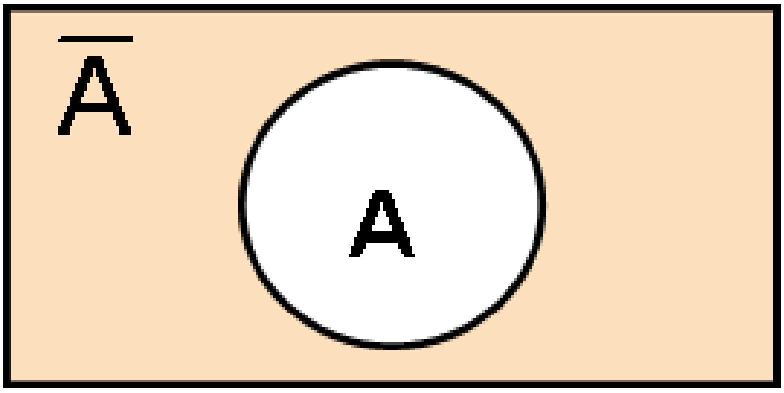
Sucesos: conceptos básicos

Sucesos triviales:

- Suceso seguro Ω: conjunto $=$ espacio muestral
- Suceso imposible $\emptyset$: conjunto $=$ conjunto vacío

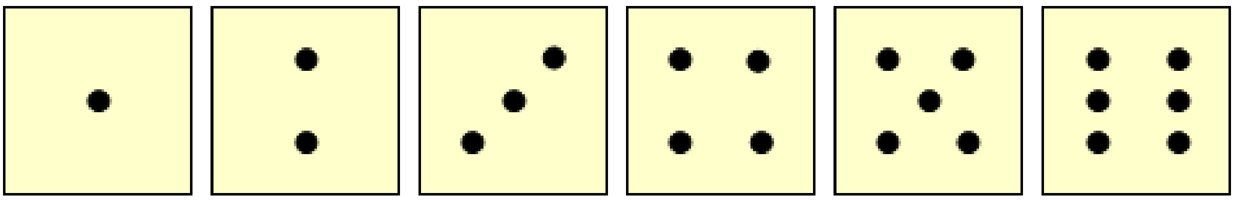
Complementario

El complementario de un suceso A es el conjunto de todos los sucesos elementales de Ω que no están en A.



Ejemplo: lanzamiento de un dado

Consideremos el experimento aleatorio "resultado observado al lanzar un dado":



- suceso elemental: el 1, el 2, el 3, el 4, el 5, el 6
- espacio muestral: $\Omega = \{1, 2, 3, 4, 5, 6\}$
- suceso: $A = \{2, 4, 6\}$ $\qquad B = \{4, 5, 6\}$

El suceso A es "sale un número par".
El suceso B es "sale un número mayor que tres".

Ejemplo: lanzamiento de un dado

$$\Omega = \{1,2,3,4,5,6\} \qquad A = \{2,4,6\} \qquad B = \{4,5,6\}$$

- Complementario:
$$\bar{A} = \{1,3,5\} \qquad \bar{B} = \{1,2,3\}$$
- Intersección:
$$A \cap B = \{4,6\} \qquad \bar{A} \cap \bar{B} = \{1,3\}$$
- Unión:
$$A \cup B = \{2,4,5,6\} \qquad \bar{A} \cup \bar{B} = \{1,2,3,5\}$$
$$A \cup \bar{A} = \{1,2,3,4,5,6\} = \Omega$$
- Sucesos incompatibles:
$$A \cap \bar{A} = \emptyset$$

Probabilidad. Intuición

La probabilidad es una medida subjetiva sobre la incertidumbre de que suceda cierto suceso.

Al tirar un dado:

- la probabilidad de que salga un 1 es más pequeña que la probabilidad de que salga un número mayor que uno
- la probabilidad de que salga un 4 es igual que la probabilidad de que salga un 6.
- la probabilidad de que salga un 7 es cero
- la probabilidad de que salga un número positivo es uno

Tres enfoques/interpretaciones

Probabilidad clásica (regla de Laplace): Sea Ω el espacio muestral asociado a cierto experimento aleatorio formado por $n(\Omega) < \infty$ puntos muestrales equiprobables (igualmente probables). Si A es un suceso formado por $n(A)$ puntos muestrales, entonces se define la probabilidad de A como

$$P(A) = \frac{n(A)}{n(\Omega)}.$$

Enfoque frecuentista: Si repetiéramos el experimento muchas veces, la frecuencia con que ocurre el suceso sería una aproximación de la probabilidad.

$\boxed{\text{Probabilidad como el valor límite de la frecuencia}}$

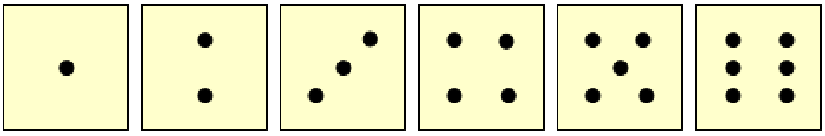
Probabilidad subjetiva: Depende de la información que tengamos en ese momento.

$\boxed{\text{Probabilidad como creencia o certeza de que ocurra}}$

Propiedades de la probabilidad

- La probabilidad es una aplicación $P : \mathcal{F} \to [0, 1]$, que asigna a cada suceso $A \in \mathcal{F}$ un valor numérico $P(A)$.
- $0 \leq P(A) \leq 1$.
- Sea $A = \{e_1, e_2, \ldots, e_n\}$, (recordemos que e_i son los puntos muestrales), entonces $P(A) = \sum_{i=1}^{n} P(e_i)$.
- $P(\Omega) = 1$ y $P(\emptyset) = 0$.
- Complementario: $P(\bar{A}) = 1 - P(A)$.
- Unión: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
- Si A y B son incompatibles, entonces $P(A \cup B) = P(A) + P(B)$.

Ejemplo: lanzamiento de un dado



- Probabilidad de un suceso elemental: $P(e_i) = \frac{1}{6}$
- Probabilidad de que salga par: $A = \{2, 4, 6\}$, luego

$$P(A) = P("2") + P("4") + P("6") = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$$

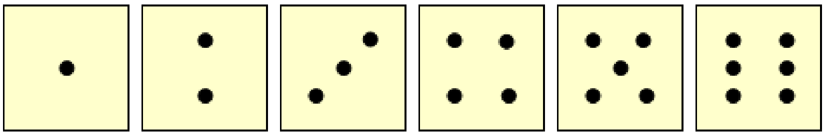
- Probabilidad de que salga mayor que 3: $B = \{4, 5, 6\}$, luego

$$P(B) = P("4") + P("5") + P("6") = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$$

- Probabilidad de que salga impar

$$P(\bar{A}) = 1 - P(A) = 1 - \frac{1}{2} = \frac{1}{2}$$

Ejemplo: lanzamiento de un dado



- Probabilidad de que salga par o mayor que tres

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Como $A \cap B = \{4, 6\}$, entonces $P(A \cap B) = \frac{2}{6} = \frac{1}{3}$

$$P(A \cup B) = \frac{1}{2} + \frac{1}{2} - \frac{1}{3} = \frac{4}{6} = \frac{2}{3}$$

- Probabilidad de que salga par o igual a uno.
Los sucesos $A = \{2, 4, 6\}$ y $C = \{1\}$ son incompatibles ($A \cap C = \emptyset$) por tanto

$$P(A \cup C) = P(A) + P(C) = \frac{1}{2} + \frac{1}{6} = \frac{4}{6} = \frac{2}{3}$$

Ejemplo: probabilidad condicional

Se clasifica un grupo de 100 ejecutivos de acuerdo a su peso y a si sufren o no de hipertensión. La tabla muestra el número de ejecutivos en cada categoría.

	Insuficiente	Normal	Sobrepeso	Total
Hipertenso	2	8	10	20
Normal	20	45	15	80
Total	22	53	25	100

- Si se elige un ejecutivo al azar, ¿cuál es la probabilidad de que tenga hipertensión?
$$P(H) = \frac{20}{100} = 0,2$$
- Si se elige a una persona al azar, y se descubre que tiene sobrepeso, ¿cuál es la probabilidad de que tenga hipertensión? ¿Es la misma que antes?

Ejemplo: probabilidad condicional

Probabilidad de que sea hipertenso, sabiendo que tiene sobrepeso:

$$P(H|S)$$

Para calcularla, nos fijamos sólo en los ejecutivos con sobrepeso:

$$P(H|S) = \frac{10}{25} = 0.4$$

¿Por qué? es como si eligiese la persona al azar sólo entre los que tienen sobrepeso.

La probabilidad condicional, (o probabilidad condicionada) es la probabilidad de que ocurra un evento, dado que otro evento ha ocurrido.

Probabilidad condicional

Probabilidad condicional (o condicionada)
Sean A y B dos sucesos con $P(B) > 0$, se define la probabilidad de A condicionada a B como:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Ley de la multiplicación (fórmula de las probabilidades compuestas)
Es útil para calcular la probabilidad de una intersección de sucesos, a partir de la noción de probabilidad condicionada.
$P(A \cap B) = P(A|B)\, P(B)$, siempre que $P(B > 0)$.
$P(A \cap B \cap C) = P(A)\, P(B|A)\, P(C|A \cap B)$, siempre que $P(A \cap B) > 0$.
Independencia estocástica
Dos sucesos A y B son independientes si $P(A \cap B) = P(A)\, P(B)$.
De forma equivalente, si $P(B) > 0$, A y B son independientes si $P(A|B) = P(A)$. O equivalentemente, si $P(A) > 0$, A y B son independientes si $P(B|A) = P(B)$.

Ejemplo

Se extraen dos cartas de una baraja española. Probabilidad de que:

- la primera carta sea copa: $P(A) = \frac{12}{48}$.
- la segunda sea copa, sabiendo que la primera lo fue: $P(B|A) = \frac{11}{47}$.
- las dos cartas sean copas: $P(A \cap B) = P(B|A)\, P(A) = \frac{11}{47}\,\frac{12}{48}$.

Se lanzan dos dados. Probabilidad de que:

- en el primer dado salga un uno: $P(C) = \frac{1}{6}$.
- en el segundo dado salga un uno, sabiendo que en el primero salió uno: $P(D|C) = P(D) = \frac{1}{6}$.
- en el primer dado salga un uno, si en el segundo salió uno: $P(C|D) = P(C) = \frac{1}{6}$.
- en los dos dados salga uno: $P(C \cap D) = P(D)\, P(C) = \frac{1}{6}\,\frac{1}{6}$ (sucesos independientes)

Ley de la probabilidad total

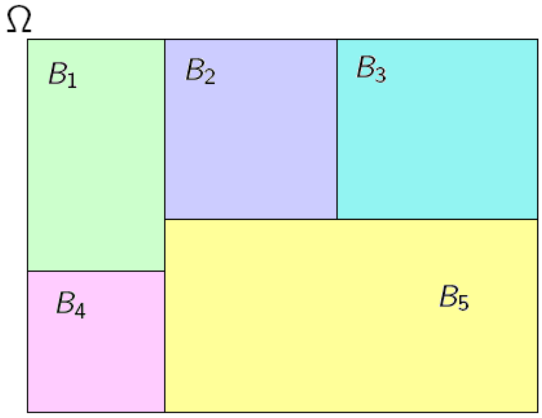
Un conjunto de sucesos $B_1, B_2, \ldots, B_k$ son mútuamente excluyentes si

$$B_i \cap B_j = \emptyset, \qquad \forall i \neq j.$$

Si además cumplen que

$$\Omega = B_1 \cup B_2 \cup \ldots \cup B_k,$$

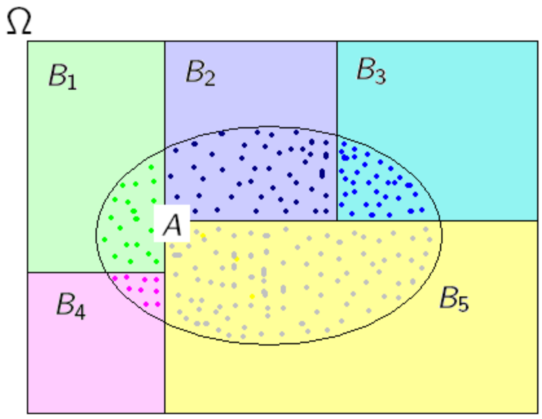
se dice que $B_1, B_2, \ldots, B_k$ forman una partición del espacio muestral.



Ley de probabilidad total

Si $B_1, B_2, \ldots, B_k$ es una partición del espacio muestral tal que $P(B_i) \neq 0$, $i = 1, \ldots, k$, y A es un suceso cualquiera, entonces

$$P(A) = P(A \cap B_1) + P(A \cap B_2) + \ldots + P(A \cap B_k) =$$
$$= P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \ldots + P(A|B_k)P(B_k).$$



Ejemplo: probabilidad total

En una fábrica se embalan galletas en cuatro cadenas de montaje: A_1, A_2, A_3, y A_4. El 35% de la producción total se embala en la cadena A_1, el 20%, 24% y 21% en las cadenas A_2, A_3, y A_4, respectivamente.
Los datos indican que no se embalan correctamente un porcentaje pequeño de las cajas: el 1% en la cadena de montaje A_1, el 3% en A_2, el 2.5% en A_3 y el 2% en A_4.
¿Cuál es la probabilidad de que una caja elegida al azar de la producción total sea defectuosa (suceso D)?

$$
\begin{aligned}
P(D) &= P(D \cap A_1) + P(D \cap A_2) + P(D \cap A_3) + P(D \cap A_4) \\
&= P(D|A_1)P(A_1) + P(D|A_2)P(A_2) + P(D|A_3)P(A_3) + P(D|A_4)P(A_4) \\
&= 0.01 \times 0.35 + 0.03 \times 0.20 + 0.025 \times 0.24 + 0.02 \times 0.21 = 0.0197.
\end{aligned}
$$

Inversión de las condiciones: Teorema de Bayes

Para dos sucesos A y B se tiene que

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Ejemplo: (continuación del anterior) Supongamos que descubrimos una caja defectuosa, ¿cuál es la probabilidad de que la caja haya sido embalada en la cadena de montaje A_1?

$$P(A_1|D) = \frac{P(D|A_1)P(A_1)}{P(D)} = \frac{0.01 \times 0.35}{0.0197} = 0.17766$$

Variables aleatorias

- Variable aleatoria.
- Variables discretas:
  - Función de probabilidad (f. de masa)
  - Función de distribución
- Variables continuas:
  - Función de densidad
  - Función de distribución
- Esperanza, varianza.

Variables aleatorias

Sea Ω el espacio muestral asociado a cierto experimento aleatorio y $\mathcal{F}$ el correspondiente conjunto de sucesos.

Se denomina **variable aletoria** (v.a.) a una función $X : \Omega \to \mathbb{R}$, que a cada elemento $e_i \in \Omega$ le asigna un valor numérico $X(e_i) = x_i \in \mathbb{R}$.

Intuitivamente, una variable aleatoria es una medida o cantidad que varía en función del resultado concreto e_i que se observa al realizar el experimento aleatorio.

La v.a. se denota con letras mayúsculas, mientras que las letras minúsculas indican el valor concreto que toma la v.a. cuando se evalúa en un punto muestral.

Ejemplo:
Lanzar un dado una vez. Considerar la v.a. $X =$"resultado de la tirada".
¿Cuántos sucesos elementales hay? ¿Qué valores puede tomar X?

Variables aleatorias

V.a. discreta

Si X toma valores sobre un conjunto $S \subseteq \mathbb{R}$ finito o infinito numerable, se dice que X es una variable aleatoria discreta.

V.a. continua

Si X toma valores sobre un conjunto $S \subseteq \mathbb{R}$ infinito no numerable (por ejemplo, en intervalo o una unión de intervalos), se dice que X es una variable aleatoria continua.

Ejemplos

- "resultado al tirar un dado", "número de coches que pasan por cierto peaje en una semana" son v.a. discretas. En el primer caso $S = \{1, 2, 3, 4, 5, 6\}$ es finito, mientras que en el segundo caso, $S = \{0, 1, 2, \ldots\} = \mathbb{N} \cup \{0\}$ es infinito numerable.
- "altura de una persona", "el tiempo de reacción a cierto medicamento" son v.a. continuas. En ambos casos puede considerarse $S = [0, +\infty)$.

Variables aleatorias discretas

Función de probabilidad

Sea X una variable aleatoria discreta que toma valores en el conjunto $S = \{x_1, x_2, \ldots\}$, finito o infinito numerable, con probabilidades $p_1 = P(X = x_1)$, $p_2 = P(X = x_2)$, $\ldots$.
Se define la función de probabilidad de X o función de masa de X como

$$P(X = x) = \begin{cases} p_i, & \text{si } x = x_i \in S, \\ 0, & \text{si } x \notin S. \end{cases}$$

Ejemplo

$X =$"resultado de lanzar un dado". La función de probabilidad es

x	1	2	3	4	5	6
$P(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Variables aleatorias discretas

Función de probabilidad. Propiedades

X variable aleatoria discreta que toma valores en el conjunto $S = \{x_1, x_2, \ldots\}$ con probabilidades $p_1 = P(X = x_1)$, $p_2 = P(X = x_2)$, ...

- $0 \leq P(X = x) \leq 1$, para todo $x \in \mathbb{R}$.
- $\displaystyle\sum_{x \in S} P(X = x) = \sum_i P(X = x_i) = \sum_i p_i = 1.$
- $\displaystyle P(X \in A) = \sum_{x \in A} P(X = x).$
- $\displaystyle P(X \leq x) = P(X \in (-\infty, x]) = \sum_{i, x_i \leq x} P(X = x_i) = \sum_{i, x_i \leq x} p_i.$
- $P(X > x) = 1 - P(X \leq x)$.

Variables aleatorias discretas

Función de distribución

La función de distribución o función de probabilidad acumulada de una variable aleatoria X es una aplicación $F : \mathbb{R} \to [0, 1]$, que a cada valor $x \in \mathbb{R}$ le asigna la probabilidad $F(x) = P(X \leq x) = P(X \in (-\infty, x])$.

Atención! $F(x)$ está definida para todo $x \in \mathbb{R}$ y no sólo para los $x \in S$.

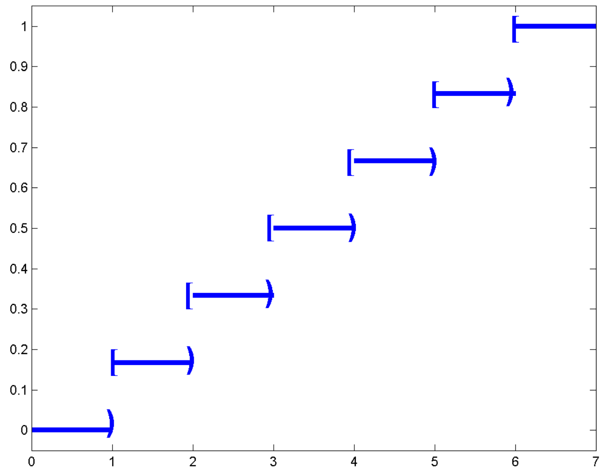
Propiedades

- $0 \leq F(x) \leq 1$ para todo $x \in \mathbb{R}$.
- $F(y) = 0$ para todo $y < \min S$. Por tanto, $F(-\infty) = 0$.
- $F(y) = 1$ para todo $y \geq \max S$. Por tanto, $F(\infty) = 1$.
- Si $x_1 \leq x_2$, entonces $F(x_1) \leq F(x_2)$, es decir, $F(x)$ es no decreciente.
- Para todo $a, b \in \mathbb{R}$, $P(a < X \leq b) = P(X \in (a, b]) = P(X \in (-\infty, b]) - P(X \in (-\infty, a]) = F(b) - F(a)$.

Variables aleatorias discretas

Ejemplo $X =$"resultado de lanzar un dado". La función de distribución es

$$F(x) = \begin{cases} 0, & \text{si } x < 1, \\ 1/6, & \text{si } 1 \leq x < 2, \\ 2/6, & \text{si } 2 \leq x < 3, \\ 3/6, & \text{si } 3 \leq x < 4, \\ 5/6, & \text{si } 4 \leq x < 5, \\ 1, & \text{si } x \geq 6. \end{cases}$$



Si X es una v.a. discreta, su función de distribución es de tipo escalón (discontinuidades de salto). Cada escalón corresponde a un $x_i \in S$ y el salto correspondiente es la probabilidad $P(X = x_i) = p_i$.

Variables aleatorias continuas

Función de densidad

Las probabilidades de una variable aleatoria continua se calculan a partir de una función $f : \mathbb{R} \to [0, +\infty)$ denominada función de densidad. Esta función cumple las propiedades siguientes:

Propiedades

- $f(x) \geq 0$ para todo $x \in \mathbb{R}$.
- $\int_{-\infty}^{\infty} f(x)\, dx = 1$, es decir, el área total de la función de densidad es 1.
- Para todo $a, b \in \mathbb{R}$, $P(a \leq X \leq b) = P(X \in [a, b]) = \int_a^b f(x)\, dx$ es el área que determina la función de densidad de X sobre el intervalo $[a, b]$.
- Los intervalos $[a, b]$, (a, b), $(a, b]$ y $[a, b)$ tienen la misma probabilidad.

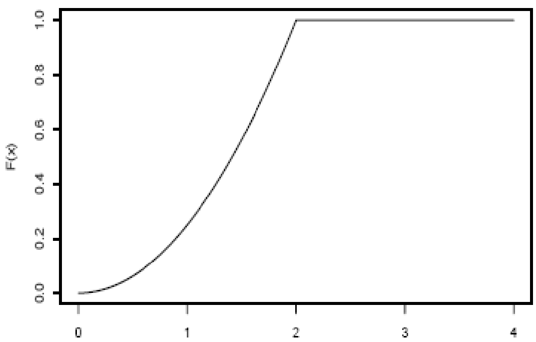
Atención! La función de densidad juega el mismo papel que la función de probabilidad para v.a. discretas. Pero, en el caso continuo, solamente tiene sentido calcular probabilidades de intervalos, puesto que $P(X = x) = 0$ para todo $x \in R$.

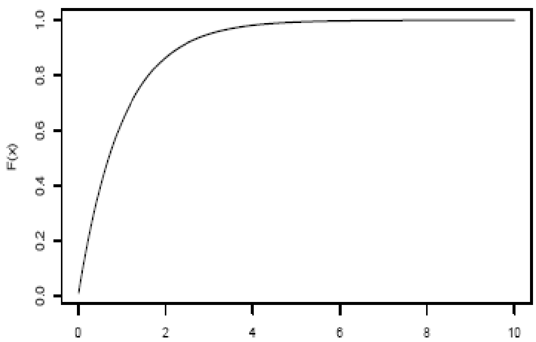
Variables aleatorias continuas

Función de distribución

Para una v.a. continua X, la función de distribución se define como la función $F(x) = P(X \leq x) = P(X \in (-\infty, x]) = \int_{-\infty}^{x} f(t)\, dt$, para todo $x \in \mathbb{R}$.

Igual que en el caso discreto, la función $F(x)$ da las probabilidades acumuladas hasta el punto $x \in \mathbb{R}$, pero ahora se trata de una función continua y no de tipo escalón. Dos ejemplos son:





Variables aleatorias continuas

Propiedades

- $0 \leq F(x) \leq 1$, para todo $x \in \mathbb{R}$.
- $F(-\infty) = 0$.
- $F(\infty) = 1$.
- Si $x_1 \leq x_2$, entonces $F(x_1) \leq F(x_2)$, es decir, $F(x)$ es no decreciente.
- Para todo $a, b \in \mathbb{R}$, $P(a \leq X \leq b) = F(b) - F(a)$.
- La función de densidad de X se obtiene derivando la función de distribución, es decir, $f(x) = F'(x)$.

Variables aleatorias continuas

Ejemplo

Una variable aleatoria X tiene función de densidad

$$f(x) = \begin{cases} 3\,x^2, & \text{si } x \in (0,1), \\ 0, & \text{si } x \notin (0,1) \end{cases}$$

¿Cómo es la gráfica de la función de densidad de X?

Indicar cuál es el área asociada a la probabilidad $P(X > 1/2)$.

Calcular la probabilidad $P(X > 1/2)$.

Obtener la función de distribución de X.

Variables aleatorias continuas

Ejemplo

Una variable aleatoria X tiene función de densidad

$$f(x) = \begin{cases} 12x^2(1-x), & \text{si } 0 < x < 1, \\ 0, & \text{en otro caso.} \end{cases}$$

$$P(X \leq 0.5) = \int_{-\infty}^{0.5} f(u)du = \int_{0}^{0.5} 12u^2(1-u)du = 0.3125$$

$$P(0.2 \leq X \leq 0.5) = \int_{0.2}^{0.5} f(u)du = \int_{0.2}^{0.5} 12u^2(1-u)du = 0.2853$$

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)du = \begin{cases} 0, & \text{si } x \leq 0, \\ 12\left(\frac{x^3}{3} - \frac{x^4}{4}\right), & \text{si } 0 < x \leq 1, \\ 1, & \text{si } x > 1. \end{cases}$$

Variables aleatorias continuas

Ejemplo. Solución con R/Rcommander

Se define la función a integrar:

```
integrando <- function(x) {12 * x^2 * (1-x)}
```

Se define una función de R que hace integración numérica:

```
mi.integrate <- function(foo, a, b, ...  )
   integrate(function(asdf) foo(asdf,...), a, b)
```

Se calculan las integrales:

```
mi.integrate(integrando, 0, 0.5)
mi.integrate(integrando, 0.2, 0.5)
```


Esperanza y varianza de una variable aleatoria

Sea X una v.a. discreta que toma valores en $S = \{x_1, x_2, \ldots\}$ con probabilidades $p_1 = P(X = x_1)$, $p_2 = P(X = x_2)$, $\ldots$. Entonces:

$$\mathrm{E(X)} = \sum_{x \in S} x\, P(X = x) = \sum_i x_i\, P(X = x_i) = \sum_i x_i\, p_i$$

$$\mathrm{var(X)} = E[(X - E(X))^2] = \sum_{x \in S} (x - E(X))^2\ P(X = x)$$

$$= \sum_i (x_i - E(X))^2\ p_i = \sum_i x_i^2\, p_i - E(X)^2$$

Sea X una v.a. continua que toma valores en $S \subseteq \mathbb{R}$ con función de densidad $f(x)$. Entonces:

$$\mathrm{E(X)} = \int_S x\, f(x)\, dx$$

$$\mathrm{var(X)} = E[(X - E(X))^2] = \int_S (x - E(X))^2\ f(x)\, dx$$

$$= \int_S x^2\, f(x)\, dx - E(X)^2$$

Desigualdad de Chebyschev

Este resultado es útil para estimar una probabilidad cuando se desconoce la distribución de probabilidad (o ley) de una v.a. X.

Si X es una v.a. con esperanza y varianza finitas, entonces para todo $k \geq 1$

$$P(|X - E(X)| \geq k) \leq \frac{var(X)}{k^2},$$

o equivalentemente,

$$P(|X - E(X)| < k) \geq 1 - \frac{var(X)}{k^2}.$$

Atención! La cota que proporciona la desigualdad de Chebyschev es "demasiado gruesa" y sólo debe utilizarse cuando no se disponga de la ley de la v.a. X.

Esperanza y varianza

Ejemplo

$X =$"resultado de lanzar un dado". La función de probabilidad es

x	1	2	3	4	5	6
$P(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

El conjunto S donde X toma valores es $S = \{1, 2, 3, 4, 5, 6\}$.

Calculamos su esperanza:

$$E(X) = \sum_{x \in S} x\, P(X = x)$$
$$= 1 \cdot \tfrac{1}{6} + 2 \cdot \tfrac{1}{6} + 3 \cdot \tfrac{1}{6} + 4 \cdot \tfrac{1}{6} + 5 \cdot \tfrac{1}{6} + 6 \cdot \tfrac{1}{6}$$
$$= \tfrac{1+2+3+4+5+6}{6} = \tfrac{21}{6} = 3.5$$

Esperanza y varianza

Ejemplo

x	1	2	3	4	5	6
$P(X=x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Calculamos su varianza:

$$var(X) = E[(X - E(X))^2]$$

$$= \sum_{x\in S} (x - E(X))^2 \; P(X = x)$$

$$= (1 - 3.5)^2 \cdot \frac{1}{6} + (2 - 3.5)^2 \cdot \frac{1}{6} + (3 - 3.5)^2 \cdot \frac{1}{6}$$

$$+(4 - 3.5)^2 \cdot \frac{1}{6} + (5 - 3.5)^2 \cdot \frac{1}{6} + (6 - 3.5)^2 \cdot \frac{1}{6}$$

$$= \frac{(-2.5)^2+(-1.5)^2+\ldots+2.5^2}{6} = \frac{17.5}{6} = 2.9167$$

Esperanza y varianza

Ejemplo

x	1	2	3	4	5	6
$P(X=x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Otra forma de calcular la varianza:

$$var(X) = E(X^2) - (E(X))^2$$

$$= \sum_{x\in S} x^2 \, P(X = x) - E(X)^2 =$$

$$= (1)^2 \cdot \frac{1}{6} + (2)^2 \cdot \frac{1}{6} + (3)^2 \cdot \frac{1}{6} + (4)^2 \cdot \frac{1}{6} + (5)^2 \cdot \frac{1}{6} + (6)^2 \cdot \frac{1}{6} - 3.5^2$$

$$= \frac{1^2+2^2+\ldots+6^2}{6} - 3.5^2 = 2.9167$$

Esperanza y varianza

Ejemplo

$X =$"número de caras al tirar una moneda dos veces".

El espacio muestral asociado al experimento aleatorio "lanzamiento de dos monedas" es $\Omega = \{(cara, cara), (cara, cruz), (cruz, cara), (cruz, cruz)\}$. La variable X toma valores en $S = \{0, 1, 2\}$ con probabilidades $P(X = 0) = 1/4$, $P(X = 1) = 1/4 + 1/4 = 1/2$, $P(X = 2) = 1/4$.

Por tanto, la función de probabilidad de X es

x	0	1	2
$P(X = x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Calculamos su esperanza y varianza:
$E(X) = \sum_{x \in S} x\, P(X = x) = 0\, \frac{1}{4} + 1\, \frac{1}{2} + 2\, \frac{1}{4} = 1$,
$var(X) = \sum_{x \in S} (x - E(X))^2\ P(X = x) =$
$(0 - 1)^2\, \frac{1}{4} + (1 - 1)^2\, \frac{1}{2} + (2 - 1)^2\, \frac{1}{4} = \frac{1}{2}$.

Esperanza y varianza

Ejemplo

Una variable aleatoria X tiene función de densidad

$$f(x) = \begin{cases} 12x^2(1 - x), & \text{si } 0 < x < 1, \\ 0, & \text{en otro caso.} \end{cases}$$

Calculamos su esperanza:

$$E(X) = \int_{\mathbb{R}} x \cdot f(x) dx$$
$$= \int_0^1 x \cdot 12x^2(1 - x) dx$$
$$= \int_0^1 \left(12(x^3 - x^4)\right) dx = 12 \left(\frac{1}{4}x^4 - \frac{1}{5}x^5\right) \Big|_0^1$$
$$= 12 \left(\frac{1}{4} - \frac{1}{5}\right) = \frac{3}{5}$$

Esperanza y varianza

Ejemplo

$$f(x) = \begin{cases} 12x^2(1-x), & \text{si } 0 < x < 1, \\ 0, & \text{en otro caso.} \end{cases}$$

Calculamos su varianza:

$$\begin{aligned}
& var(X) = E[(X - E(X))^2] \\
& = \int_{\mathbb{R}} (x - E(X))^2 \cdot f(x) dx \\
& = \int_0^1 \left(x - \tfrac{3}{5}\right)^2 \cdot 12x^2(1-x) dx \\
& = \int_0^1 12 \left(-x^5 + \tfrac{11}{5}x^4 - \tfrac{39}{25}x^3 + \tfrac{9}{25}x^2\right) dx \\
& = 12 \left(-\tfrac{1}{6}x^6 + \tfrac{11}{5}\tfrac{1}{5}x^5 - \tfrac{39}{25}\tfrac{1}{4}x^4 + \tfrac{9}{25}\tfrac{1}{3}x^3\right) \Big|_0^1 \\
& = 12 \left(-\tfrac{1}{6} + \tfrac{11}{5}\tfrac{1}{5} - \tfrac{39}{25}\tfrac{1}{4} + \tfrac{9}{25}\tfrac{1}{3}\right) = 0.04
\end{aligned}$$

Esperanza y varianza

Ejemplo. Solución con R/Rcommander

Se definen las funciones a integrar:

```
integrando2 <- function(x) {x * 12 * x^2 * (1-x)}
media <- mi.integrate(integrando2, 0, 1)
media

integrando3 <- function(x) {(x-media[[1]])^2 * 12 * x^2 *
(1-x)}
varianza <- mi.integrate(integrando3, 0, 1)
varianza
```

Ejemplo de repaso: distribución uniforme en (3,5)

Algunas probabilidades

Una variable aleatoria X que sigue una distribución uniforme en el intervalo $(3,5)$ tiene función de densidad

$$f(x) = \begin{cases} \frac{1}{5-3} = \frac{1}{2}, & \text{si } x \in (3,5) \\ 0, & \text{si } x \notin (3,5). \end{cases}$$

Calculamos algunas probabilidades:

$$P(X \leq 0.5) = \int_{-\infty}^{0.5} f(u)du = 0$$

$$P(X \leq 4) = \int_{-\infty}^{4} f(u)du = \int_3^4 \tfrac{1}{2}du = \tfrac{1}{2}u|_3^4 = \tfrac{1}{2}$$

$$P(3.5 \leq X \leq 4.5) = \int_{3.5}^{4.5} f(u)du = \int_{3.5}^{4.5} \tfrac{1}{2}du = \tfrac{1}{2}$$

Ejemplo de repaso: distribución uniforme en (3,5)

Función de distribución

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)du = \ldots$$

- Si $x \leq 3$ entonces $F(x) = P(X \leq x) = 0$.
- Si $3 \leq x < 5$ entonces $F(x) = P(X \leq x) = \int_3^x \frac{1}{2}du = \frac{u}{2}\,|_3^x\, = \frac{x-3}{2}$.
- Si $x \geq 5$ entonces $F(x) = P(X \leq x) = \int_3^5 \frac{1}{2}du = \frac{u}{4}\,\Big|_3^5\, = \frac{5-3}{2} = 1$.

$$F(x) = \begin{cases} 0, & \text{si } x \leq 3, \\ \frac{x-3}{2}, & \text{si } 3 < x < 5, \\ 1, & \text{si } x \geq 5. \end{cases}$$

Ejemplo de repaso: distribución uniforme en (3,5)

Esperanza

$$E(X) = \int_{\mathbb{R}} x \cdot f(x)dx = \int_3^5 x \cdot \frac{1}{2}dx = \left.\frac{x^2}{4}\right|_3^5 = \frac{5^2-3^2}{4} = 4$$

Varianza

$$var(X) = \int_{\mathbb{R}} x^2 \cdot f(x)dx - E^2[X]$$

$$= \int_3^5 \frac{x^2}{2}dx - 4^2 = \left.\frac{x^3}{6}\right|_3^5 - 16 = 0.33$$

Algunos modelos probabilísticos

Modelos discretos

- Ensayos de Bernoulli
- Distribución Binomial
- Distribución de Poisson

Modelos continuos

- Distribución uniforme
- Distribución exponencial
- Distribución normal
- Distribuciones asociadas a la normal

Modelo Bernoulli

Descripción / Definición

Es una forma de modelar estadísticamente cualquier experimento aleatorio que tenga solamente dos resultados posibles, mútuamente excluyentes, que suelen llamarse *éxito* y *fracaso*, con la condición de que la probabilidad de estos dos resultados se mantenga constante en cada realización del experimento (experimentos o ensayos de Bernoulli).

Si la probabilidad de éxito es p (por tanto, la de fracaso es $1-p$), se define la variable aleatoria de Bernoulli como

$$X = \begin{cases} 1, & \text{si se observa un éxito,} \\ 0, & \text{si se observa un fracaso.} \end{cases}$$

La v.a. X toma valores en $S = \{0, 1\}$ con probabilidades $P(X = 0) = 1 - p$, $P(X = 1) = p$.

Para denotar que X sigue una ley de Bernoulli de parámetro p escribiremos $X \sim Ber(p)$.

Modelo Bernoulli

Ejemplo

Resultado de lanzar una moneda al aire

$$X = \begin{cases} 1, & \text{sale cara,} \\ 0, & \text{si sale cruz.} \end{cases}$$

Es un ensayo Bernoulli, donde se ha considerado como *éxito* el observar una cara. X sigue una distribución Bernoulli de parámetro $1/2$ (si la moneda no está trucada).

Ejemplo

Una línea aérea estima que los pasajeros que compran un billete no se presentan al embarque con una probabilidad de 0.05.
Definimos

$$Y = \begin{cases} 1, & \text{si el pasajero se presenta,} \\ 0, & \text{si el pasajero no se presenta.} \end{cases}$$

Y sigue una distribución Bernoulli con parámetro 0.95.

Modelo Bernoulli

Función de Probabilidad:

$$P(X=0)=1-p \quad P(X=1)=p$$

Función de distribución:

$$F(x)=\left\{\begin{array}{cl} 0 & \text{si } x<0 \\ 1-p & \text{si } 0\le x<1 \\ 1 & \text{si } x\ge 1 \end{array}\right.$$

Propiedades

- $E(X)=0\,P(X=0)+1\,P(X=1)=0\,(1-p)+1\,p=p$
- $E(X^2)=0^2\,P(X=0)+1^2\,P(X=1)=0^2\,(1-p)+1^2\,p=p$
- $var(X)=E(X^2)-E(X)^2=p-p^2=p(1-p)$

Modelo Binomial

Descripción / Definición

Se realizan n experimentos e Bernoulli con la misma probabilidad de éxito p. La v.a. X que cuenta el número de éxitos observados en estos n experimentos se dice que sigue una distribución Binomial de parámetros n y p y se escribe $X\sim B(n,p)$.

La v.a. X toma valores en $S=\{0,1,2,\ldots,n\}$ y su función de probabilidad viene dada por la fórmula

$$P(X=x)=\binom{n}{x}\,p^x\,(1-p)^{n-x}, \quad x=0,1,\ldots,n, \quad 0\le p\le 1,$$

donde $\binom{n}{x}=\frac{n!}{x!(n-x)!}$, para $0\le x\le n$. Recordad que, por convenio, $0!=1$.

Propiedades

$E(X)=np, \quad var(X)=np(1-p).$

Modelo Binomial

Ejemplo

La línea aérea del ejemplo anterior ha vendido 80 billetes para un vuelo. La probabilidad de que un pasajero no se presente al embarque es de 0.05. Definimos $X =$ *número de pasajeros que se presentan al embarque*. Entonces (suponiendo independencia)

$$X \sim B(80, 0.95)$$

- La probablidad de que los 80 pasajeros se presenten es

$$P(X = 80) = \binom{80}{80} 0.95^{80} \times (1 - 0.95)^{80-80} = 0.0165$$

- La probabilidad de que al menos un pasajero no se presente es

$$P(X < 80) = 1 - P(X = 80) = 1 - 0.0165 = 0.9835$$

Distribución de Poisson: sucesos raros

Descripción / Definición

Cuenta el número de sucesos *raros* que ocurren en una determinada unidad de tiempo o de espacio. Por ejemplo, llamadas de teléfono en una hora, erratas en una página, accidentes de tráfico a la semana, ...

Una v.a. X sigue una distribución de Poisson de parámetro λ, y se denotará por $X \sim Pois(\lambda)$, si su función de probabilidad es

$$P(X = x) = e^{-\lambda} \frac{\lambda^x}{x!}, \qquad \text{para } x = 0, 1, 2, \ldots$$

Observad que X toma valores en $S = \{0, 1, 2, \ldots\} = \mathbb{N} \cup \{0\}$.

Propiedades

$E(X) = \lambda, \quad var(X) = \lambda.$

λ representa el número medio de sucesos que se producen por unidad.

Distribución de Poisson: sucesos raros

Propiedad de la Poisson

Si $X \sim Pois(\lambda)$ y representa el número de sucesos raros en una unidad de tiempo o de espacio, e Y es una variable aleatoria que representa el número de dichos sucesos raros en s unidades, se tiene que:

$$Y \sim Pois(s\lambda)$$

Distribución de Poisson: sucesos raros

Ejemplo

El número medio de erratas por transparencias es de 0.2. Sea X es la v.a. que cuenta el número de erratas por transparencia, entonces

$$X \sim Pois(0.2)$$

¿Cuál es la probabilidad de que en una transparencia no haya erratas?

$$P(X=0) = e^{-0.2}\,\frac{0.2^0}{0!} = e^{-0.2} = 0.8187.$$

¿Cuál es la probabilidad de que en 4 transparencias haya exactamente una errata?
Sea Y la v.a. que cuenta el número de erratas en 4 transparencias. Entonces:

$$Y \sim Pois(0.2 \cdot 4) = Pois(0.8)$$

$$P(Y=1) = e^{-0.8}\,\frac{0.8^1}{1!} = e^{-0.8}\,0.8 = 0.3595.$$

Distribución uniforme

Descripción / Definición

Se dice que una variable X sigue una distribución uniforme en el intervalo (a,b), y se denota por $X \sim \mathcal{U}(a,b)$, si su función de densidad es

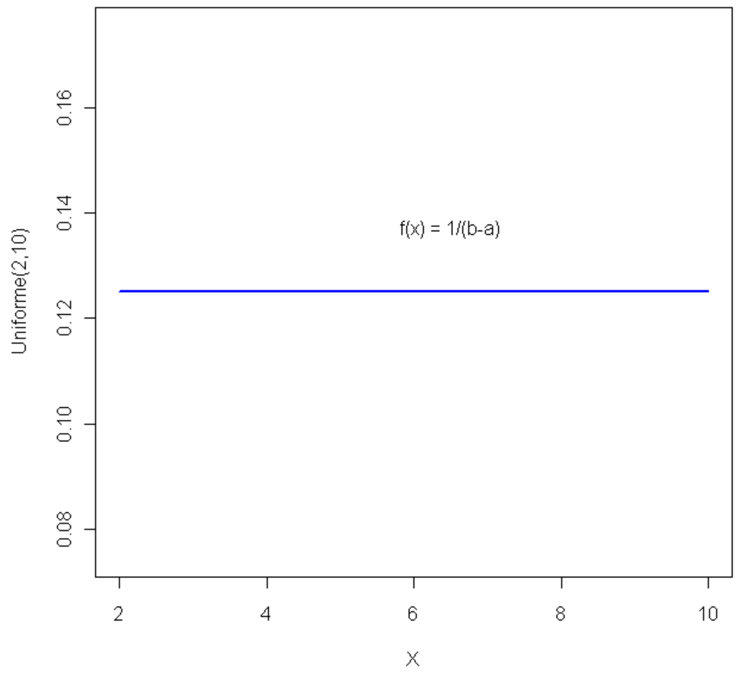
$$f(x) = \begin{cases} \frac{1}{b-a}, & \text{si } x \in (a,b), \\ 0, & \text{si } x \notin (a,b). \end{cases}$$

Esta v.a. queda definida por los extremos del intervalo, es decir, a y b son sus parámetros.

Propiedades

$E(X) = \frac{a+b}{2}, \quad var(X) = \frac{(b-a)^2}{12}.$

Distribución uniforme



Distribución exponencial

Descripción / Definición

La distribución exponencial es aquella que modela el tiempo transcurrido entre dos sucesos que se producen de forma independiente, separada y uniforme en el tiempo.

Se dice que una v.a. X sigue una distribución exponencial de parámetro λ, y se denota por $X \sim exp(\lambda)$, si su función de densidad es

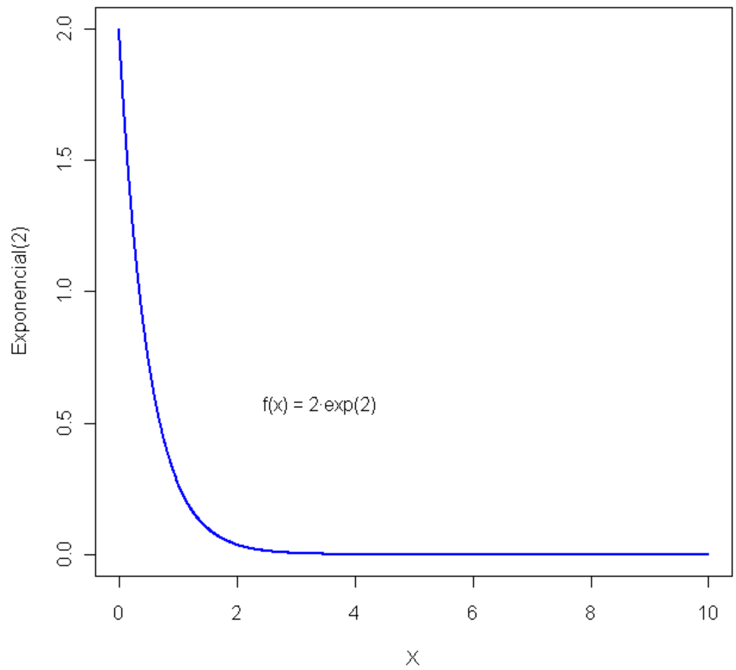
$$f(x) = \lambda \, e^{-\lambda x}, \qquad \text{para } x \geq 0.$$

Observad que X toma valores en el conjunto $S = [0, +\infty)$.

Ejemplos

- Tiempo entre llegadas de camiones al punto de descarga.
- Tiempo entre llamadas de emergencia.
- Tiempo de vida de una bombilla.

Distribución exponencial



Distribución exponencial

Propiedades

- $E(X) = \frac{1}{\lambda}$
- $var(X) = \frac{1}{\lambda^2}$
- Función de distribución:

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & \text{si } x \geq 0, \\ 0, & \text{si } x < 0. \end{cases}$$

- Está relacionada con la distribución de Poisson.
- λ es el número medio de ocurrencias del suceso por unidad de tiempo.

Distribución exponencial

Ejemplo

Hemos observado que en cierta provincia se producen, en promedio, 50 incendios serios cada año. Suponemos que estos incendios se producen de forma independiente y decidimos modelar el número de incendios por año mediante una distribución Poisson.

- ¿Cuál es el tiempo medio que transcurre entre dos incendios consecutivos?
- Si acaba de ocurrir un incendio ¿cuál es la probabilidad de que el próximo se produzca al cabo de dos semanas?

Sabemos que:

- El número de incendios por año $N \sim Pois(\lambda)$ con $\lambda = 50$.
- El tiempo entre dos incendios $X \sim exp(\lambda)$ con $\lambda = 50$.
- El tiempo medio entre dos incendios $E(X) = \frac{1}{\lambda} = 1/50$ años, 7.3 días.
- Dos semanas, en años son: $\frac{2 \cdot 7}{365} = 0.03836$,
- $P[X > 0.03836] = 1 - P[X \leq 0.03836] = 1 - (1 - e^{-50 \cdot 0.03836}) = 0.147$.

Distribución normal

Descripción / Definición

La distribución (o ley) normal describe una variable aleatoria "ideal". Se trata de un modelo teórico que aproxima bien muchas situaciones reales.
La inferencia estadística se fundamenta básicamente en la ley normal y en distribuciones que se derivan de ella.

Se dice que una v.a. X sigue una distribución normal o gausiana con parámetros μ y σ, y se denota por $X \sim \mathcal{N}(\mu, \sigma)$, si su función de densidad es

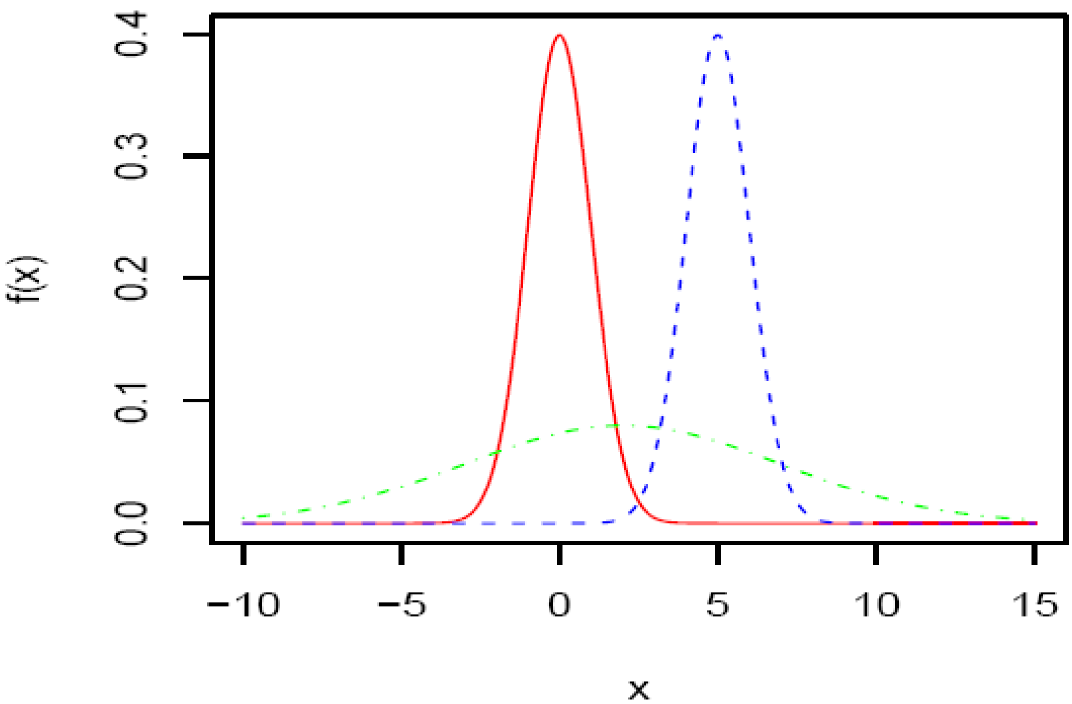
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}$$

Propiedades

$E(X) = \mu, \quad var(X) = \sigma^2.$

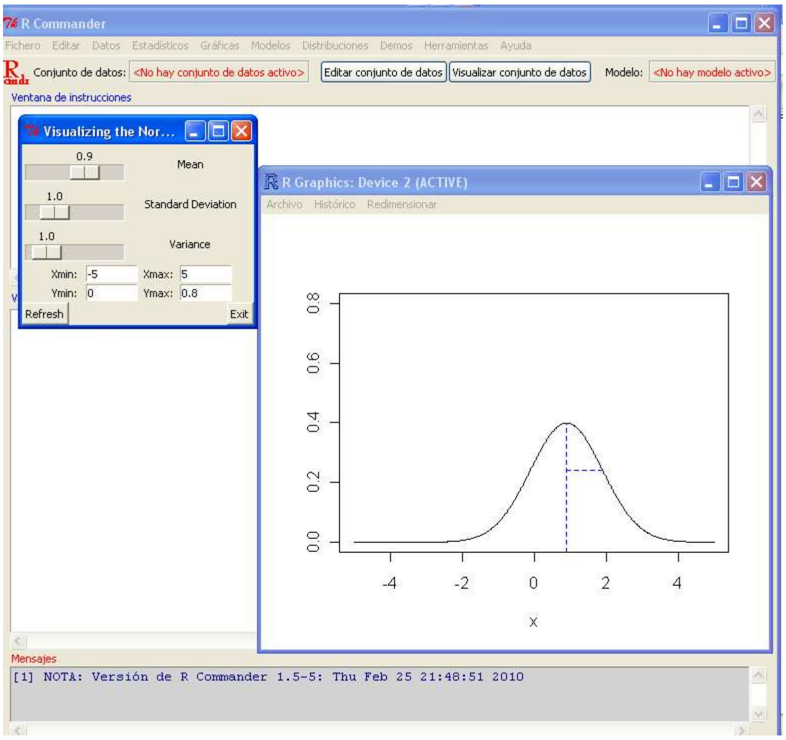
Distribución normal

Función de densidad para 3 valores distintos de μ y σ



Visualización con R/RCommander

Con el *plug-in* de `TeachingDemos`



Distribución normal

Propiedad

Si $X \sim \mathcal{N}(\mu, \sigma)$, entonces:

- $P(\mu - \sigma < X < \mu + \sigma) \approx 0.683$
- $P(\mu - 2\sigma < X < \mu + 2\sigma) \approx 0.955$
- $P(\mu - 3\sigma < X < \mu + 3\sigma) \approx 0.997$

Cota de Chebyshev

Si $X \sim \mathcal{N}(\mu, \sigma)$,

$$P(\mu - k\sigma < X < \mu + k\sigma) \geq 1 - \frac{1}{k^2}$$

Distribución normal

Transformación lineal

$$Y = a + bX \sim \mathcal{N}(a + b\mu, |b|\sigma)$$

Estandarización

Si $X \sim \mathcal{N}(\mu, \sigma)$, considero

$$Z = \frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1)$$

Se llama distribución normal estándar.

Tablas de la $\mathcal{N}(0, 1)$

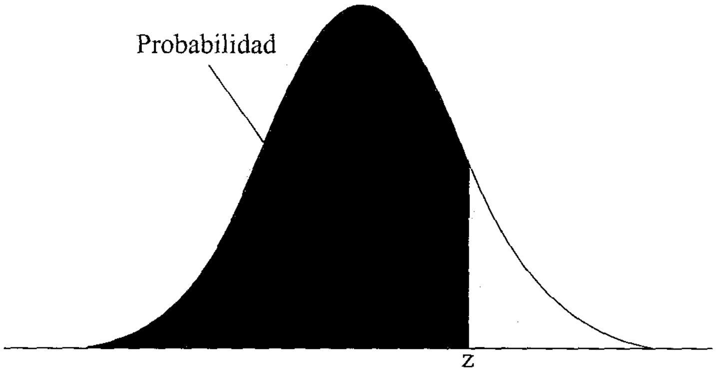


Tabla 3. (*continuación*) *Probabilidad de que una variable normal de media cero y desviación típica uno tome un valor menor que z*

z	**0,00**	**0,01**	**0,02**	**0,03**	**0,04**	**0,05**	**0,06**	**0,07**	**0,08**	**0,09**
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830

Distribución normal: Ejemplo

Sea $Z \sim N(0,1)$. Calculemos algunas probabilidades:

$$P(Z < 1.5) = 0.9332 \quad \textbf{(Ver tabla)}$$

```
pnorm(1.5)
```

$$\begin{aligned} &P(Z > -1.5) = P(Z < 1.5) = 0.9332 \quad \textbf{¿por qué?} \\ &P(Z < -1.5) = P(Z > 1.5) = 1 - P(Z < 1.5) = \\ &= 1 - 0.9332 = 0.0668 \quad \textbf{¿por qué no} \leq \textbf{?} \end{aligned}$$

```
pnorm(-1.5)
```

$$\begin{aligned} P(-1.5 &< Z < 1.5) = P(Z < 1.5) - P(Z < -1.5) = \\ &= 0.9332 - 0.0668 = 0.8664 \end{aligned}$$

```
diff(pnorm(c(-1.5,1.5),0,1))
```


Distribución normal: Ejemplo

Sea $X \sim N(\mu = 2, \sigma = 3)$.
Calcular $P(X < 4)$.
En este caso, tipificamos la variable original:

$$P(X < 4) = P\left(\frac{X-2}{3} < \frac{4-2}{3}\right) = P\left(Z < 0.66\dot{6}\right) \approx 0.7454$$

donde $Z \sim N(0,1)$

¿Cuál es $P(-1 < X < 3.5)$?

$$\begin{aligned} &P(-1 < X < 3.5) = P(-1-2 < X-2 < 3.5-2) = \\ &P\left(\frac{-1-2}{3} < \frac{X-2}{3} < \frac{3.5-2}{3}\right) = P(-1 < Z < 0.5) = \\ &P(Z < 0.5) - P(Z < -1) = 0.6915 - 0.1587 = 0.5328 \end{aligned}$$

donde $Z \sim N(0,1)$

Distribución normal: otro ejemplo

Es difícil etiquetar la carne empaquetada con su peso correcto debido a los efectos de pérdida de líquido (definido como porcentaje del peso original de la carne). Supongamos que la pérdida de líquido en un paquete de pechuga de pollo puede modelarse mediante una ley normal con media 4% y desviación típica 1%.
Sea X la pérdida de líquido de un paquete de pechuga de pollo elegido al azar.

- ¿Cuál es la probabilidad de que $3\% < X < 5\%$?
- ¿Cuál es el valor de x para que un 90% de paquetes tengan pérdidas de líquido menores que x?
- En una muestra de 4 paquetes, hallar la probabilidad de que todos tengan pérdidas de peso de entre 3% y 5%.

Sexauer, B. (1980) *Drained-Weight Labelling for Meat and Poultry: An Economic Analysis of a Regulatory Proposal*, Journal of Consumer Affairs, 14, 307-325.

Distribución normal: otro ejemplo

$$P(3 < X < 5) = P\left(\frac{3-4}{1} < \frac{X-4}{1} < \frac{5-4}{1}\right) = P(-1 < Z < 1)$$
$$= P(Z < 1) - P(Z < -1) = 0.8413 - 0.1587 = 0.6827$$

Queremos $P(X < x) = 0.9$. Entonces

$$P\left(\frac{X-4}{1} < \frac{x-4}{1}\right) = P(Z < x-4) = 0.9$$

Mirando las tablas, tenemos $x - 4 \approx 1.28$ que implica que un 90% de las paquetes tienen pérdidas de menos de $x = 5.28\%$.

Para un paquete $p = P(3 < X < 5) = 0.6827$. Sea Y el número de paquetes en la muestra que tienen pérdidas de entre 3% y 5%. Luego $Y \sim B(4, 0.6827)$.

$$P(Y = 4) = \left(\begin{array}{c} 4 \\ 4 \end{array}\right) 0.6827^4 (1 - 0.6827)^4 = 0.2172$$

Teorema Central del Límite (TCL)

El siguiente teorema nos habla de la distribución de la media de un conjunto de v.a. independientes e idénticamente distribuidas (i.i.d.), es decir, todas con la misma ley de probabilidad,

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

y nos dice que, para n grande, la ley de la media de v.a. independientes e igualmente distribuidas es normal, sea cual sea la ley de las v.a.
De aquí el papel "central" que juega la ley normal o de Gauss.

Teorema

Sean $X_1, X_2, \ldots, X_n$ v.a. i.i.d. con media μ y desviación típica σ (ambas finitas). Si n es suficientemente grande, se tiene que

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$$

Aproximaciones

Binomial (Teorema de De Moivre-Laplace)

Si $X \sim B(n,p)$ con n suficientemente grande

$$\frac{X - np}{\sqrt{np(1-p)}} \sim \mathcal{N}(0,1)$$

Poisson

Si $X \sim Pois(\lambda)$ con λ suficientemente grande

$$\frac{X - \lambda}{\sqrt{\lambda}} \sim \mathcal{N}(0,1)$$

TCL y aproximaciones: Ejemplo

Sea $X \sim B(100, 1/3)$. Estimar $P(X < 40)$.

Calculamos primero la media y varianza de X.

$$E(X) = 100 \times \frac{1}{3} = 33.\dot{3}$$
$$var(X) = 100 \times \frac{1}{3} \times \frac{2}{3} = 22.\dot{2}$$
$$D.T.(X) = \sqrt{22.\dot{2}} = 4.714$$

Usamos la aproximación normal

$$P(X < 40) = P\left(\frac{X - 33.\dot{3}}{4.714} < \frac{40 - 33.\dot{3}}{4.714}\right)$$
$$\approx P\left(Z < 1.414\right) \approx 0.921,$$
$$\text{donde } Z \sim N(0, 1).$$

Distribución asociada a la normal

t de Student

Sean $Y, X_1, X_2, \ldots, X_n$ v.a. i.i.d. con ley $\mathcal{N}(0, 1)$. La distribución de

$$t_n = \frac{Y}{\sqrt{\sum_{i=1}^{n} X_i^2 / n}}$$

se llama distribución t de Student con n grados de libertad.

- $E(t_n) = 0$
- $var(t_n) = \frac{n}{n-2}$