Variables aleatorias

Definición, propiedades y funciones

Fabián Abarca Calderón

IE0405 – Modelos Probabilísticos de Señales y Sistemas

4

Tema I

UNIVERSIDAD DE COSTA RICA

EIE Escuela de Ingeniería Eléctrica

Las variables aleatorias facilitan una manipulación numérica más robusta de los fenómenos aleatorios, y permiten extender el análisis a muchos más casos que los vistos hasta ahora. Herramientas como las funciones de densidad y de distribución acumulativa de probabilidad proveen descripciones completas de los modelos probabilísticos.

Variables aleatorias

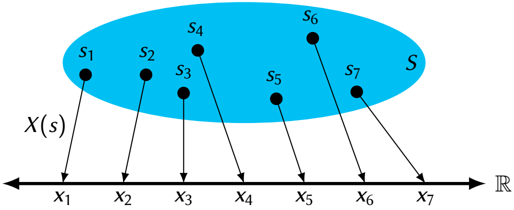
Los conjuntos $S_1 = \{$ todos los equipos del campeonato nacional $\}$ o $S_2 = \{$ los colores favoritos de los estudiantes de esta clase $\}$ son útiles en la descripción de ciertos eventos, pero no permiten la *manipulación numérica*.

- Este espacio de eventos contiene conjuntos **abstractos**.
- Necesitamos **números** para sumar, restar, multiplicar, dividir...
- Necesitamos **funciones** para diferenciar, integrar...

Aquí es útil, entonces, la variable aleatoria...

Definición

Para un espacio de eventos S, una **variable aleatoria** es cualquier regla que asocia cada resultado elemental de S con **un número**. Es decir, es una **función** cuyo dominio es el espacio (quizá abstracto) de eventos o muestras, y cuyo rango es algún subconjunto de los números reales.



La notación $X(s) = x$ significa que x es el valor asociado por X con el evento elemental s.

Las variables aleatorias *ni son variables, ni son aleatorias*[1].

$X(s)$ es una **función** y es *determinística*, pero describe el comportamiento de un fenómeno aleatorio subyacente.

[1] Por tanto, se trata de un nombre poco apropiado. En inglés: *misnomer*.

Requisitos para la construcción de variables aleatorias I

Un espacio de probabilidades (S, P) para un experimento, que contiene todos los eventos elementales S y sus probabilidades asociadas P.

Una función de mapeo X que mapea cada $s \in S$ a un único punto $x \in \mathbb{R}$ (la recta real).

Una relación de dualidad de la probabilidad, esto es, que si B es un subconjunto de $\mathbb{R}$, la probabilidad del evento $X(s) = B$ es equivalente a la del conjunto $A = X^{-1}(B) \in S$, que contiene todos los $s \in S$ que se mapean a B bajo la función X.

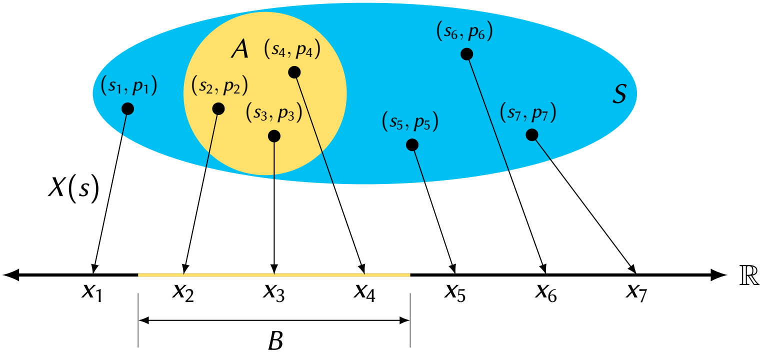


Figura: Mapeo de un subconjunto del espacio de eventos S en un segmento de la recta real $\mathbb{R}$.

$$X(s \in A) = B \qquad A = X^{-1}(B) \in S \qquad P(B) = P(A)$$

Nota: VA será la abreviación de “variable aleatoria”, de aquí en adelante. En inglés se utiliza **rv**, de *random variable*.

Algunas condiciones debe cumplir $X(s)$ para ser una VA.

1. Una variable aleatoria es una función **no multivaluada**. Es decir, todo punto en S corresponde a solo un valor de la VA.

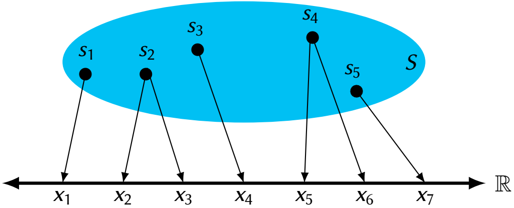
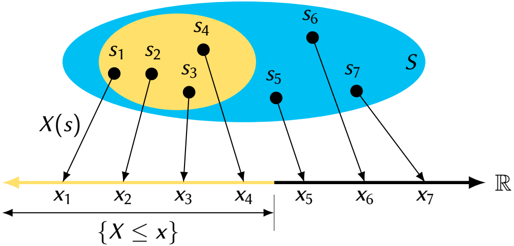


Figura: Esto **no** (atención: **no**) representa el mapeo de una VA.

Condiciones para que una función sea variable aleatoria II

2. El conjunto $\{X \leq x\}$ existe y es un evento para cualquier número real x. Corresponde a los puntos s en el espacio de muestras S para los que la VA $X(s)$ no excede el número x.



La probabilidad $P\{X \leq x\}$ es igual a la suma de las probabilidades de todos los eventos elementales s correspondientes a $\{X \leq x\}$.

3. Las probabilidades de los eventos $\{X = \infty\}$ y $\{X = -\infty\}$ son cero, es decir:

$$P\{X = -\infty\} = 0 \qquad P\{X = \infty\} = 0$$

$X(s)$ puede ser ya sea $-\infty$ o ∞ para algunos valores de s, pero su probabilidad será cero.

Como se especifica más adelante, esto es necesario para que su "función de densidad" tenga un área total finita.

Una variable aleatoria discreta es una VA cuyos valores posibles constituyen o un *conjunto* ***finito*** o un *conjunto infinito* ***enumerable***.

Ejemplos

- Las caras de un dado
- La población mundial
- Otros ejemplos que se mapean en $\mathbb{N}$

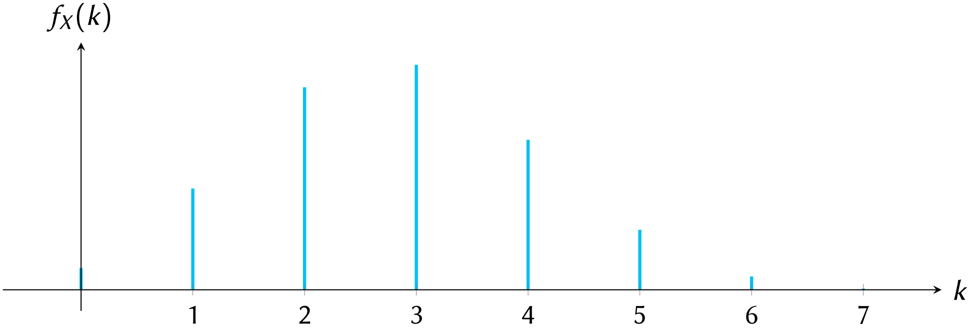


Figura: En una variable aleatoria discreta, la asignación de probabilidades se hace en valores discretos, ya sea finitos o infinitos enumerables. En este caso de ejemplo, están mapeados en un subconjunto de $\mathbb{N}$.

Una VA es continua si las siguientes condiciones aplican:

1. Su conjunto de valores posibles consiste en todos los números en un solo intervalo en la recta real (por ejemplo, de 0 a $+\infty$) o todos los números en la unión disjunta de intervalos.
2. Ningún valor posible de la variable aleatoria continua tiene probabilidad positiva, es decir, $P(X = c) = 0$ para cualquier valor de c.

Ejemplo

La estatura de los habitantes de un país es una VA continua. Pero, ¿cuál es la probabilidad de que alguien mida **exactamente** 52π cm?

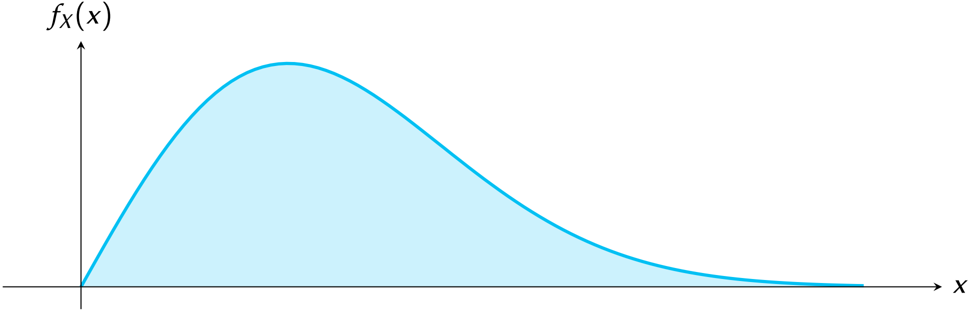


Figura: En una variable aleatoria continua, la asignación de probabilidades se hace en una sección o toda la recta real. Debido a que hay infinitos posibles resultados, la probabilidad de ocurrencia de exactamente un valor en particular es cero.

Las funciones de distribución y de densidad

- Las funciones de distribución y de densidad de una variable aleatoria son **descripciones completas** de su *modelo probabilístico*.
- Es la generalización de la **asignación de probabilidades** a todos los posibles resultados del experimento, que ahora están mapeados en la recta real y por eso podemos crear una función sobre ellos.

La más famosa de las funciones de probabilidad es la *normal* (o *gaussiana*), que describe la noción de que las probabilidades de ocurrencia de un evento están concentradas alrededor de un valor central.

La función de probabilidad acumulativa, $F_X(x)$

La premisa de la función de probabilidad acumulativa es que, por definición, $F_X(x)$ representa "la probabilidad de que el resultado del experimento sea un valor menor a x". Conforme x aumenta se "acumula" toda la probabilidad anterior (pues está contenida dentro del nuevo intervalo), y de ahí su nombre.

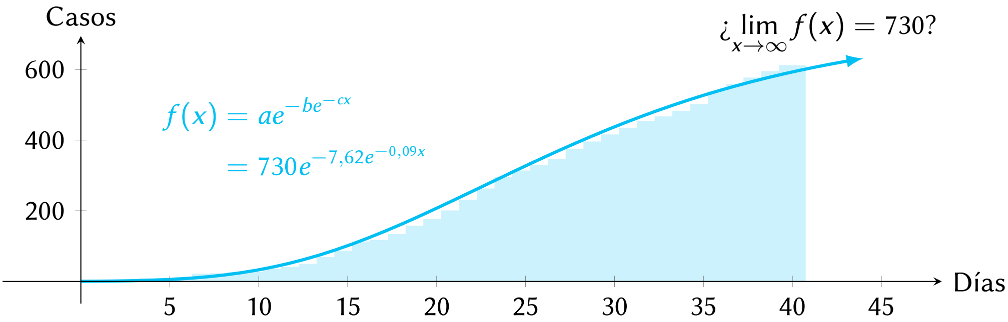


Figura: Casos acumulados de Covid-19 en los primeros 40 días de contagio en Costa Rica, modelados con la curva sigmoidal tipo Gompertz, que es un caso especial de la curva logística y utilizado en crecimiento de poblaciones. Es un ejemplo de función *acumulativa* pero no de *probabilidad*. Datos y modelado del profesor Víctor Granados.

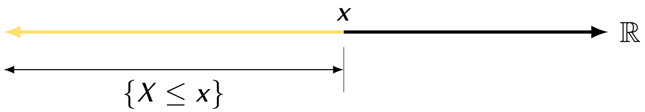
La función de probabilidad acumulativa, $F_X(x)$

Definición

Si la probabilidad $P\{X \leq x\}$ es la probabilidad del evento $\{X \leq x\}$, entonces a la función de x

$$\boxed{F_X(x) = P\{X \leq x\}} \tag{1}$$

se le llama *función de distribución de probabilidad acumulativa*[2] de la variable aleatoria X. El argumento x es cualquier número real entre $-\infty$ y ∞.



[2]A menudo llamada simplemente **función acumulativa** de X. En inglés llamada *Cumulative Distribution Function*, **CDF**.

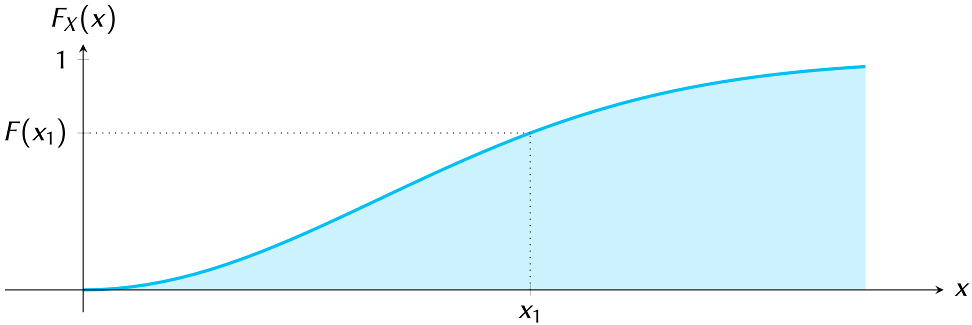


Figura: En una función acumulativa de probabilidad de una VA X particular, los valores $x_1 = 4$ y $F_X(4) = 0{,}6811$ se interpretan como que hay un 68,11 % de probabilidad de que el valor de x está por debajo de 4, es decir, $P(X < 4) = 0{,}6811$

La función de distribución de probabilidad acumulativa presenta las siguientes propiedades.

1. La probabilidad de que X tenga un valor menor a $-\infty$ es cero

$$\boxed{F_X(-\infty) = 0}$$

   Esto tiene relación con el hecho conocido de que $P(\overline{S}) = 0$.

2. La probabilidad de que X tenga un valor menor a $+\infty$ (dentro de la recta real) es uno

$$\boxed{F_X(\infty) = 1}$$

   Esto tiene relación con el hecho conocido de que $P(S) = 1$.

3. $F_X(x) = P(X < x)$ es una probabilidad por sí misma, con valores acotados según el primer y segundo axioma de la probabilidad

$$\boxed{0 \leq F_X(x) \leq 1}$$

Esto tiene relación directa con las dos propiedades anteriores.

4. $F_X(x)$ es una función no-decreciente de x

$$\boxed{F_X(x_1) \leq F_X(x_2) \quad \text{si} \quad x_1 < x_2}$$

Esto es claro por el hecho de que la función es "acumulativa".

5. La probabilidad de que X tenga valores más grandes que algún número x_1 pero que no exceda otro número x_2, es igual a la diferencia en $F_X(x)$ evaluada en tales puntos.

$$\boxed{P(x_1 < X \leq x_2) = F_X(x_2) - F_X(x_1)}$$

   Esta propiedad permite hacer cálculos numéricos de probabilidades a partir de la CDF.

6. $F_X(x)$ tiene continuidad por la derecha

$$\boxed{F_X(x^+) = F_X(x)}$$

   donde x^+ significa $x + \epsilon$, y $\epsilon > 0$ es infinitesimalmente pequeño; es decir, $\epsilon \to 0$. Esto es útil con VA discretas.

Si X es una variable aleatoria discreta, $F_X(x)$ debe tener una forma escalonada. La amplitud de un escalón igualará la probabilidad de ocurrencia del valor de X donde el escalón ocurre.

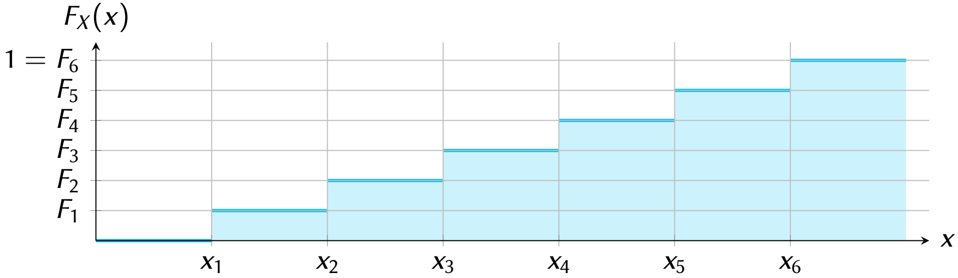


Figura: Función de distribución acumulativa discreta, donde se utiliza la notación $F_X(x_i) = F_i$.

Si los valores de X se denotan x_i, $F_X(x)$ se escribe como:

$$\boxed{F_X(x) = \sum_{i=1}^{N} P\{X = x_i\} u(x - x_i)} \tag{2}$$

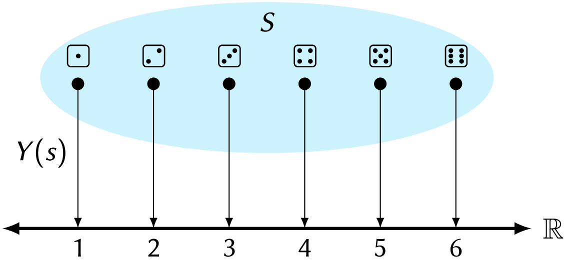
donde $u(\cdot)$ es la función escalón unitario. Si se introduce la notación abreviada $P(x_i) = P\{X = x_i\}$, se puede reescribir como

$$\boxed{F_X(x) = \sum_{i=1}^{N} P(x_i) u(x - x_i)} \tag{3}$$

Ejemplo con un dado I

Función acumulativa discreta

Un experimento consiste en lanzar un dado. Hay seis resultados posibles: las caras del dado, identificadas como $S = \{$⚀, ⚁, ⚂, ⚃, ⚄, ⚅$\}$. Todas tienen igual probabilidad de ocurrencia. Se define una variable aleatoria $Y(s)$ que mapea cada cara a un número en la recta real igual a la cantidad de puntos en la cara del dado.



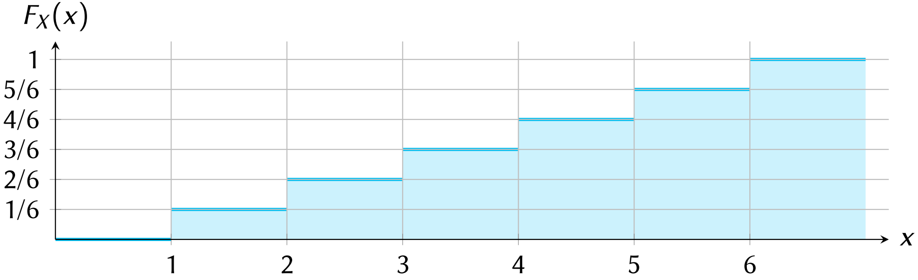
Dado que son igualmente probables, las probabilidades son

$$P(x_i) = P(X = x_i) = \frac{1}{6}$$

i	1	2	3	4	5	6
x_i	1	2	3	4	5	6
$P(x_i)$	1/6	1/6	1/6	1/6	1/6	1/6

Y así su función de distribución es

$$F_X(x) = \sum_{i=1}^{N} P(x_i)u(x - x_i) = \sum_{i=1}^{6} \frac{1}{6}u(x - x_i)$$



¿Cuál es la probabilidad $P(1 < x \leq 5)$? Utilizando directamente la propiedad 5:

$$
\begin{aligned}
P\{x_1 < X \leq x_2\} &= F_X(x_2) - F_X(x_1) \\
&= F_X(5) - F_X(1) = \frac{5}{6} - \frac{1}{6} = \frac{4}{6}
\end{aligned}
$$

La función de densidad de probabilidad, $f_X(x)$

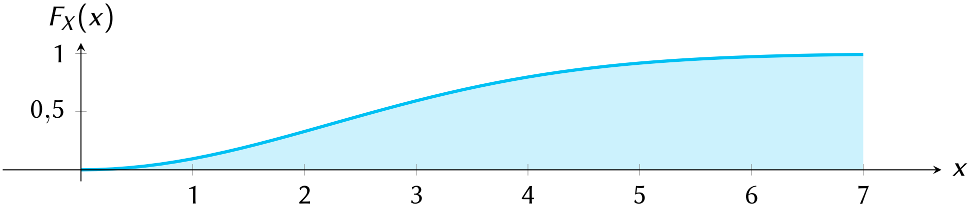
La función de densidad de probabilidad es más intuitiva que la función acumulativa. Visualmente, es un descripción directa de la *distribución* de la probabilidad en la recta real. Matemáticamente, es la base para una multitud de cálculos numéricos útiles en el análisis del fenómeno aleatorio.

La función de densidad probabilística[3] de la VA X, denotada por $f_X(x)$, está dada por la derivada de la función acumulativa:

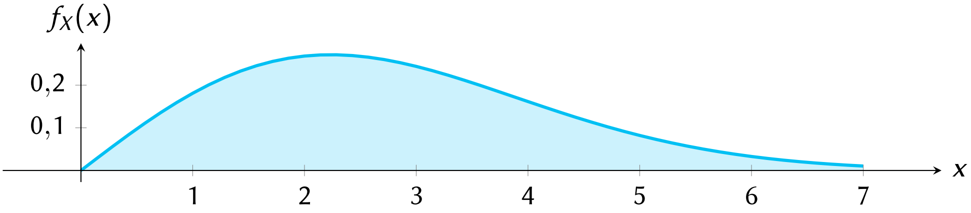
$$\boxed{f_X(x) = \frac{\mathrm{d}F_X(x)}{\mathrm{d}x}} \tag{4}$$

A $f_X(x)$ a menudo se le llama simplemente *función de densidad* o *función de distribución*.

[3] En inglés llamada *Probability Density Function*, **PDF**.



$F_X(x)$
1
0,5
1 2 3 4 5 6 7
x



$f_X(x)$
0,2
0,1
1 2 3 4 5 6 7
x

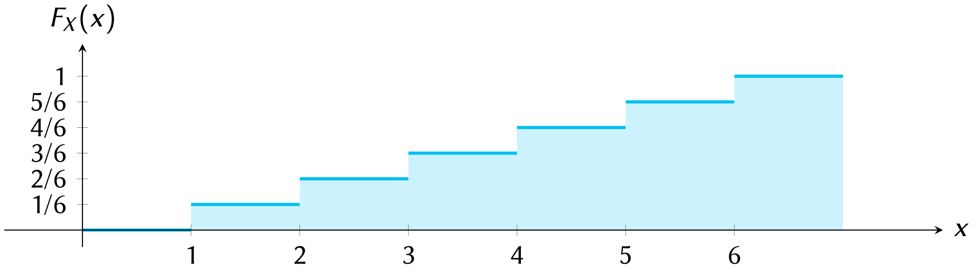
La función de densidad de probabilidad discreta, $f_X(x)$

Para una variable aleatoria discreta, la función de densidad discreta[4] se obtiene después de derivar la $F_X(x)$ discreta:

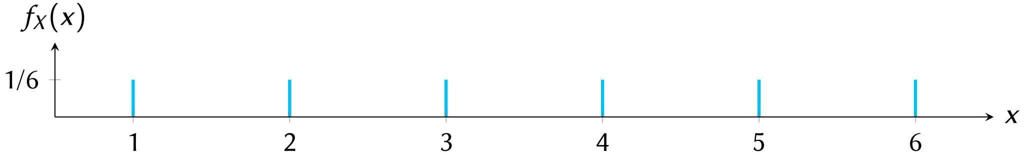
$$
\begin{aligned}
f_X(x) &= \frac{\mathrm{d}}{\mathrm{d}x}\sum_{i=1}^{N} P(x_i)u(x - x_i) \\
\boxed{f_X(x) = \sum_{i=1}^{N} P(x_i)\delta(x - x_i)}
\end{aligned}
\tag{5}
$$

En algunos textos, la PMF utiliza la notación $P_X(k)$ donde P_X es la función de la variable aleatoria X y $k = \{1, 2, 3 \ldots\}$.

[4] En inglés llamada *Probability Mass Function*, **PMF**.



$F_X(x)$
1
5/6
4/6
3/6
2/6
1/6
1
2
3
4
5
6
x



$f_X(x)$
1/6
1
2
3
4
5
6
x

Las funciones de densidad presentan las siguientes propiedades:

1. **La función $f_X(x)$ es siempre positiva**

$$f_X(x) \geq 0 \text{ para todo } x$$

   Esto a partir del primer axioma de la probabilidad. Por analogía con la densidad de masa: no existe masa negativa.

2. **El área bajo la curva es unitaria**

$$\int_{-\infty}^{\infty} f_X(x)\, \mathrm{d}x = 1$$

   A partir del segundo axioma de la probabilidad. Esta es una “normalización” necesaria para que una función sea una PDF.

3. **La probabilidad de la VA en un intervalo es el área bajo la curva en ese intervalo**

$$P\{x_1 < X \leq x_2\} = \int_{x_1}^{x_2} f_X(x)\, \mathrm{d}x$$

Esto permite hacer cálculos numéricos de probabilidad, y es lo que da origen a su nombre. Más adelante, con VA múltiples (en N dimensiones), será "el volumen bajo la superficie", etc.

4. **La función acumulativa se obtiene de la de densidad**

$$F_X(x) = \int_{-\infty}^{x} f_X(\xi)\, \mathrm{d}\xi$$

Esto a partir de la definición de la función de distribución acumulativa, $F_X(x) = P(X < x)$, y de la propiedad anterior.

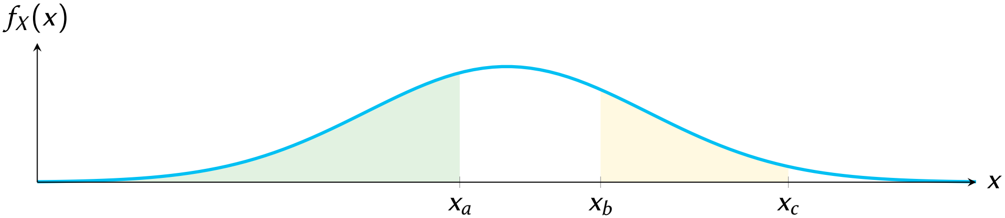


Figura: Función de densidad de probabilidad

1. $f_X(x) \geq 0$
2. $\int_{-\infty}^{\infty} f_X(x) \; \mathrm{d}x = 1$
3. $P(x_b < X \leq x_c) = \int_{x_b}^{x_c} f_X(x)\, \mathrm{d}x$
4. $F_X(x_a) = \int_{-\infty}^{x_a} f_X(x)\, \mathrm{d}x$

Algunas funciones de distribución probabilística de aplicación común

La función gaussiana tiene la *forma* general de la función

$$f(x) = e^{-x^2}$$

En particular, una VA X es normal si su función de densidad es

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma_X^2}} \exp\left[\frac{-(x-\mu_X)^2}{2\sigma_X^2}\right] \tag{6}$$

donde $\sigma_X > 0$ y $-\infty < \mu_X < \infty$ son constantes reales, conocidas como desviación estándar y media, respectivamente. La diferencia entre las ecuaciones anteriores es que (6) aplica la normalización (el área bajo la curva es 1) y el desplazamiento en el eje real (¡y por eso se ve tan complicada, pero la forma es la misma!).

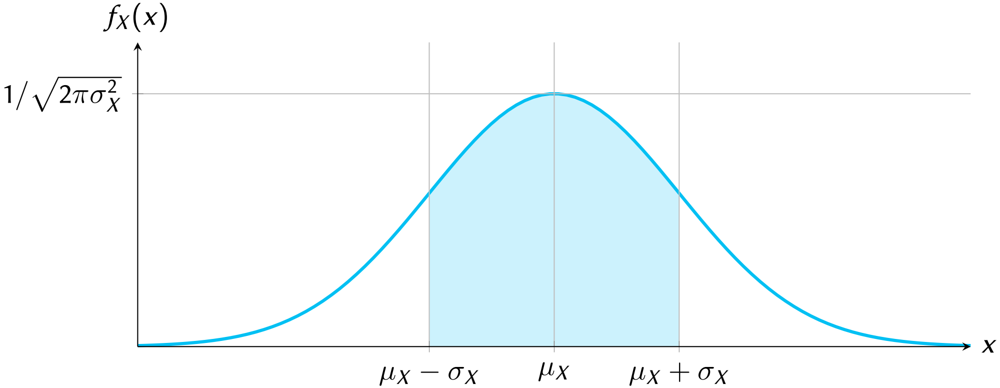
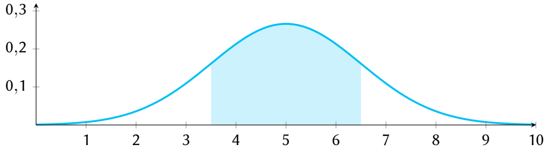


Figura: Función de distribución de probabilidad gaussiana

- Cumple con las propiedades $f_X(x) \geq 0$ para todo x, y $\int_{-\infty}^{\infty} f_X(x)\,\mathrm{d}x = 1$.
- Su valor máximo $1/\sqrt{2\pi\sigma_X^2}$ ocurre en $x = \mu_X$.
- Su dispersión (es decir, la forma particular en que se distribuyen los valores de la función) alrededor de $x = \mu_X$ está relacionado con σ_X.
- La función disminuye a 0,607 veces su máximo en $x = \mu_X + \sigma_X$ y en $x = \mu_X - \sigma_X$.



Como la variable gaussiana es tan común tiene una notación especial. Al decir:

$$X : \mathcal{N}(\mu, \sigma^2) \quad \text{o} \quad X \sim \mathcal{N}(\mu, \sigma^2)$$

significa que la VA X tiene una distribución gaussiana o "normal" con media μ y varianza σ_X^2 (la desviación estándar es σ_X).

¿Dónde aparece?

En el ruido térmico que afecta a la electrónica y en cierto tipo de interferencia que posa un canal inalámbrico sobre las comunicaciones. Además de innumerables otros fenómenos físicos, económicos y sociales, algo que quizá se explicará con el *teorema del límite central*, más adelante.

Las funciones de densidad probabilística y acumulativa uniforme están definidas por:

$$f_X(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{para otros valores de x} \end{cases} \tag{7}$$

$$F_X(x) = \begin{cases} 0 & x < a \\ \frac{(x-a)}{(b-a)} & a \leq x < b \\ 1 & b \leq x \end{cases} \tag{8}$$

para constantes reales $-\infty < a < \infty$ y $b > a$.

Figura: Función de densidad de probabilidad uniforme

¿Dónde aparece?

Cuando no se tiene mayor información sobre el comportamiento de la VA puede asumirse un comportamiento uniforme. **Ejemplo** (un mal ejemplo): una llamada de oficina sucederá entre las 8:00 am y las 5:00 pm con igual probabilidad.

Las funciones de distribución y de densidad exponencial son:

$$f_X(x) = \begin{cases} \frac{1}{b} \exp\left[-\left(\frac{x-a}{b}\right)\right] & x > a \\ 0 & x < a \end{cases} \tag{9}$$

$$F_X(x) = \begin{cases} 1 - \exp\left[-\left(\frac{x-a}{b}\right)\right] & x > a \\ 0 & x < a \end{cases} \tag{10}$$

para números reales $-\infty < a < \infty$ y $b > 0$.

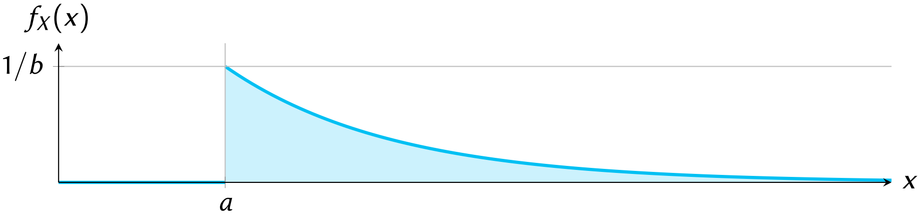


Figura: Función de densidad de probabilidad exponencial

¿Dónde aparece?

Ocurre en problemas de tiempo de espera o en el cálculo de la vida útil de maquinaria.

Las funciones de distribución y de densidad Rayleigh son:

$$f_X(x) = \begin{cases} \frac{2}{b}(x-a)\exp\left[-\left(\frac{(x-a)^2}{b}\right)\right] & x \geq a \\ 0 & x < a \end{cases} \tag{11}$$

$$F_X(x) = \begin{cases} 1-\exp\left[-\left(\frac{(x-a)^2}{b}\right)\right] & x \geq a \\ 0 & x < a \end{cases} \tag{12}$$

para constantes reales $-\infty < a < \infty$ y $b > 0$.

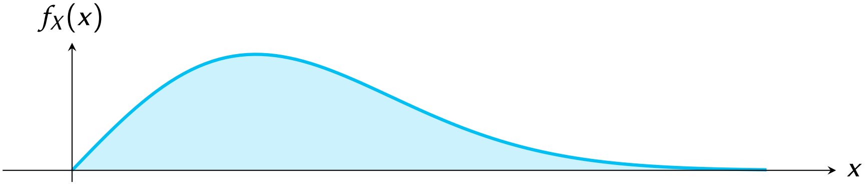


Figura: Función de densidad de probabilidad de Rayleigh

¿Dónde aparece?

Aparece en errores de aterrizaje de cohetes, fluctuaciones aleatorias de la envolvente de ciertas formas de onda, la distribución radial de los errores en un tablero de dardos, o los tiempos de llegada de las señales de múltiples trayectorias en transmisión inalámbrica.

Sea $0 < p < 1$ y $N = 1, 2, \ldots$, entonces la función

$$f_X(x) = \sum_{k=0}^{N} \binom{N}{k} p^k (1-p)^{N-k} \delta(x-k) \tag{13}$$

se llama la función de densidad binomial. La cantidad $\binom{N}{k}$ es el coeficiente binomial

$$\binom{N}{k} = \frac{N!}{k!(N-k)!} \tag{14}$$

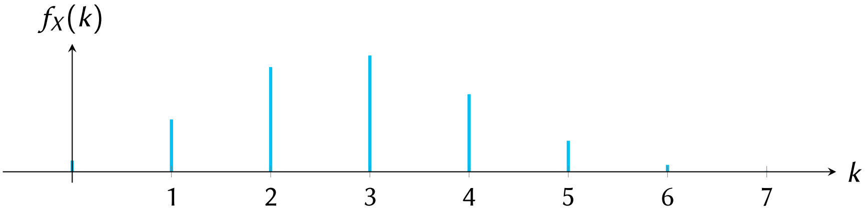


Figura: Función de densidad de probabilidad discreta binomial

¿Dónde aparece?

Modela la posibilidad de superar un umbral aceptable dada una tasa de "errores" (u ocurrencias de un evento). Se usa en análisis de riesgo, estimación de personal necesario según demanda de servicios, o número de defectos en un lote de producción.

La variable aleatoria de Poisson X tiene una densidad y distribución dadas por

$$f_X(x) = e^{-b}\sum_{k=0}^{\infty}\frac{b^k}{k!}\delta(x-k) \tag{15}$$

$$F_X(x) = e^{-b}\sum_{k=0}^{\infty}\frac{b^k}{k!}u(x-k) \tag{16}$$

donde $b > 0$ es una constante real. Cuando son graficadas, estas funciones parecen similares a la variable aleatoria binomial. De hecho, si $N \to \infty$ y $p \to 0$ para el caso binomial de tal manera que $Np = b$, una constante, entonces resulta la función de densidad de Poisson.

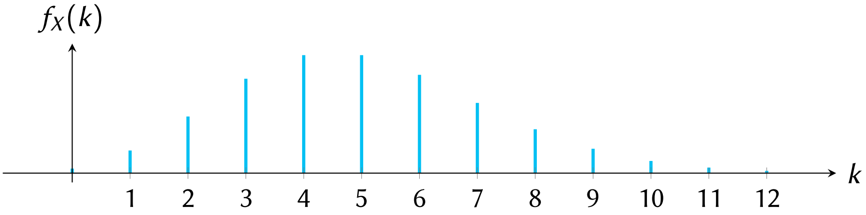


Figura: Función de de densidad de probabilidad de Poisson

¿Dónde aparece?

Se utiliza para describir eventos esporádicos en una población grande, como la mutación de una célula, o los errores de bits en una transmisión de datos.

¡Y hay muchas (muchas) distribuciones más!

- Bernoulli
- Rademacher
- Beta-binomial
- Degenerada en x_0
- Hipergeométrica
- Benford
- Binomial negativa
- Geométrica
- Boltzmann
- Gibbs
- Elíptica asimétrica
- Fractal parabólica
- Polya-Eggenberger
- Skellam
- Yule–Simon
- Zeta
- Zipf–Mandelbrot
- Arcoseno
- Beta
- Coseno elevado
- Irwin-Hall
- Kent
- Kumaraswamy
- Logit-normal
- Normal truncada
- Triangular
- U-cuadrática
- Von Mises-Fisher
- Wigner
- Beta prima
- Birnbaum
- Chi, χ
- χ^2
- Dagum
- F (mi favorita)
- Fréchet
- Gamma
- Erlang
- Gamma-Gompertz
- Gumbel tipo-2
- Lévy
- Log-Cauchy
- Log-gamma
- Log-normal
- Mittag–Leffler
- Nakagami
- Wald
- Pareto
- Tipo III de Pearson
- Bi-exponencial
- Bi-Weibull
- Rice
- T^2 de Hotelling
- Rosin-Rammler
- Z de Fisher
- Behrens–Fisher
- Cauchy
- Chernoff
- Lévy
- Estable geométrica
- Fisher–Tippett
- Gumbel
- Holtsmark
- Landau
- Linnik
- Map-Airy
- Dirichlet
- Ewens
- Balding–Nichols
- Multivariante
- ...

- Para describir el número de eventos que ocurren en un periodo de tiempo
- Para describir el número de intentos necesarios hasta conseguir el primer acierto
- Para predecir tiempos de espera en sistemas telefónicos
- Para estimar la esperanza de vida poblacional
- Para modelar procesos farmacocinéticos (relacionados con la acción de los medicamentos en el cuerpo)
- Para describir la distribución de tamaños de determinadas partículas
- Para análisis genético de poblaciones
- Para pronosticar fenómenos atmosféricos
- Para pronosticar movimientos de la bolsa de valores
- Para calcular los recursos necesarios en una epidemia
- …

- **Variables aleatorias discretas y continuas | Estadística UNED**
  FísicayMates, https://youtu.be/n0T_HcJ7oak
- **Understanding Random Variables - Probability Distributions 1**
  Dr Nic's Maths and Stats, https://youtu.be/lHCpYeFvTs0
- **The Galton Board** (una visualización interesante de cómo aparece la distribución normal en experimentos aleatorios)
  PhysicsFun, https://youtu.be/Vo9Esp1yaC8
- **[*Name of the distribution*]**
  Wikipedia, https://en.wikipedia.org/wiki/Name_of_the_distribution
- **Why "probability of 0" does not mean "impossible"**
  3Blue1Brown, https://youtu.be/ZA4JkHKZM50