Introducción a la probabilidad y estadística CM274

César Lara Avila

17 de septiembre de 2017

https://github.com/C-Lara

7 . Principales variables aleatorias continuas.

Introducción

Aquí introducimos algunas distribuciones continuas fundamentales. Para cada distribución, damos el rango, el pdf, el cdf, y una breve descripción de las situaciones que modela. Todas estas distribuciones dependen de los parámetros que especificamos.

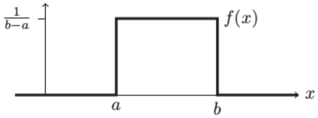
Al mirar a través de cada distribución no trates de memorizar todos los detalles, es preferible centrarse en la forma de cada distribución y lo que hace.

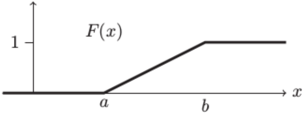
Aunque viene hacia el final, se debe observar con atención la distribución normal, es la distribución más importante definida aquí.

Distribución uniforme

1. Paramétros: a, b.
2. Rango :$[a, b]$.
3. Notación: Uniforme(a, b).
4. Densidad: $f(x) = \frac{1}{(b-a)}$ para $a \leq x \leq b$.
5. Distribución : $F(x) = (x-a)/(b-a)$ para $a \leq x \leq b$.
6. Modelos: Todos los resultados en el rango tienen igual probabilidad (más precisamente todos los resultados tienen la misma densidad de probabilidad).

Gráficos del PDF y el CDF de la distribución uniforme (a, b):





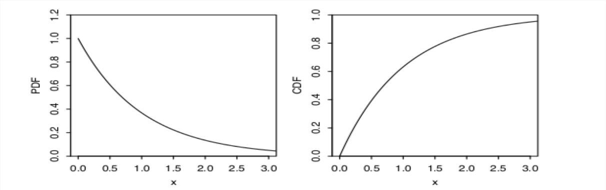
Ejemplos

1. Supongamos que tenemos una cinta métrica con marcas en cada milímetro. Si se mide la longitud de los elementos que miden aproximadamente un metro de longitud, el error de redondeo se distribuirá uniformemente entre -0.5 y 0.5 milímetros.
2. Muchos juegos de mesa utilizan flechas giratorias (spinners) para introducir aleatoriedad. Cuando se hace girar, la flecha se detiene en un ángulo que está uniformemente distribuido entre 0 y 2π radianes.
3. En la mayoría de los generadores de números pseudoaleatorios, el generador básico simula una distribución uniforme y todas las demás distribuciones se construyen mediante la transformación del generador básico.

Distribución exponencial

1. Paramétros: λ.
2. Rango: $[a, \infty)$.
3. Notación: Exponencial(λ).
4. Densidad: $f(x) = \lambda e^{-\lambda x}$ para $0 \leq x$.
5. Distribución : $F(x) = 1 - e^{-\lambda x}$ para $x \geq 0$.
6. Distribución de la cola derecha: $\mathbb{P}(X > x) = 1 - F(x) = e^{-\lambda x}$.
7. Modelos: El tiempo de espera para un proceso continuo para cambiar el estado.

Gráficos del PDF y el CDF de la distribución Exponencial(1):



Ejemplos

1. Si salgo 77 Mass Ave despus de clases y espero el próximo taxi, mi tiempo de espera en minutos se distribuye exponencialmente. Veremos que en este caso λ está dado por uno sobre el promedio de taxis que pasan por minuto (en las tardes de los dias de semana).
2. La distribución exponencial modela el tiempo de espera hasta que un isótopo inestable sufra desintegración nuclear. En este caso, el valor de λ está relacionado con la semivida del isótopo.

Memorylessness

Hay otras distribuciones que también modelan los tiempos de espera, pero la distribución exponencial tiene la propiedad adicional de que no tiene memoria. En el contexto del primer ejemplo, supongamos que la probabilidad de que un taxi llegue dentro de los primeros cinco minutos es p. Si espero cinco minutos y de hecho no llega ningún taxi, entonces la probabilidad de que un taxi llegue dentro de los próximos cinco minutos sigue siendo p.

Formalmente, **memoryless** significa que la probabilidad de esperar t más minutos no se ve afectada por haber esperado s minutos sin incidentes.

En símbolos, $\mathbb{P}(X > s + t | X > s) = \mathbb{P}(X > t)$.

En efecto, desde que $(X > s + t) \cap (X > s) = (X > s + t)$, tenemos

$$\mathbb{P}(X > s + t | X > s) = \frac{\mathbb{P}(X > s + t)}{\mathbb{P}(X > s)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} = \mathbb{P}(X > t).$$

Memorylessness

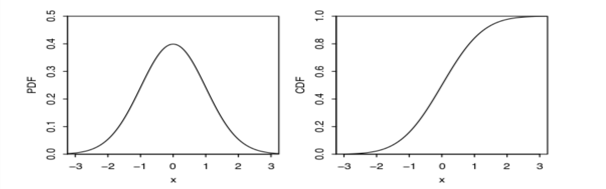
Por el contrario, supongamos que vamos a la estación de metro Miraflores y esperamos al tren entrante. Dado que los trenes están coordinados para seguir un horario (por ejemplo, aproximadamente 12 minutos entre trenes), si esperamos cinco minutos sin ver un tren, entonces hay una probabilidad mucho mayor de que un tren llegará en los próximos cinco minutos.

En particular, el tiempo de espera para el metro no es sin memoria y un mejor modelo modelo sería la distribución uniforme en el rango $[0, 12]$.

Distribución normal

1. Paramétros: μ, σ.
2. Rango: $(-\infty, \infty)$.
3. Notación: $N(\mu, \sigma)$.
4. Densidad: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$.
5. Distribución : $F(x)$ no tiene fórmula, así usamos tablas o funciones como **pnorm** en R, para calcular $F(x)$.
6. Modelos: Error de medición, inteligencia,habilidad, altura, promedios de muchos datos..

Gráficos del PDF y el CDF de la distribución normal estándar $N(0, 1)$:



Distribución normal estándar

La **distribución normal estándar** $N(0,1)$ tiene media 0 y varianza 1. Reservamos Z para la variable aleatoria estándar, $\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ para la densidad normal estándar y $\Phi(z)$ para la distribución normal estándar.

Hay varias propiedades de simetría importantes que se pueden deducir del PDF y CDF de la distribución normal estándar.

1. Simetría del PDF: φ es una función par.
2. Simetría de áreas de cola:

   El área bajo la curva del PDF a la izquierda de -2, que es $P(Z \leq 2) = \Phi(-2)$ por definición, es igual al área a la derecha de 2, que es $\mathbb{P}(Z \geq 2) = 1 - \Phi(2)$. En general, tenemos

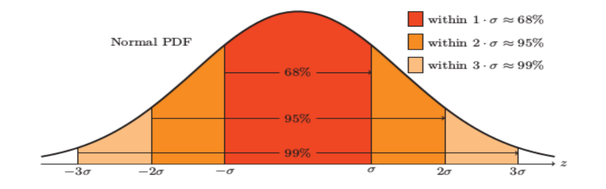
$$\Phi(z) = 1 - \Phi(-z)$$

3. Simetría de Z y $-Z$: Si $Z \sim N(0,1)$ entonces $-Z \sim N(0,1)$.

Probabilidades normales

Para hacer aproximaciones es útil recordar la siguiente regla empírica para tres probabilidades aproximadas

$$\mathbb{P}(-1 \leq Z \leq 1) \approx 0.68, \ \mathbb{P}(-2 \leq Z \leq 2) \approx 0.95, \ \mathbb{P}(-3 \leq Z \leq 3) \approx 0.99,$$



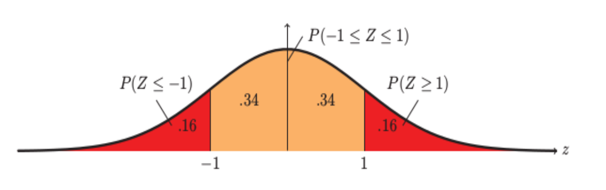
Podemos usar la simetría de la distribución normal estándar alrededor de $x = 0$ para hacer algunos cálculos.

Ejemplos

Calculemos $\Phi(1)$,

Usando el resultado,$\mathbb{P}(-1 \leq Z \leq 1) \approx 0.68$, $\Phi(1) = \mathbb{P}(Z \leq 1)$.

En la figura, las dos colas (en rojo) han combinado el área $1 - 0.68 = 0.32$. Por simetría la cola izquierda tiene área 0.16 (la mitad de 0.32), así que $\mathbb{P}(Z \leq 1) \approx 0.68 + 0.16 = 0.84$.



Distribución de Pareto

1. Paramétros: $m > 0$ y $\alpha > 0$.
2. Rango: $[m, \infty)$.
3. Notación: Pareto(m, α).
4. Densidad: $f(x) = \frac{\alpha m^\alpha}{x^{\alpha+1}}$.
5. Distribución :

$$F(x) = 1 - \frac{m^\alpha}{x^\alpha}, \text{ para } x \geq m.$$

6. Distribución de la cola: $\mathbb{P}(X > x) = m^\alpha / x^\alpha$, para $x \geq \alpha$.
7. Modelos: La distribución de Pareto modela una ley de potencia, donde la probabilidad de que ocurra un evento varía como una potencia de algún atributo del evento.

   Muchos fenómenos siguen una ley de potencia, como el tamaño de los meteoros, los niveles de ingresos en una población y los niveles de población en las ciudades.