

Intro a la Probabilidad y estadística



Martes y Jueves Aula B17
Dra Ana Georgina Flesia

Cuantiles

Definición

1. Sea $p \in (0, 1)$ y sea X una variable aleatoria continua sobre (Ω, \mathcal{A}, P) con función de distribución acumulada F entonces se llama cuantil p de X al valor $\eta(p)$ tal que

$$p = F(\eta(p))$$

2. Si f es una función de densidad de X , entonces

$$p = F(\eta(p)) = \int_{-\infty}^{\eta(p)} f(t) dt$$

Cuantiles

Propiedades

Los cuantiles suelen usarse por grupos que dividen la distribución en partes iguales; entendidas estas como intervalos que comprenden la misma proporción de valores. Los más usados son:

1. Los cuartiles, que dividen a la distribución en cuatro partes (corresponden a los cuantiles 0,25; 0,50 y 0,75);
2. Los quintiles, que dividen a la distribución en cinco partes (corresponden a los cuantiles 0,20; 0,40; 0,60 y 0,80);
3. Los deciles, que dividen a la distribución en diez partes;
4. Los percentiles, que dividen a la distribución en cien partes.

Percentiles y cuantiles

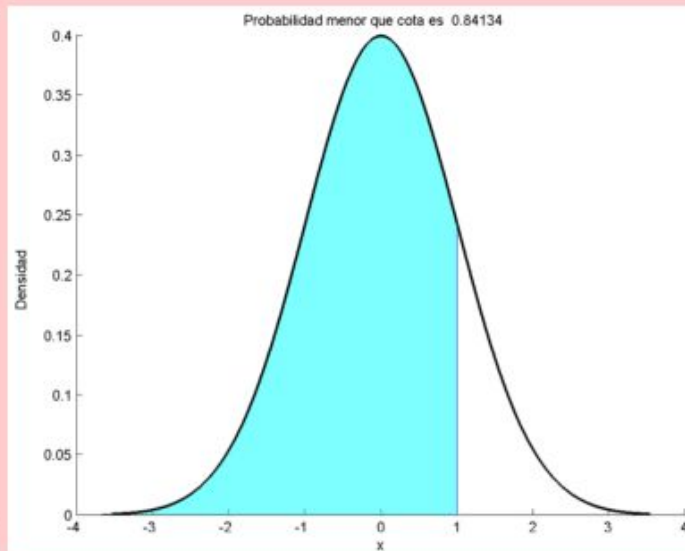
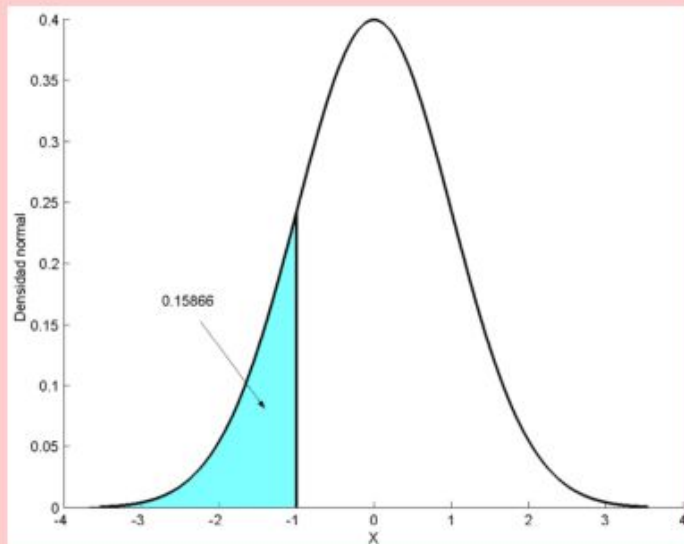
Propiedad

Sea X una variable aleatoria continua con función de distribución acumulada F , y sea $\eta_X(p)$ el cuantil p de X , con $p \in [0; 1]$. Si $Y = aX + b$ con $a \neq 0$ y $b \in \mathbb{R}$, entonces

$$\eta_Y(p) = \begin{cases} a\eta_X(p) + b & \text{si } a > 0 \\ a\eta_X(1 - p) + b & \text{si } a < 0 \end{cases}$$

Ejemplo

Cuantiles



Esperanza y Varianza

- Recordemos que si X era una variable aleatoria discreta, tal que $\sum_x |x|p_X(x) < \infty$ entonces X tenía esperanza y

$$E(X) = \sum_x xp_X(x)$$

- En forma análoga, definimos esperanza de variables aleatorias continuas de la siguiente forma.

Esperanza y Varianza

Definición

Sea X un variable aleatoria continua con densidad f . Decimos que X tiene esperanza finita si

$$\int_{-\infty}^{\infty} |x|f(x)dx < \infty$$

y en ese caso se define la esperanza como

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx < \infty$$

y la varianza como

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

La desviación estándar es la raíz cuadrada positiva de la varianza,
 $\sigma = \sqrt{V(X)}$.

Esperanza y Varianza

Propiedades

Sea X un variable aleatoria continua sobre (Ω, \mathcal{A}, P) con densidad de probabilidad f y $E(X^2) < \infty$, entonces

1. $0 \leq V(X) = E(X^2) - \mu^2$ donde $\mu = E(X)$.
2. $E(aX + b) = aE(X) + b$ y $V(aX + b) = a^2V(X)$, para todo $a, b \in \mathbb{R}$
3. Si $h(X)$ es una variable aleatoria continua, entonces

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

siempre que

$$\int_{-\infty}^{\infty} |h(x)|f(x)dx < \infty$$

Distribución uniforme

Distribución Uniforme

Distribución Uniforme

Sean a y b constantes con $a < b$. La densidad uniforme sobre el intervalo (a, b) es la densidad f definida por

$$f(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \text{en otro caso} \end{cases}$$

La distribución correspondiente a f está dada por

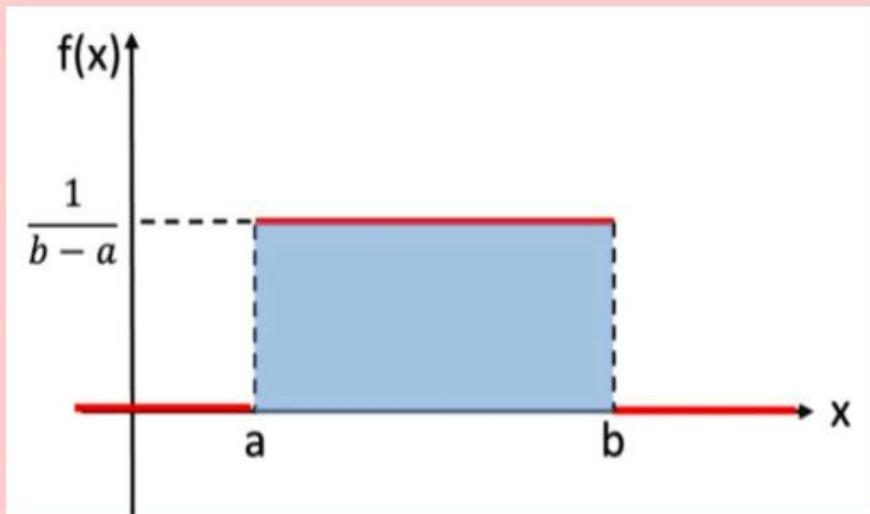
$$F(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x \leq b \\ 1 & x > b \end{cases}$$

La esperanza y varianza de X con distribución uniforme es

$$E(X) = \frac{a+b}{2} \quad \text{Var}(X) = \frac{(b-a)^2}{12}$$

Distribución Uniforme

Densidad $\mathcal{U}[a, b]$



Distribución Uniforme

Ejemplo

Supongamos que Mario usualmente va a visitar a María entre las 8 y 8:30 de la noche. Si el horario de llegada es aleatorio, ¿cuál es la probabilidad de que Mario llegue después de las 8:25?.

Distribución Uniforme

Ejemplo: Resolución

En este caso, queremos calcular la probabilidad de que X , la variable que modela el tiempo de llegada de Mario, tome valores entre las 8:25 y las 8:30 hs. Si X tiene distribución uniforme entre las 8 y las 8:30 hs, su densidad puede escribirse como

$$f(x) = \begin{cases} \frac{1}{30} & 0 < x < 30 \\ 0 & \text{en otro caso} \end{cases}$$

donde x son los minutos después de las 8 en los que Mario llega. Por lo tanto

$$P(25 \leq X \leq 30) = \int_{25}^{30} \frac{1}{30} dx = \frac{30 - 25}{30} = \frac{5}{30} = \frac{1}{6} = 0.166$$

Distribución Uniforme

Ejemplo

Sea X con distribución uniforme en el $(0, 1)$. Encuentre la densidad de $Y = -\frac{1}{\lambda} \log(1 - X)$ para todo $\lambda > 0$.

Distribución Uniforme

Ejemplo: Resolución

Observemos primero que Y es una variable positiva y por lo tanto $F_Y(y) = 0$ para todo $y \leq 0$. Para $y > 0$ se tiene

$$\begin{aligned}F_Y(y) = P(Y \leq y) &= P\left(-\frac{1}{\lambda} \log(1 - X) \leq y\right) \\&= P(\log(1 - X) \geq -\lambda y) \\&= P(1 - X \geq e^{-\lambda y}) \\&= 1 - e^{-\lambda y}\end{aligned}$$

por lo cual $f_Y(y) = F'_Y(y) = \lambda e^{-\lambda y}$ para $y > 0$. La densidad de Y está dada por

$$f_Y(y) = \begin{cases} \lambda e^{-\lambda y} & y > 0 \\ 0 & y \leq 0 \end{cases}$$

Distribución Uniforme

Propiedad

Sea X una variable aleatoria continua con función densidad f_X y función de distribución acumulada F_X , tal que F_X es estrictamente creciente en un intervalo I (que puede ser no acotado), nula a la izquierda de I y vale 1 a la derecha de I . Entonces $F^{-1}(x)$ está bien definido para cada valor $x \in I$.

1. Sea $Z = F(X)$, entonces Z tiene distribución uniforme en $[0, 1]$.
2. Si U tiene distribución uniforme en el $[0, 1]$, entonces $Y = F^{-1}(U)$ tiene como distribución acumulada a la función F .

Distribución Uniforme

Propiedad: Resolución

Observemos que si $z \in [0, 1]$

$$P(Z \leq z) = P(F(X) \leq z) = P(X \leq F^{-1}(z)) = F(F^{-1}(z)) = z$$

y como ésta es la función de distribución acumulada de una variable uniforme, sale el resultado.

Observemos ahora que

$$P(Y \leq y) = P(F^{-1}(U) \leq y) = P(U \leq F(y)) = F(y)$$

por lo cual X e Y tienen la misma función de distribución acumulada.

Distribución Uniforme

Propiedad

Si $X \sim [0, 1]$ entonces

- ▶ $\eta(p) = p$ con $p \in (0, 1)$
- ▶ $E(X) = \frac{1}{2}$ y $V(X) = \frac{1}{12}$
- ▶ Para cualquier $z < b$ números reales sea una variable aleatoria $Y = (b - a)X + a$. Entonces
 - $Y \sim U[a, b]$
 - $\eta_Y(p) = (b - a)p + a$ para cualquier $p \in (0, 1)$
 - $E(Y) = \frac{a+b}{2}$ y $Var(Y) = \frac{(b-a)^2}{12}$

Distribución Uniforme

Ejemplo

El tiempo (en minutos) que tarda un camión en realizar un viaje (ida y vuelta), para transportar concreto hacia una obra en construcción, tiene distribución uniforme en el intervalo $[50; 70]$.

1. ¿Cual es la probabilidad de que la duración del viaje de un camión sea menor de 65 minutos?
2. ¿Cual es la probabilidad de que la duración del viaje de un camión sea mayor de 65 minutos dado que se sabe que fue mayor a 55 minutos?
3. Dar el valor medio y desviación estándar del tiempo que tarda un camión en realizar un viaje (ida y vuelta) para transportar concreto hacia la obra en construcción.
4. Suponga que los tiempos que tardan cada uno de tres camiones en realizar un viaje (ida y vuelta) para transportar concreto hacia la obra en construcción son independientes entre si. ¿Cual es la probabilidad de que por lo menos uno de ellos tarde mas de 65 minutos?

Ejemplo: Resolución

Vamos a resolver cada una de las preguntas paso a paso. Sabemos que el tiempo de viaje tiene una distribución uniforme en el intervalo $[50, 70]$.

1. Probabilidad de que la duración del viaje sea menor de 65 minutos

- ▶ Para una distribución uniforme en el intervalo $[a, b]$, la probabilidad de que una variable X sea menor que un valor x dentro del intervalo se calcula como:

$$P(X < x) = \frac{x - a}{b - a}$$

- ▶ Aquí, $a = 50$, $b = 70$, y queremos la probabilidad de que $X < 65$:

$$P(X < 65) = \frac{65 - 50}{70 - 50} = \frac{15}{20} = 0.75$$

- ▶ Así que la probabilidad de que la duración del viaje sea menor de 65 minutos es 0.75 o 75%.

Ejemplo: Resolución

2. Probabilidad de que la duración del viaje sea mayor de 65 minutos dado que se sabe que fue mayor a 55 minutos

- Queremos encontrar $P(X > 65 \mid X > 55)$. Usamos la fórmula de la probabilidad condicional:

$$P(X > 65 \mid X > 55) = \frac{P(X > 65 \cap X > 55)}{P(X > 55)}$$

- Sabemos que $P(X > 65 \cap X > 55) = P(X > 65)$. Primero, calculemos $P(X > 65)$:

$$P(X > 65) = \frac{70 - 65}{70 - 50} = \frac{5}{20} = 0.25$$

Ejemplo: Resolución

- ▶ Ahora calculemos $P(X > 55)$:

$$P(X > 55) = \frac{70 - 55}{70 - 50} = \frac{15}{20} = 0.75$$

- ▶ Entonces:

$$P(X > 65 \mid X > 55) = \frac{0.25}{0.75} = \frac{1}{3} \approx 0.3333$$

La probabilidad de que la duración del viaje sea mayor de 65 minutos dado que se sabe que fue mayor a 55 minutos es $\frac{1}{3}$ o aproximadamente 33.33%.

Ejemplo: Resolución

3. Valor medio y desviación estándar del tiempo de viaje

- ▶ Para una distribución uniforme en el intervalo $[a, b]$, el valor medio μ y la desviación estándar σ se calculan como:
- ▶ **Valor medio (esperanza):**

$$\mu = \frac{a + b}{2}$$

- ▶ **Desviación estándar:**

$$\sigma = \sqrt{\frac{(b - a)^2}{12}}$$

Ejemplo: Resolución

- Para $a = 50$ y $b = 70$:

$$\mu = \frac{50 + 70}{2} = \frac{120}{2} = 60$$

$$\sigma = \sqrt{\frac{(70 - 50)^2}{12}} = \sqrt{\frac{400}{12}} = \sqrt{33.\bar{3}} \approx 5.77$$

- El valor medio del tiempo de viaje es 60 minutos y la desviación estándar es aproximadamente 5.77 minutos.

Ejemplo: Resolución

4. Probabilidad de que al menos uno de los tres camiones tarde más de 65 minutos

- ▶ Primero, calculamos la probabilidad de que un camión tarde 65 minutos o menos. Sabemos que $P(X \leq 65) = 0.75$.
- ▶ Entonces, la probabilidad de que un camión tarde más de 65 minutos es $1 - 0.75 = 0.25$.
- ▶ Para tres camiones independientes, la probabilidad de que ninguno tarde más de 65 minutos es:

$$P(\text{ninguno tarda más de 65}) = (0.75)^3 = 0.421875$$

Ejemplo: Resolución

- ▶ La probabilidad de que al menos uno de los tres camiones tarde más de 65 minutos es:

$$\begin{aligned}P(\text{al menos uno tarda más de 65}) &= 1 - P(\text{ninguno tarda más de 65}) \\ &= 1 - 0.421875 = 0.578125\end{aligned}$$

- ▶ La probabilidad de que al menos uno de los tres camiones tarde más de 65 minutos es aproximadamente 0.578 o 57.8%.

distribución exponencial

Distribución Exponencial

Definición

Sea λ un valor real positivo. La densidad exponencial es la densidad f definida por

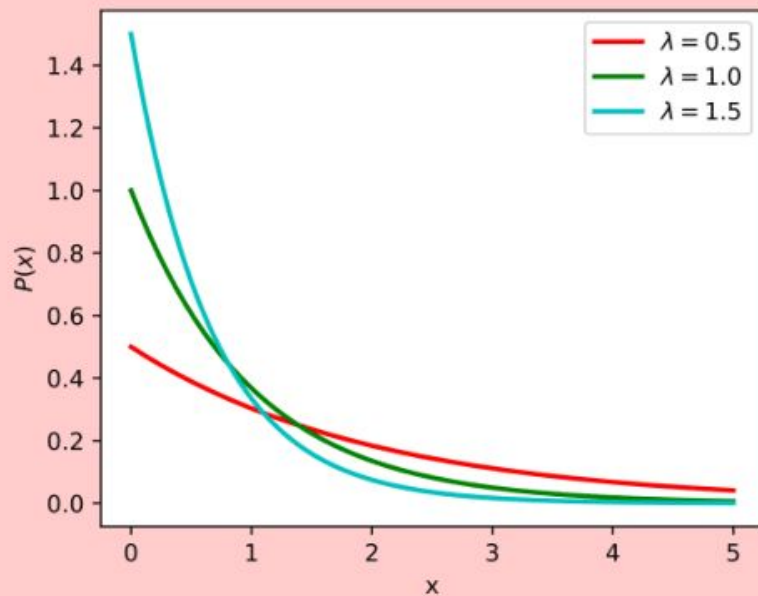
$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

La esperanza y varianza de X con distribución exponencial de parámetro λ es

$$E(X) = \frac{1}{\lambda} \quad \text{Var}(X) = \frac{1}{\lambda^2}$$

Distribución Exponencial

Densidades



Distribución Exponencial

Ejemplo

Como vimos en el ejemplo anterior, es la densidad de la variable $Y = -\frac{1}{\lambda} \log(1 - X)$, cuando X es uniforme en $(0, 1)$.

La función de distribución acumulada correspondiente a f es

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

Distribución Exponencial

Ejemplo

Sea X la duración en horas de las lámparas almacenadas en un depósito, y supongamos que pueden modelarse con la distribución exponencial con parámetro 0.01.

1. Hallar la función de distribución acumulada de X , la duración de una lámpara elegida al azar del lote.
2. ¿Cuál es la probabilidad de que una lámpara dure más de 120 horas?
item ¿Cuántas horas se espera que dure una lámpara?
3. Si se tiene un lote de 10000 lámparas ¿Cuál es la probabilidad de que 10 lámparas elegidas al azar del lote duren menos de 70 horas?

Ejemplo Resolución

- ▶ La distribución acumulada de X es la distribución de una exponencial, y se calcula aplicando directamente la definición:

$$P(X \leq x) = \int_0^x \lambda \exp(-\lambda t) dt = \lambda \frac{\exp(-\lambda t)}{-\lambda} \Big|_0^x = 1 - \exp(-\lambda x) \quad x > 0$$

- ▶ En el caso de $\lambda = 0.01$ resulta

$$P(X \leq x) = 1 - \exp(-0.01x)$$

- ▶ Usando esta fórmula podemos calcular la probabilidad de que una lámpara elegida al azar del depósito dure más de 120 hs.

$$\begin{aligned} P(X > 120) &= 1 - P(X \leq 120) = 1 - (1 - \exp(-0.01120)) \\ &= \exp(-0.01120) = \exp(-1.2) = 0.3011 \end{aligned}$$

Ejemplo Resolución

Duración esperada de las lámparas

- ▶ Para una distribución exponencial de parámetro λ , el valor medio μ se calculan como:
- ▶ **Valor medio (esperanza):**

$$\mu = \frac{1}{\lambda}$$

- ▶ Entonces se espera que dure 100 horas.

Ejemplo Resolución

Si se tiene un lote de 10000 lámparas ¿Cuál es la probabilidad de que 10 lámparas elegidas al azar del lote duren menos de 70 horas?

- ▶ Lote es realmente grande, comparado con la muestra
- ▶ Suponer que extraemos con reposición, (pues la probabilidad de elegir dos veces la misma lámpara es casi nula).
- ▶ Entonces podemos suponer que tenemos un experimento binomial, donde en cada extracción tenemos éxito si la lámpara dura menos de 70 horas.

Ejemplo Resolución

- ▶ La variable Y , número de lámparas que duran menos de 70 horas, es una variable binomial de parámetros $n = 10$ y $p = P(X < 70)$.
- ▶ Calculemos esta probabilidad,

$$P(X < 70) = 1 - \exp(-0.0170) = 1 - \exp(-0.7) = 0.5034$$

- ▶ Entonces, la probabilidad de que las 10 lámparas elegidas al azar del lote duren menos de 70 horas es

$$P(Y = 10) = \binom{10}{10} 0.5034^{10} (1 - 0.5034)^{10-10} = 0.5034^{10} = 0.00104$$

Distribución Exponencial

Ejemplo

Sea X una variable que tiene densidad exponencial con parámetro λ . Encuentre la densidad de $Y = X^{1/\beta}$, con $\beta \neq 0$.

Ejemplo Resolución

- ▶ Si X tiene distribución exponencial de parámetro λ entonces su densidad es $f_X(x) = \lambda e^{-\lambda x}$, si $x > 0$, y cero si no.
- ▶ Supongamos $\beta > 0$, entonces $g(x) = x^\beta$ es una función creciente y si $y > 0$

$$P(Y \leq y) = P(X^{1/\beta} \leq y) = P(X \leq y^\beta) = F_X(y^\beta) = 1 - e^{-\lambda y^\beta}$$

- ▶ entonces

$$f_Y(y) = F'_X(y) = \beta \lambda y^{\beta-1} e^{-\lambda y^\beta}$$