Variables Aleatorias Parte II

Variables aleatorias continuas

- Como ya hemos dicho las variables aleatorias continuas toman valores en intervalos o áreas.
- Lo más habitual es que estas variables tengan función de distribución continua y derivable (salvo a los más en una cantidad finita o numerable de puntos:-)).
- En lo que sigue supondremos que la función de distribución de variables aleatorias continuas cumplen estas propiedades.
- Notemos que si X es una v.a. con función de distribución continua se tiene que $P(X = x_0) = F_X(x_0) - F(x_0^-) = 0$. Por lo que no tiene sentido definir ''*función de probabilidad*''.

- En general tendremos que $P(X < x_0) = P(X \leqslant x_0)$.
- Por otra parte podemos utilizar una regla parecida del cociente entre casos favorables y casos posibles de Laplace pero en este caso el conteo se hace por la "medida" de los casos posibles partida por la "medida" de los casos favorables.
- Veamos un ejemplo de v.a. continua, que ampliaremos en el tema siguiente, en el que se utilizan todos estos conceptos.

Ejemplo: Distribución uniforme en el intervalo unidad.

Supongamos que lanzamos un dardo a una diana de radio 1, de forma que sea "*equiprobable*" cualquier distancia al centro[1]. Consideremos la v.a. continua X =distancia al centro.

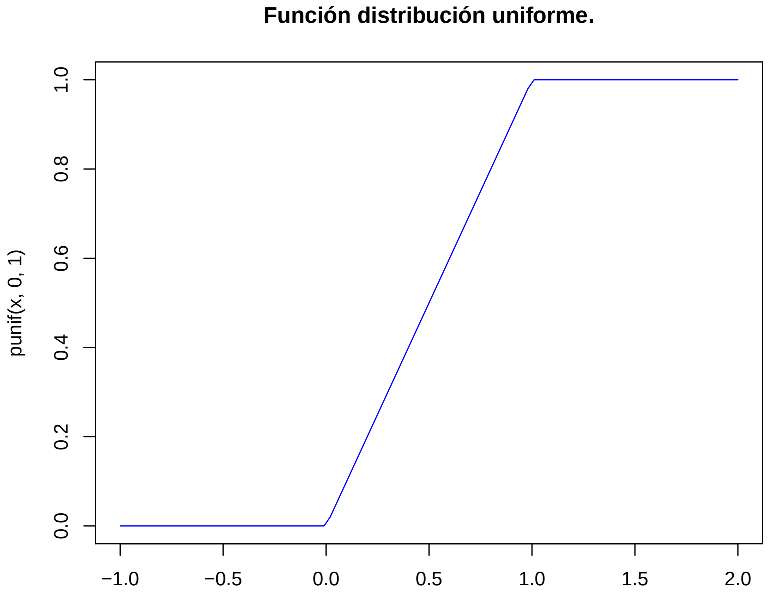
$$F_X(x) = \begin{cases} 0 & \text{si } x \leqslant 0 \\ x & \text{si } 0 < x < 1 \\ 1 & \text{si } x \geqslant 1 \end{cases}$$

Ya que

- C.F. ''*longitud favorable*'' $x - 0$
- C.P. ''*longitud posible*'' $1 - 0$
- Luego $P(X \leqslant x) = \frac{x-0}{1-0} = x$

[1] !Cuidado! esto no es equivalente a que cualquier punto de la diana sea "*equiprobable*":-).

```r
curve(punif(x,0,1),xlim=c(-1,2),col="blue",
      main="Función distribución uniforme.")
```



En las variables continuas los sucesos del tipo $\{X \leqslant x\}$ y $\{X < x\}$ tendrán la misma probabilidad, y otros tipos de sucesos similares también, algunas de estas propiedades se explicitan en la siguiente proposición.

Propiedades

Dada una v.a. continua X se tiene que:

a) $P(X \leqslant b) = P(X < b)$
b) $P(X < b) = P(X < a) + P(a < X < b)$
c) $P(a < X < b) = P(X < b) - P(X < a)$

Demostración:

b) $\{X < a\} \cap \{a < X < b\} = \emptyset$
$\{X < a\} \cup \{a < X < b\} = \{X < a\}$ entonces

$$\begin{aligned} P(X \leqslant b) =& P(\{X < a\} \cup \{a < X < b\}) \\ &= P(X < a) + P(a < X < b) \end{aligned}$$

a) $P(X < b) = P(X < b) + P(X = b) = P(X < b)$
c) Ídem que b) aplicando a).

Las propiedades anteriores y combinaciones de ellas se pueden escribir utilizando la función de distribución de X:

Propiedades

Dada una variable aleatoria continua se tiene que:

a) $F_X(b) = F_X(a) + P(a < X < b)$
b) $P(a < X < b) = F_X(b) - F_X(a)$
c) $P(a \leqslant X \leqslant b) = F_X(b) - F_X(a)$

Ejemplo

Demostración: ejercicio.
En los dardos:

$$P(0.25 < X < 0.3) = F_X(0.3) - F_X(0.25) =$$

$$= 0.3 - 0.25 = 0.05$$

Una función $f : \mathbb{R} \to \mathbb{R}$ es una función de densidad sobre $\mathbb{R}$ si cumple que

a) $f_X(x) \geqslant 0$ para todo $x \in \mathbb{R}$.
b) f es continua salvo a lo más en una cantidad finita de puntos sobre cada intervalo acotado de $\mathbb{R}$.
c) $\int\limits_{-\infty}^{+\infty} f_X(x)dx = 1$.

Definición

Sea X una v.a. con función de distribución F_X. Sea $f : \mathbb{R} \to \mathbb{R}$ una función de densidad tal que

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt. \text{ para todo } x \in \mathbb{R}.$$

Entonces X es una variable aleatoria continua y f_X es la densidad de la v.a. X.
El conjunto $D_X = \{x \in \mathbb{R} | f_x(x) > 0\}$ recibe el nombre de soporte o dominio de la variable aleatoria continua y se interpreta su conjunto de resultados posibles.

En nuestra diana la función f es una densidad

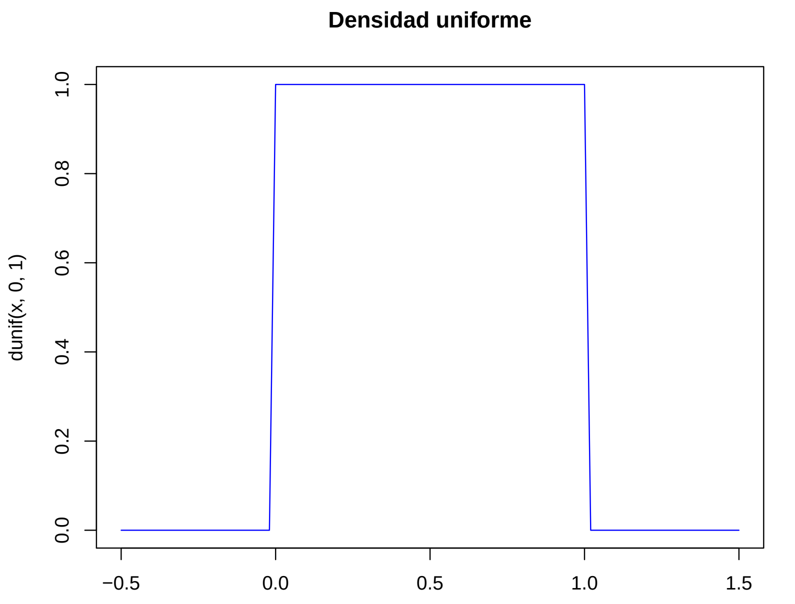
$$f_X(x) = \begin{cases} 0 & \text{si } x \leqslant 0 \\ 1 & \text{si } 0 < x < 1 \\ 0 & \text{si } 1 \leqslant x \end{cases}$$

que es la densidad de X, en efecto:

- Si $x \leqslant 0$ entonces $\int_{-\infty}^{x} f_X(t)dt = 0$.
- Si $0 \leqslant x \leqslant 1$ entonces $\int_{-\infty}^{x} f_X(t)dt = \int_0^x 1dt = x$.
- Si $x \geqslant 1$ entonces $\int_{-\infty}^{x} f_X(t)dt = \int_0^1 1dt = 1$.

Por lo tanto $F_X(x) = \int_{-\infty}^{x} f_X(t)dt$ para todo $x \in \mathbb{R}$.


```
curve(dunif(x,0,1),xlim=c(-0.5,1.5),col="blue",
      main="Densidad uniforme")
```



La función de densidad nos permite calcular diversas probabilidades.

Propiedades

Sea X una v.a. continua con función de distribución F_X y de densidad f_X, entonces

a)

$$P(a < X < b) = P(a < X \leqslant b) = P(a \leqslant X < b) = P(a \leqslant X \leqslant b) = \int_a^b f_X(x)dx.$$

b) *Si A es un conjunto adecuado de $\mathbb{R}$ entonces* $P(X \in A) = \int_A f(x)dx = \int_{A\cap D_X} f(x)dx$.

Propiedades

Sea X una v.a. continua con función de distribución F_X y de densidad f_X, entonces:

a) *Si f_x es continua en un punto x, F_X es derivable en ese punto y $F'_X(x) = f_X(x)$.*

b) $P(X = x) = 0$ *para todo* $x \in \mathbb{R}$.

Ejemplo

Sea X = tiempo de ejecución de un proceso. Se supone que X sigue una distribución uniforme en dos unidades de tiempo, si tarda más el proceso se cancela. Entonces

$$F_X(x) = P(X \leqslant x) = \frac{CF}{CP} = \frac{x}{2}$$

Luego su función de distribución es:

$$F_X(x) = \begin{cases} 0 & \text{si } x \leqslant 0 \\ \frac{x}{2} & \text{si } 0 < x < 2 \\ 1 & \text{si } 2 \leqslant x \end{cases}$$

mientras que su función de densidad es:

$$f_X(x) = F'_X(x) = \begin{cases} 0 & \text{si } x \leqslant 0 \\ \frac{1}{2} & \text{si } 0 < x \leqslant 2 \\ 0 & \text{si } 2 \leqslant x \end{cases}$$

Efectivamente

- $f_X(x) \geqslant 0$, y tiene un conjunto finito de discontinuidades.
- $F_X(x) = \int_{-\infty}^{x} f_X(t)dt$. para todo $x \in \mathbb{R}$ (ejercicio, resolverlo gráficamente.)
- $\int\limits_{-\infty}^{+\infty} f_X(x)dx = \int\limits_{0}^{2} \frac{1}{2}dx = \frac{x}{2} \mid_0^2 == \frac{2}{2} - \frac{0}{2} = 1.$

Ejercicio: Calcular la probabilidad de que uno de nuestros procesos tarde más de una unidad de tiempo en ser procesado. Calcular también la probabilidad de que dure entre 0.5 y 1.5 unidades de tiempo.

Los mismos comentarios y definiciones que se dieron en la sección correspondiente del tema de estadística descriptiva son aplicables aquí. Así que sólo daremos las definiciones, la forma de cálculo y algunos ejemplos.

Sea X una v.a. continua con función de densidad $f_X(x)$ entonces:

- su esperanza es : $E(X) = \int_{-\infty}^{+\infty} x f_X(x) dx$.
- Si $g(x)$ es una función de la variable X entonces

$$E(g(X)) = \int_{-\infty}^{+\infty} g(x) f_X(x) dx.$$

- $Var(X) = \sigma_X^2 = E((X-\mu_X)^2) = \int_{-\infty}^{+\infty} (x-\mu_X)^2 f_X(x) dx$.
- A $\sigma_X = +\sqrt{\sigma_X^2}$ se le denomina desviación típica de X.

Propiedades

- $\sigma_X^2 \geqslant 0$
- $Var(cte) = E(cte^2) - (E(cte))^2 = cte^2 - cte^2 = 0$
- $Var(x) = E(X^2) - \mu_X^2 = \int_{-\infty}^{+\infty} x^2 f_X(x) dx - \mu_X^2$.
- El mínimo de $E((X-C)^2)$ se alcanza cuando $C = E(X)$ y es $Var(X)$.

Ejemplos Calcular μ_X y σ_X^2 en el dardo.
Resultado $\mu_X = \frac{1}{2}$, $E(X^2) = \frac{1}{3}$, $Var(X) = \frac{1}{12}$.

Sea X una v.a. continua con $E(X) = \mu_X$ y $Var(X) = \sigma_X^2$ sea $Y = a + bX$, donde $a, b \in \mathbb{R}$, es una nueva v.a. continua obtenida mediante una transformación lineal de X. Se verifican las mismas propiedades que en el caso discreto:

- $E(Y) = E(a + bX) = a + bE(X)$
- $Var(Y) = Var(a + bX) = b^2 Var(X)$
- $\sigma_Y = |b|\sigma_X$
- $Z = \frac{X-\mu_X}{\sigma_X}$ es una transformación lineal de X de forma que
$$E(Z) = 0 \text{ y } Var(Z) = 1$$

Ejemplo En una empresa de venta de vinos por internet, sea X = número de litros de vino del país vendidos en un año. Supongamos que sabemos que $E(X) = 10000$ y que $Var(X) = 100$ Supongamos que los gastos fijos de distribución son 50000 y el beneficio por litro es de 10 pts por botella. Definimos $T = 10X - 50000$ que será el beneficio después de gastos entonces:

$$E(T) = 10E(X) - 50000 = 50000$$

y

$$Var(T) = 10^2 VAR(X) = 10000$$

Muchas variables aleatorias son funciones de otras v.a. En lo que sigue resumiremos diversas técnicas para dada una v.a. X y una transformación $Y = h(X)$ encontrar F_Y a partir de F_X.

Propiedades

Sea X una v.a. discreta con $X(\Omega) = \{x_1, x_2, \ldots, x_n, ..\}$ y sea $h : \mathbb{R} \to \mathbb{R}$ una aplicación. Entonces $Y = h(X)$ es también una v.a. discreta. Además si P_X y F_X son las funciones de probabilidad y de distribución de X entonces

a) $P_Y(y) = \displaystyle\sum_{x_i | h(x_i) = y} P_X(x_i).$

b) $F_Y(y) = \displaystyle\sum_{x_i | h(x_i) \leqslant y} P_X(x_i).$

Desafortunadamente este caso no es tan sencillo como el anterior, pues la transformación de una v.a. continua puede ser continua, discreta, mixta ...

Propiedades

Sea X una v.a. continua cuya función de densidad es f_X. Sea $h : \mathbb{R} \to \mathbb{R}$ una aplicación estrictamente monótona y derivable, tal que $h'(x) \neq 0$ para todo $x \in \mathbb{R}$. Sea $Y = h(X)$ la transformación de X por h. Entonces Y es una v.a. continua con función de densidad

$$f_Y(y) = \left. \frac{f_X(x)}{|h'(x)|} \right|_{x=h^{-1}(y)}$$

Propiedades

Sea X una v.a. continua cuya función de densidad es f_X. Si $h : \mathbb{R} \to \mathbb{R}$ es una aplicación, no necesariamente monótona, pero sí derivable con derivada no nula, y si la ecuación $h(x) = y$ tiene un número finito de soluciones $x_1, x_2, .., x_n$ entonces:

$$f_Y(y) = \sum_{k=1}^{n} \left. \frac{f_X(x)}{|h'(x)|} \right|_{x=x_k}$$

Cuando no podamos aplicar las propiedades anteriores intentaremos calcular primero la función de distribución de la transformación y luego su densidad.
Notemos que en general si $Y = g(X)$ es una v.a. transformación de la v.a. X entonces

$$F_Y(y) = P(Y \leqslant y) = P(g(X) \leqslant y)$$

Por ejemplo si g es estrictamente creciente y cont.

$$F_Y(y) = P(g(X) \leqslant y) = P(X \leqslant g^{-1}(y)) = F_X(g^{-1}(y))$$

y si g es estrictamente decreciente y cont.

$$F_Y(y) = P(g(X) \leqslant y) = P(X \geqslant g^{-1}(y)) = 1 - F_X(g^{-1}(y))$$

Desigualdades de Markov y de Chebyshef

- Veremos en esta sección distintas desigualdades que acotan determinadas probabilidades de una variable aleatoria.
- Estas desigualdades sirven en algunos casos para acotar probabilidades de determinados sucesos.
- También son útiles desde el punto de vista teórico, por ejemplo para justificar que la varianza es una mediada de la dispersión de los datos.

Propiedades

Sea X una v.a. positiva con $E(X)$ finita. Entonces $P(X \geqslant a) \leqslant \frac{E(X)}{a}$ para todo $a > 0$.

Demostración:
Si X es continua y sólo toma valores positivos

$$\begin{aligned} E(X) &= \int_{-\infty}^{+\infty} x f_X(x) dx = \int_{0}^{+\infty} x f_X(x) dx \\ &= \int_{0}^{a} x f_X(x) dx + \int_{a}^{+\infty} x f_X(x) dx \\ &\geqslant \int_{a}^{+\infty} x f_X(x) dx \geqslant a \int_{a}^{+\infty} f_X(x) dx \\ &= a \cdot P(X \geqslant a) \end{aligned}$$

de donde se sigue que

$$P(X \geqslant a) \leqslant \frac{E(X)}{a}.$$

Propiedades

Sea X una v.a. con $E(X)$ finita entonces para todo $a > 0$

$$P(|X| \geqslant a) \leqslant \frac{E(|X|)}{a}$$

Propiedades

Sea X una v.a.con $E(X) = \mu$ y $Var(X) = \sigma^2$ entonces para todo $a > 0$

$$P(|X - \mu| \geqslant a) \leqslant \frac{\sigma^2}{a^2}$$

Demostración:
Apliquemos la consecuencia de la desigualdad de Markov a la v.a. no negativa

$$Y^2 = (X - \mu)^2$$

entonces

$$P(Y^2 \geqslant a^2) \leqslant \frac{E(Y^2)}{a^2} = \frac{E((X-\mu)^2)}{a^2} = \frac{Var(X)}{a^2} = \frac{\sigma^2}{a^2}$$

Por otra parte

$$P(Y^2 \geqslant a^2) = P(|Y| \geqslant a) = P(|X - \mu| \geqslant a)$$

hecho que, junto con la desigualdad anterior, demuestra el resultado.

Observación: Supongamos que X es una v.a. con $Var(X) = 0$ entonces.
Aplicando la desigualdad anterior

$$P(|X - E(X)| \geqslant a) = 0$$

para todo $a > 0 loqueimplicaque$

$$P(X = E(X)) = 1$$

Por lo que probabilidad de que X sea constantemente $E(X)$ es 1.
Lo que nos confirma la utilidad de la varianza es una medida de la dispersión de los datos.

Ejemplo

Se sabe que el tiempo de respuesta medio y la desviación típica de un sistema multiusuario son 15 y 3 u.t. respectivamente. Entonces:

$$P(|X - 15| \geqslant 5) \leqslant \frac{9}{25} = 0.36.$$

Si substituimos a por $a \cdot \sigma$ en la desigualdad de Chebyshef.

Nos queda:

$$P(|X-\mu| \geqslant a\sigma) \leqslant \frac{\sigma^2}{(a\sigma)^2} = \frac{1}{a^2}.$$

Que es otra manera de expresar la desigualdad de Chebyshef. La desigualdad de Chebyshef también se puede escribir de al menos dos maneras más:

$$P(\mu - a \leqslant X \leqslant \mu + a) \geqslant 1 - \frac{\sigma^2}{a^2}$$

$$P(\mu - a\cdot\sigma \leqslant X \leqslant \mu + a\cdot\sigma)$$

Tomando la segunda expresión que hemos visto para la desigualdad de Chebyshef para distintos valores de $a > 0$ tenemos la siguiente tabla.

a	$P(\lvert X - E(X)\rvert \geqslant a\sigma)$
1	$\leqslant 1$
2	$\leqslant 0.25$
3	$\leqslant 0.111$
4	$\leqslant 0.0025$

Interpretación de la desigualdad

- Por ejemplo para $a = 2$ esta desigualdad se puede interpretar como que dada una v.a. X con cualquier distribución que tenga $E(X)$ y $Var(X)$ finitos la probabilidad de que un valor se aleje de la media μ más de $a = 2$ desviaciones típicas es menor o igual que 0.25.
- Es decir sólo el 25 % de los valores estarán alejados de la media más de 2σ
  ¡Sea cual sea la distribución de la v.a.!