

Tema 2: Variables aleatorias

Parte 1: Probabilidad con R y python

septiembre 2024

Lección 1

Introducción a las variables aleatorias

Introducción

- Hasta ahora nuestros sucesos han sido de varios tipos: $\{C, +\}$ en la moneda, nombres de periódicos, ángulos en una ruleta, número de veces que sale cara en el lanzamiento de una moneda etc....
- Necesitamos estandarizar de alguna manera todos estos sucesos. Una solución es asignar a cada suceso un cierto conjunto de números reales, es decir, convertir todos los sucesos en *sucesos de números reales* para trabajar con ellos de forma unificada.
- Para conseguirlo utilizaremos unas funciones que transformen los elementos del espacio muestral en números; estas funciones son las variables aleatorias.

Definición de variable aleatoria

Comenzaremos dando una definición poco rigurosa, pero suficiente, de variable aleatoria.

Variable Aleatoria (definición práctica)

Una variable aleatoria (v.a.) es una aplicación que toma valores numéricos determinados por el resultado de un experimento aleatorio

Notación:

- Normalmente representaremos las v.a. por letras mayúsculas $X, Y, Z \dots$
- Los valores que “*toman*” las v.a. los representaremos por letras minúsculas (las mismas en principio) $x, y, z \dots$

Ejemplo

Ejemplo

Lanzamos un dado convencional de parchís el espacio muestral del experimento es

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

Una v.a $X : \Omega \rightarrow \mathbb{R}$ sobre este espacio queda definida por

$$X(1) = 1, X(2) = 2, X(3) = 3, X(4) = 4, X(5) = 5, X(6) = 6.$$

- Ahora el suceso $A = \{2, 4, 6\}$, es decir “salir número par”, es equivalente a $\{X = 2, X = 4, X = 6\}$.
- El suceso $B = \{1, 2, 3\}$, es decir “salir un número inferior o igual a 3” es en términos de la v.a. $\{X = 1, X = 2, X = 3\}$ o también $\{X \leq 3\}$.

Ejemplo

Ejemplo

Consideremos el experimento lanzar una anilla al cuello de una botella. Si acertamos a ensartar la anilla en la botella el resultado del experimento es **éxito** y **fracaso** en caso contrario.

El espacio muestral asociado a este experimento será $\Omega = \{\text{éxito}, \text{fracaso}\}$. Construyamos la siguiente variable aleatoria:

$$X : \{\text{éxito}, \text{fracaso}\} \rightarrow \mathbb{R}$$

definida por

$$X(\text{éxito}) = 1 \text{ y } X(\text{fracaso}) = 0.$$

Tipos de variables aleatorias

Hay dos tipos fundamentales de variables aleatorias, las discretas y las continuas.

Damos a continuación una definición informal.

Variables Aleatorias Discretas y Continuas

- Una variable aleatoria es **discreta** si sólo puede tomar una cantidad numerable de valores con probabilidad positiva.
- Las variables aleatorias **continuas** toman valores en intervalos.
- También existen las variables aleatorias **mixtas**; con una parte discreta y otra continua.

Ejemplo

Ejemplo

Son variables *aleatorias discretas*:

- Número de artículos defectuosos en un cargamento.
- Número de clientes que llegan a una ventanilla de un banco en una hora.
- Número de errores detectados en las cuentas de una compañía.
- Número de reclamaciones de una póliza de un seguro médico.

Son variables *aleatorias continuas*:

- Renta anual de una familia.
- Cantidad de petróleo importado por un país.
- Variación del precio de las acciones de una compañía de telecomunicaciones.
- Porcentaje de impurezas en un lote de productos químicos.

Lección 2

Variables aleatorias discretas

Distribuciones de probabilidad para v.a. discretas.

- Pasamos ahora a describir el comportamiento de la v.a. Para ello utilizaremos distintas funciones que nos darán algunas probabilidades de la variable aleatoria.
- En el caso discreto estas funciones son la de probabilidad, y la función de distribución o de probabilidad acumulada.
- En el caso discreto la función de probabilidad es la que nos da las probabilidades de los sucesos elementales de la v.a. que definimos a continuación.

Función de probabilidad para variables discretas

Función de Probabilidad

La **función de probabilidad** (*probability mass function* o incluso abusando de notación *probability density function*) de una variable aleatoria discreta X a la que denotaremos por $P_X(x)$ está definida por

$$P_X(x) = P(X = x),$$

es decir la probabilidad de que X tome el valor x .

Si X no asume ese valor x , entonces $P_X(x) = 0$.

Función de probabilidad discreta

Dominio de una variable aleatoria discreta

El conjunto

$$D_X = \{x \in \mathbb{R} \mid P_X(x) > 0\}$$

recibe el nombre de **dominio** de la v.a. y son los valores posibles de esta variable.

En el caso discreto lo más habitual es que $X(\Omega) = D_X$.

Ejemplo

Ejemplo: parchís

Lanzamos un dado de parchís una vez, en esta ocasión representaremos los sucesos elementales por el número de puntos de la cara obtenida, tenemos que

$$\Omega = \{1\text{-puntos}, 2\text{-puntos}, 3\text{-puntos}, 4\text{-puntos}, 5\text{-puntos}, 6\text{-puntos}\}$$

y la variable aleatoria $X : \Omega \rightarrow \mathbb{R}$ viene definida por

$$X(i\text{-puntos}) = i \text{ para } i = 1, 2, 3, 4, 5, 6.$$

Supongamos que el dado está bien balanceado. Entonces

$$P_X(1) = P_X(2) = P_X(3) = P_X(4) = P_X(5) = P_X(6) = \frac{1}{6}; \text{ concretamente.}$$

$$P_X(x) = \begin{cases} \frac{1}{6}, & \text{si } x = 1, 2, 3, 4, 5, 6. \\ 0 & \text{en otro caso.} \end{cases}$$

Ejemplo

Ejemplo: lanzamiento moneda

Sea X la v.a. asociada al lanzamiento de una moneda. Su espacio muestral es $\Omega = \{c, +\}$, la v.a. queda definida por:

$$X(\omega) = \begin{cases} 1 & \text{si } \omega = c \\ 0 & \text{si } \omega = + \end{cases}$$

Su función de probabilidad es:

$$P_X(x) = P(X = x) = \begin{cases} \frac{1}{2}, & \text{si } x = 0, 1, \\ 0, & \text{en otro caso.} \end{cases}$$

Finalmente su dominio es $D_X = \{0, 1\}$.

Ejemplo

Ejemplo urna con bolas

Tenemos una urna con tres bolas rojas, una negra y dos blancas. Realizamos una extracción y observamos el color de la bola entonces un espacio muestral es

$$\Omega = \{roja, blanca, negra\}.$$

Una variable aleatoria asociada al experimento es:

$$X(\omega) = \begin{cases} 1, & \text{si } \omega = roja, \\ 2, & \text{si } \omega = negra, \\ 3, & \text{si } \omega = blanca. \end{cases}$$

Ejemplo

La función de probabilidad es

$$P_X(x) = \begin{cases} \frac{3}{6}, & \text{si } x = 1, \\ \frac{1}{6}, & \text{si } x = 2, \\ \frac{2}{6}, & \text{si } x = 3, \\ 0 & \text{en otro caso.} \end{cases}$$

El dominio de la v.a. X es $D_X = \{1, 2, 3\}$.

Propiedades de la función de probabilidad.

Propiedades básicas de la función de probabilidad

Sea X una v.a. discreta $X : \Omega \rightarrow \mathbb{R}$ con dominio D_X . Su función de probabilidad P_X verifica las siguientes propiedades:

- $0 \leq P_X(x) \leq 1$ para todo $x \in \mathbb{R}$,
- $\sum_{x \in D_X} P_X(x) = 1$.

Ejemplo

Ejemplo: Lanzamiento moneda

Lanzamos al aire tres veces, de forma independiente, una moneda perfecta. El espacio muestral de este experimento es

$$\Omega = \{ccc, cc+, c+c, +cc, c++ , +c+, ++c, +++\}$$

(expresados en orden de aparición).

Este espacio tiene todos los sucesos elementales equiprobables.

Consideremos la variable aleatoria asociada a este experimento:

$$X = \text{número de caras en los tres lanzamientos.}$$

Ejemplo

Su función de probabilidad es:

$$\begin{aligned}P(X = 0) &= P(\{+++ \}) = \frac{1}{8}, \\P(X = 1) &= P(\{c++ , +c+ , ++c \}) = \frac{3}{8}, \\P(X = 2) &= P(\{cc+ , c+c , +cc \}) = \frac{3}{8}, \\P(X = 3) &= P(\{ccc \}) = \frac{1}{8}.\end{aligned}$$

Ejemplo

Podemos reescribir la función de probabilidad de X de forma simplificada:

$$P_X(x) = \begin{cases} \frac{1}{8}, & \text{si } x = 0, 3, \\ \frac{3}{8}, & \text{si } x = 1, 2, \\ 0, & \text{en otro caso.} \end{cases}$$

Efectivamente los valores de la función de distribución suman 1:

$$\sum_{x=0}^3 P_X(x) = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = 1.$$

Función de distribución de variables aleatorias

Distribución de Probabilidad

La función de *distribución de probabilidad* (acumulada) de la v.a. X (de cualquier tipo; discreta o continua) $F_X(x)$ representa la probabilidad de que X tome un menor o igual que x , es decir,

$$F_X(x) = P(X \leq x).$$

Esta función también se denomina función de **distribución de probabilidad o simplemente función de distribución** de una v.a., y en inglés *cumulative distribution function* por lo que se abrevia con el acrónimo cdf.

Propiedades

Propiedades de la Función de Distribución

Sea X una v.a. y F_X su función de distribución:

- ① $P(X > x) = 1 - P(X \leq x) = 1 - F_X(x)$.
- ② Sea a y b tales que $a < b$, $P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F_X(b) - F_X(a)$.

Propiedades

Demostración:

Tenemos que el complementario de X mayor que x es: $\overline{\{X > x\}} = \{X > x\}^c = \{X \leq x\}$.
Además,

$$P(X > x) = 1 - P(\overline{\{X > x\}}) = 1 - P(X \leq x) = 1 - F_X(x),$$

lo que demuestra la primera propiedad.

Por otro lado, si X se encuentra entre dos valores a y b
 $\{a < X \leq b\} = \{X \leq b\} - \{X \leq a\}$. Ahora podemos hacer

$$\begin{aligned} P(a < X \leq b) &= P(\{X \leq b\} - \{X \leq a\}) \\ &= P(\{X \leq b\}) - P(\{X \leq a\}) \\ &= F_X(b) - F_X(a). \end{aligned}$$

Propiedades

Propiedades de la Función de Distribución

Sea F_X la función de distribución de una v.a. X entonces:

- $0 \leq F_X(x) \leq 1$.
- La función F_X es no decreciente.
- La función F_X es continua por la derecha.
- Si denotamos por $F_X(x_0^-) = \lim_{x \rightarrow x_0^-} F(x)$, entonces se cumple que $P(X < x_0) = F_X(x_0^-)$ y que $P(X = x_0) = F_X(x_0) - F_X(x_0^-)$.

Propiedades

Propiedades de la Función de Distribución

- Se cumple que $\lim_{x \rightarrow \infty} F_X(x) = 1$; $\lim_{x \rightarrow -\infty} F_X(x) = 0$.
- Toda función F verificando las propiedades anteriores es función de distribución de alguna v.a. X .

Advertencia desigualdades estrictas

En las propiedades anteriores no se pueden cambiar en general las desigualdades de estrictas o no estrictas.

Veamos que propiedades tenemos cuando se cambian estas desigualdades.

Dada una F_X una función de distribución de la v.a. X y denotamos por

$$F_X(x_0^-) = \lim_{x \rightarrow x_0^-} F_X(x),$$

,

entonces se cumplen las siguientes igualdades...

Advertencia desigualdades estrictas

Propiedades

- $P(X = x) = F_X(x) - F_X(x^-)$.
- $P(a < X < b) = F_X(b^-) - F_X(a)$.
- $P(a \leq X < b) = F_X(b^-) - F_X(a^-)$.
- $P(X < a) = F_X(a^-)$,
- $P(a \leq X \leq b) = F_X(b) - F_X(a^-)$.
- $P(X \geq a) = 1 - F_X(a^-)$.

Propiedades

Más propiedades de la función de distribución

- Si F_X es continua en x se tiene que $P(X = x) = 0$. Así que si la v.a. es continua $P(X \leq a) = P(X < a) + P(X = a) = P(X < a)$ y propiedades similares.
- Sea X una variable aleatoria discreta que con dominio D_X y que tiene por función de probabilidad $P_X(x)$ entonces su función de distribución $F_X(x_0)$ es

$$F_X(x_0) = \sum_{x \leq x_0} P_X(x),$$

donde $\sum_{x \leq x_0}$ indica que sumamos todos los $x \in D_X$ tales que $x \leq x_0$.

Propiedades

Demostración:

Si X es continua,

$$P(X = a) = F(a) - F(a^-) = F(a) - F(a) = 0$$

por lo tanto

$$P(X \leq a) = P(X < a) + P(X = a) = P(X < a) + 0 = P(X < a),$$

lo que demuestra la primera propiedad.

Para demostrar la segunda basta hacer

$$F_X(x_0) = P(X \leq x_0) = P\left(\bigcup_{x \leq x_0; x \in D_X} \{x\}\right) = \sum_{x \leq x_0} P(X = x) = \sum_{x \leq x_0} P_X(x).$$

Ejemplo

Ejemplo: dado (continuación)

En el experimento del dado se tiene que:

$$P_X(x) = \begin{cases} \frac{1}{6}, & \text{si } x = 1, 2, 3, 4, 5, 6 \\ 0, & \text{en el resto de casos.} \end{cases}$$

por lo tanto

$$F_X(x) = P(X \leq x) = \begin{cases} 0, & \text{si } x < 1, \\ \frac{1}{6}, & \text{si } 1 \leq x < 2, \\ \frac{2}{6}, & \text{si } 2 \leq x < 3, \\ \frac{3}{6}, & \text{si } 3 \leq x < 4, \\ \frac{4}{6}, & \text{si } 4 \leq x < 5, \\ \frac{5}{6}, & \text{si } 5 \leq x < 6, \\ 1, & \text{si } 6 \leq x. \end{cases}$$

Ejemplo

Calculemos más detalladamente algún valor de F_X , por ejemplo:

$$\begin{aligned}F_X(3.5) &= P(X \leq 3.5) = P(\{X = 1\} \cup \{X = 2\} \cup \{X = 3\}) \\&= P(\{X = 1\}) + P(\{X = 2\}) + P(\{X = 3\}) \\&= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2},\end{aligned}$$

o de otra forma

$$F_X(3.5) = \sum_{x \leq 3.5} P_X(x) = \sum_{x=1}^3 P(X = x) = \sum_{x=1}^3 \frac{1}{6} = 3 \cdot \frac{1}{6} = \frac{1}{2}.$$

Propiedades de la función de distribución

Propiedad

Sea X una variable con función de distribución F_X entonces:

- $0 \leq F_X(x) \leq 1$ para todo x ,
- Si $x < x'$, entonces

$$F_X(x) \leq F_X(x').$$

Es una función creciente, es decir, no necesariamente estrictamente creciente.

- $\lim_{x \rightarrow -\infty} F_X(x) = 0$ y $\lim_{x \rightarrow +\infty} F_X(x) = 1$.
- Es continua por la derecha $\lim_{x \rightarrow x_0^+} F_X(x) = F_X(x_0)$.

Lección 3

Valores esperados o esperanza

Momentos de variables aleatorias discretas

- Al igual que en la estadística descriptiva se utilizan distintas medidas para resumir los valores centrales y para medir la dispersión de una muestra, podemos definir las correspondiente medidas para variables aleatorias.
- A estas medidas se les suele añadir el adjetivo **poblacionales** mientras que a las que provienen de la muestra se las adjetiva como **muestrales**.

Por ejemplo podemos buscar un valor que resuma toda la variable. Este valor es el que “*esperamos*” que se resuma la v.a. o esperamos que las realizaciones de la v.a. queden cerca de él. Demos su definición formal.

Esperanza de un variable aleatoria discreta

Esperanza de una variable aleatoria discreta

El valor **esperado o esperanza** (*expected value* en inglés) $E(X)$ de una v.a. discreta X , se define como

$$E(X) = \sum_{x \in X(\Omega)} x \cdot P_X(x).$$

En ocasiones se denomina **media** (*mean* en inglés, *mitjana* en catalán) poblacional o simplemente media y muy frecuentemente se la denota $\mu_X = E(X)$ o simplemente $\mu = E(X)$.

Interpretación de la media aritmética para v.a. discretas

Ejemplo: lanzamiento de un dado n veces

Supongamos que lanzamos un dado n veces y obtenemos unas frecuencias absolutas n_i para el resultado i con $i = 1, \dots, 6$. Sea X la v.a. que nos representa el valor de una tirada del dado.

Calculemos la media aritmética (o media muestral) de los datos

$$\bar{x} = \frac{1 \cdot n_1 + 2 \cdot n_2 + 3 \cdot n_3 + 4 \cdot n_4 + 5 \cdot n_5 + 6 \cdot n_6}{n} = \sum_{x=1}^6 x \cdot \frac{n_x}{n}.$$

Si $n \rightarrow \infty$ se tiene que $\lim_{n \rightarrow \infty} \frac{n_x}{n} = P_X(x)$. Por lo tanto $E(X) = \lim_{n \rightarrow \infty} \sum_{x=1}^6 x \cdot \frac{n_x}{n}$.

Entonces el valor esperado en una v.a. discreta puede entenderse como el valor promedio que tomaría una v.a. en un número grande de repeticiones.

Ejemplo

Ejemplo: Erratas en un texto

Sea X = número de erratas en una página de un texto con dominio $D_X = \{0, 1, 2\}$.

Resulta que

$$P(X = 0) = 0.42, \quad P(X = 1) = 0.4, \quad P(X = 2) = 0.18, \quad \text{por lo tanto,}$$
$$E(X) = 0 \cdot 0.42 + 1 \cdot 0.4 + 2 \cdot 0.18 = 0.76.$$

Elegida una página del texto al azar esperamos encontrar 0.76 errores por página.

Supongamos que el editor nos paga 2 euros por cada página que encontremos con 1 error y 3 euros por cada página con dos errores (y nada por las páginas correctas) ¿Cuánto *esperamos* cobrar si analizamos una página?

Esperanzas de funciones de variables aleatorias discretas

Esperanzas de funciones de variables aleatorias discretas

Sea X una v.a. discreta con función de probabilidad P_X y de distribución F_X . Entonces el *valor esperado de una función $g(x)$* es:

$$E(g(X)) = \sum_x g(x) \cdot P_X(x).$$

Propiedades de los valores esperados

Propiedades

- $E(k) = k$ para cualquier constante k .
- Si $a \leq X \leq b$ entonces $a \leq E(X) \leq b$.
- Si X es una v.a. discreta que toma valores enteros no negativos entonces
$$E(X) = \sum_{x=0}^{+\infty} (1 - F_X(x)).$$

Ejercicio

La demostración de las propiedades anteriores se deja como ejercicio.

Ejemplo

Ejemplo: paleta de colores aleatoria

Supongamos que estamos sentados delante de nuestro ordenador con un amigo y le decimos que en dos minutos podemos programar una paleta para poner colores a unos gráficos.

Queremos que la paleta tenga dos botones con las opciones color rojo y color azul. Como hemos programado a gran velocidad resulta que el programa tiene un error; cada vez que se abre la paleta los colores se colocan al azar (con igual probabilidad) en cada botón, así que no sabemos en qué color hemos de pinchar.

Además, como nos sobraron 15 segundos para hacer el programa y pensando en la comodidad del usuario, la paleta se cierra después de haber seleccionado un color y hay que volverla a abrir de nuevo.

La pregunta es ¿cuál es el valor esperado del número de veces que hemos pinchar el botón de color azul antes de obtener este color?

Ejemplo

Llamemos X al número de veces que pinchamos en el botón azul (y nos sale rojo) hasta obtener el primer azul. La variable X toma valores en los enteros no negativos. Su función de probabilidad queda determinada por

$$P_X(x) = P(X = x) = P(\overbrace{\text{rojo}, \text{rojo}, \dots, \text{rojo}}^{x \text{ veces}}, \text{azul}) = \left(\frac{1}{2}\right)^{x+1}.$$

Series geométricas

Series geométricas

- Una **progresión geométrica** de razón r es una sucesión de la forma

$$r^0, r^1, \dots, r^n, \dots$$

- La serie geométrica es la suma de todos los valores de la progresión geométrica $\sum_{k=0}^{+\infty} r^k$.
- Las sumas parciales desde el término n_0 al n de una progresión geométrica valen

$$\sum_{k=n_0}^n r^k = \frac{r^{n_0} - r^{n+1}}{1 - r}.$$

Propiedades de las series geométricas

Propiedades

- Si $|r| < 1$ la serie geométrica es convergente y

$$\sum_{k=0}^{+\infty} r^k = \frac{1}{1-r}$$

- En el caso en que se comience en n_0 se tiene que

$$\sum_{k=n_0}^{+\infty} r^k = \frac{r^{n_0}}{1-r}.$$

Propiedades de las series geométricas

Propiedades

- Si $|r| < 1$ también son convergentes las derivadas, respecto de r , de la serie geométrica y convergen a la derivada correspondiente. Así tenemos que

$$\left(\sum_{k=0}^{+\infty} r^k \right)' = \sum_{k=1}^{+\infty} k r^{k-1}; \quad \left(\frac{1}{1-r} \right)' = \frac{1}{(1-r)^2}$$

$$\left(\sum_{k=0}^{+\infty} r^k \right)'' = \sum_{k=2}^{+\infty} k(k-1) r^{k-2}; \quad \left(\frac{1}{1-r} \right)'' = \frac{2}{(1-r)^3}$$

Ejemplo

Ejemplo: paleta de colores (continuación)

Si seguimos con el ejemplo de la paleta de colores, su esperanza es:

$$E(X) = \sum_{x=0}^{+\infty} x \cdot P(X = x) = \sum_{x=0}^{+\infty} x \cdot \left(\frac{1}{2}\right)^{x+1} = \left(\frac{1}{2}\right)^2 \sum_{x=1}^{+\infty} x \cdot \left(\frac{1}{2}\right)^{x-1} = \left(\frac{1}{2}\right)^2 \cdot \frac{1}{\left(1 - \frac{1}{2}\right)^2} = 1.$$

Ahora calculemos su función de distribución

$$F_X(x) = P(X \leq x) = \sum_{k=0}^x P(X = k) = \sum_{k=0}^x \left(\frac{1}{2}\right)^{k+1} = \frac{\frac{1}{2} - \frac{1}{2}^{x+1}}{1 - \frac{1}{2}} \cdot \frac{1}{2} = 1 - \left(\frac{1}{2}\right)^{x+1}.$$

Ejemplo

Como la variable toma valores enteros positivos, podemos calcular su valor esperado de esta otra manera

$$E(X) = \sum_{x=0}^{+\infty} (1 - F_X(x)) = \sum_{x=0}^{+\infty} \left(\frac{1}{2}\right)^{x+1} = \frac{1}{2} \cdot \frac{1}{1 - \frac{1}{2}} = 1.$$

Ejercicio

Calculad el valor esperado de la variable

Y = número de intentos para conseguir el color azul.

Momentos de una variable aleatoria

Momentos de orden m

Llamaremos **momento de orden** m respecto al punto C a

$$E((X - C)^m).$$

- Cuando $C = 0$ los momentos reciben el nombre de **momentos respecto al origen**.
- Cuando $C = E(X)$ reciben el nombre de **momentos centrales o respecto de la media**. Luego la esperanza es el momento de orden 1 respecto al origen. Estos momentos son la versión poblacional de los momentos que vimos en el curso de estadística descriptiva, recibiendo estos último el nombre de momentos muestrales.

Resumen de conceptos

- Hemos descrito el comportamiento aleatorio de una v.a. discreta mediante sus funciones de probabilidad P_X y de distribución F_X .
- También tenemos un valor central; el valor esperado $E(X)$.
- Como medida básica nos queda definir una medida de lo lejos que están los datos del valor central $E(X)$ una de estas medidas es la varianza de X .

Lección 4

Medidas de la variabilidad

Medidas de la variabilidad

Varianza

Sea X una v.a. Llamaremos **varianza** de X a

$$Var(X) = E((X - E(X))^2).$$

Por lo tanto, la varianza es el momento central de orden 2.

De forma frecuente se utiliza la notación

$$\sigma_X^2 = Var(X).$$

A la raíz cuadrada positiva de la varianza

$$\sigma_X = +\sqrt{Var(X)}.$$

se la denomina desviación típica o estándar de X .

Propiedades de la varianza

Propiedad

- Si X es una v.a. discreta con función de probabilidad P_X su varianza es

$$\sigma_X^2 = Var(X) = E((X - E(X))^2) = \sum_x (x - E(X))^2 \cdot P_X(x).$$

- Sea X una v.a.

$$Var(X) = E(X^2) - (E(X))^2 = \sum_x x^2 \cdot P_X(X) - (E(X))^2$$

Demostración

Demostración de b)

$$\begin{aligned} \text{Var}(X) &= \sum_x (x - E(X))^2 P_X(x) = \sum_x (x^2 - 2 \cdot x \cdot E(X) + (E(X))^2) \cdot P_X(x) \\ &= \sum_x x^2 \cdot P_X(x) - E(X) \sum_x 2 \cdot x \cdot P_X(x) + (E(X))^2 \cdot \sum_x P_X(x) \\ &= E(X^2) - 2E(X) \cdot E(X) + (E(X))^2 = E(X^2) - (E(X))^2. \end{aligned}$$

Ejemplo

Ejemplo: número de errores (continuación)

Calculemos en el ejemplo anterior la varianza del número de errores.

Recordemos que:

$$P(X = 0) = 0.42, \quad P(X = 1) = 0.4, \quad P(X = 2) = 0.18,$$

y que

$$E(X) = 0.76.$$

Entonces:

$$Var(X) = E(X^2) - (E(X))^2 = E(X^2) - (0.76)^2.$$

Ejemplo

Ahora necesitamos calcular

$$E(X^2) = 0^2(0.41) + 1^2(0.4) + 2^2(0.18) = 0.4 + 0.72 = 1.12$$

y por lo tanto

$$Var(X) = E(X^2) - (0.76)^2 = 1.12 - 0.5776 = 0.542$$

y

$$\sqrt{Var(X)} = \sqrt{0.542}$$

En resumen $\sigma_X^2 = 0.542$ y $\sigma_X = \sqrt{0.542}$

Propiedades de la varianza

Propiedades de la varianza

- $Var(X) \geq 0$.
- $Var(cte) = E(cte^2) - (E(cte))^2 = cte^2 - cte^2 = 0$.
- El mínimo de $E((X - C)^2)$ se alcanza cuando $C = E(X)$ y es $Var(X)$. Esta propiedad es una de las que hace útil a la varianza como medida de dispersión.

Ejercicio

Se deja como ejercicio la demostración de estas propiedades.

Lección 5

Esperanza y varianza de transformaciones lineales.

Transformaciones lineales.

Transformación lineal

Un **cambio de variable lineal** o **transformación lineal** de una v.a. X es otra v.a. $Y = a + b \cdot X$ donde $a, b \in \mathbb{R}$.

Esperanza de una transformación lineal

Sea X una v.a. con $E(X) = \mu_X$ y $Var(X) = \sigma_X^2$ y $a, b \in \mathbb{R}$. Entonces si $Y = a + b \cdot X$:

- $E(Y) = E(a + bX) = a + bE(X) = a + b \cdot \mu_X$.
- $Var(Y) = Var(a + bX) = b^2 Var(X) = b^2 \cdot \sigma_X^2$ -
- $\sigma_Y = \sqrt{Var(Y)} = \sqrt{b^2 Var(X)} = |b| \cdot \sigma_X$ -

Demostración

Demostración:

$$\begin{aligned} E(Y) &= E(a + bX) = \sum_x (a + b \cdot x) \cdot P_X(x) \\ &= a \sum_x P_X(x) + b \sum_x x \cdot P_X(x) \\ &= a + b \cdot E(X) = a + b \cdot \mu_X. \end{aligned}$$

Ejercicio

Las demostración de las demás propiedades se dejan como ejercicio.

Lección 6

Variables aleatorias continuas

Variables aleatorias continuas definición.

- 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*.

Variables aleatorias continuas

- En general tendremos que $P(X < x_0) = P(X \leq x_0)$.
- 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 $[0, 1]$.

Ejemplo: distancia dardo centro de la diana

Supongamos que lanzamos un dardo a una diana de radio 1, de forma que sea *equiprobable* cualquier distancia al centro (¡Cuidado! esto no es equivalente que cualquier punto de la diana sea *equiprobable*).

Consideremos la v.a. continua $X =$ distancia al centro de la diana.

Su función de distribución es

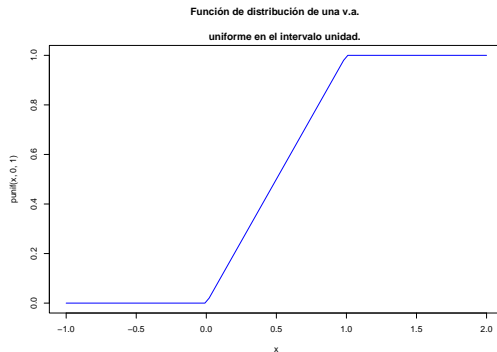
$$F_X(x) = \begin{cases} 0, & \text{si } x \leq 0, \\ x, & \text{si } 0 < x < 1, \\ 1, & \text{si } x \geq 1. \end{cases}$$

consideremos

- C.F. *longitud favorable* que es $x - 0$,
- C.P. *longitud posible* que es $1 - 0$

Gráfica de la función de distribución uniforme

```
curve(punif(x,0,1),xlim=c(-1,2),col="blue",  
      main="Función de distribución de una v.a. \n  
      uniforme en el intervalo unidad.")
```



Propiedades

En las variables continuas los sucesos del tipo $\{X \leq x\}$ 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:

- $P(X \leq b) = P(X < b).$
- $P(X < b) = P(X < a) + P(a < X < b).$
- $P(a < X < b) = P(X < b) - P(X < a).$

Demostración

Demostración:

La primera es evidente $P(X \leq b) = P(X < b) + P(X = b) = P(X < b)$.

Para demostrar la segunda, tenemos

$$\{X \leq a\} \cap \{a < X < b\} = \emptyset$$

$$\{X \leq a\} \cup \{a < X < b\} = \{X < b\},$$

entonces

$$P(X < b) = P(\{X \leq a\} \cup \{a < X < b\}) = P(X \leq a) + P(a < X < b) = P(X < a) + P(a < X < b).$$

La demostración de la tercera propiedad es similar a la segunda pero aplicando la primera. La dejamos como ejercicio.

Propiedades de la función de distribución

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

Propiedades de la Función de Distribución

Dada una variable aleatoria continua se tiene que:

- $F_X(b) = F_X(a) + P(a < X < b)$.
- $P(a < X < b) = F_X(b) - F_X(a)$.
- $P(a \leq X \leq b) = F_X(b) - F_X(a)$.

Ejercicio

Se deja la demostración como ejercicio

Ejemplo

Ejemplo: diana (continuación)

En el ejemplo de la diana:

$$P(0.25 < X < 0.3) = F_X(0.3) - F_X(0.25) = 0.3 - 0.25 = 0.05.$$

Función de densidad

Función de densidad

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

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

Función de distribución de una variable aleatoria.

Función de distribución de una variable aleatoria

Sea X una v.a. con función de distribución F_X . Sea $f : \mathbb{R} \rightarrow \mathbb{R}$ 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 .

Dominio de una variable aleatoria continua

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 como su conjunto de resultados posibles.

Ejemplo: diana (continuación)

En nuestra ejemplo de la diana, la función f es una densidad

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

que es la densidad de X , en efecto:

Densidad diana

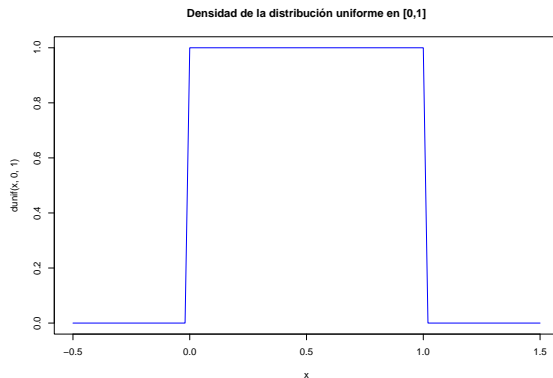
$$f_X(x) = \begin{cases} 0, & \text{si } x \leq 0, \\ 1, & \text{si } 0 < x < 1, \\ 0, & \text{si } 1 \leq x. \end{cases}$$

- Si $x \leq 0$ entonces $\int_{-\infty}^x f_X(t)dt = 0$.
- Si $0 \leq x \leq 1$ entonces $\int_{-\infty}^x f_X(t)dt = \int_0^x 1dt = x$.
- Si $x \geq 1$ 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 $x \in \mathbb{R}$.

Densidad diana

```
curve(dunif(x,0,1),xlim=c(-0.5,1.5),col="blue",  
      main="Densidad de la distribución uniforme en [0,1]")
```



Utilidad de la función de densidad

La función de densidad nos permite calcular diversas probabilidades.

Propiedades de la función de densidad

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

$$\begin{aligned}P(a < X < b) &= P(a < X \leq b) = P(a \leq X < b) = \\P(a \leq X \leq b) &= \int_a^b f_X(x)dx.\end{aligned}$$

- Si A es un subconjunto adecuado de \mathbb{R} entonces

$$P(X \in A) = \int_A f(x)dx = \int_{A \cap D_X} f(x)dx.$$

Utilidad de la función de densidad

Propiedades de la función de densidad

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

- Si f_x es continua en un punto x , F_X es derivable en ese punto y $F'_X(x) = f_X(x)$.
- $P(X = x) = 0$ para todo $x \in \mathbb{R}$.

Ejercicio

Comprobar estas propiedades en el ejemplo de la diana.

Ejemplo tiempo ejecución de un proceso

Ejemplo: tiempo ejecución de un proceso.

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.

Calculemos la función de densidad y de distribución de la v.a X .

Entonces

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

Luego su función de distribución es:

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

Ejemplo tiempo ejecución de un proceso

Su función de densidad por su lado es:

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

Efectivamente

- $f_X(x) \geq 0$, y tiene un conjunto finito de discontinuidades: en 0 y en 2
- $F_X(x) = \int_{-\infty}^x f_X(t)dt$, para todo $x \in \mathbb{R}$ (Ejercicio: resolverlo gráficamente.)
- $\int_{-\infty}^{+\infty} f_X(x)dx = \int_0^2 \frac{1}{2}dx = \left[\frac{x}{2}\right]_0^2 = \frac{2}{2} - \frac{0}{2} = 1.$

Ejemplo tiempo ejecución de un proceso

Ejercicio: Tiempo de un proceso:

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.

Lección 7

Esperanza y varianza para variables aleatorias continuas

Esperanza y varianza para variables aleatorias continuas

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.

En lo que sigue, salvo que indiquemos lo contrario, X es una v.a. continua con función de densidad $f_X(x)$

Esperanza y varianza para variables aleatorias continuas

Esperanza v.a. continuas

- Su esperanza es:

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

- Si $g(x)$ es una función de la variable X entonces:

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

Esperanza y varianza para variables aleatorias continuas

Varianza v.a. continuas

- Su varianza es:

$$Var(X) = \sigma_X^2 = E((X - \mu_X)^2) = \int_{-\infty}^{+\infty} (x - \mu_X)^2 \cdot f_X(x) dx.$$

- Su desviación típica es:

$$\sigma_X = \sqrt{\sigma_X^2}.$$

Propiedades

Propiedades

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

Ejemplo

Ejemplo: diana (continuación)

Calcular μ_X y σ_X^2 en el ejemplo de la diana

Resultado

$$\mu_X = \frac{1}{2},$$

$$E(X^2) = \frac{1}{3},$$

$$Var(X) = \frac{1}{12}.$$

Esperanza de trans. lineales de v.a. continuas

Proposición

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:

- $E(Y) = E(a + b \cdot X) = a + b \cdot E(X)$.
- $Var(Y) = Var(a + b \cdot X) = b^2 \cdot Var(X)$.
- $\sigma_Y = |b| \cdot \sigma_X$.
- $Z = \frac{X - \mu_X}{\sigma_X}$ es una transformación lineal de X de forma que

$$E(Z) = 0 \text{ y } Var(Z) = 1$$

Ejemplo

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 50.000 € y el beneficio por litro es de 10 € por botella. Definimos $T = 10 \cdot X - 50000$, que será el beneficio después de gastos.

Entonces la esperanza del beneficio es

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

y

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

Lección 8

Transformaciones de variables aleatorias

Transformaciones de variables aleatorias

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 .

Propiedad

Transformaciones de v.a. discretas

Sea X una v.a. discreta con $X(\Omega) = \{x_1, x_2, \dots, x_n, \dots\}$ y sea $h : \mathbb{R} \rightarrow \mathbb{R}$ 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

- $P_Y(y) = \sum_{x_i | h(x_i)=y} P_X(x_i).$
- $F_Y(y) = \sum_{x_i | h(x_i) \leq y} P_X(x_i).$

Propiedades

Desafortunadamente para variables no discretas el resultado no es tan sencillo como el anterior, pues la transformación de, por ejemplo, una v.a. continua puede ser continua, discreta, mixta,...

Transformación de v.a. continuas en continuas

Sea X una v.a. continua cuya función de densidad es f_X . Sea $h : \mathbb{R} \rightarrow \mathbb{R}$, una aplicación estrictamente monótona y derivable, por lo tanto $h'(x) \neq 0$ 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

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

Propiedades

Densidad de una transformación de una v.a. continua

Sea X una v.a. continua cuya función de densidad es f_X . Sea

$$h : \mathbb{R} \rightarrow \mathbb{R}$$

una aplicación, no necesariamente monótona tal que :

- sea derivable con derivada no nula
- la ecuación $h(x) = y$ tiene un número finito de soluciones x_1, x_2, \dots, x_n

entonces:

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

Método general de transformación de v.a.

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

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y).$$

Método general del transformación de variables aleatorias

Por ejemplo, si g es estrictamente creciente y continua,

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

y si g es estrictamente decreciente y continua,

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

Lección 9

Desigualdades básicas: Markov y Chebychev

Desigualdades de Markov y de Chebychev

- 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 medida de la dispersión de los datos.

Desigualdad de Markov

Desigualdad de Markov

Sea X una v.a. positiva con $E(X)$ finita. Entonces

$$P(X \geq a) \leq \frac{E(X)}{a} \text{ para todo } a > 0.$$

Demostración:

Si X es continua y solo toma valores positivos

$$\begin{aligned} E(X) &= \int_{-\infty}^{+\infty} x \cdot f_X(x) dx = \int_0^{+\infty} x \cdot f_X(x) dx = \int_0^a x \cdot f_X(x) dx + \int_a^{+\infty} x \cdot f_X(x) dx \\ &\geq \int_a^{+\infty} x \cdot f_X(x) dx \geq a \int_a^{+\infty} f_X(x) dx = a \cdot P(X \geq a), \end{aligned}$$

Desigualdad de Markov

Corolario

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

$$P(|X| \geq a) \leq \frac{E(|X|)}{a}.$$

Ejercicio

Demuestra el corolario anterior a partir de la desigualdad de Markov.

Desigualdad de Chebychev

La **desigualdad de Chebychev** también se denomina de Chebyshev y en inglés *Chebyshev*.

Desigualdad de Chebychev

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

$$P(|X - \mu| \geq a) \leq \frac{\sigma^2}{a^2}.$$

Demostración desigualdad de Chebychev

Demostración

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

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

entonces

$$P(Y^2 \geq a^2) \leq \frac{E(Y^2)}{a^2} = \frac{E((X - \mu)^2)}{a^2} = \frac{Var(X)}{a^2} = \frac{\sigma^2}{a^2}.$$

Por otra parte

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

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

Uso de la desigualdad de Chebychev

Utilidad básica de la desigualdad de Chebychev

Supongamos que X es una v.a. con $Var(X) = 0$, entonces, aplicando la desigualdad anterior

$$P(|X - E(X)| \geq a) = 0 \text{ para todo } a > 0,$$

lo que implica que

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

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

Ejemplo

Ejemplo: tiempo de respuesta

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

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

Si sustituimos a por $a \cdot \sigma$ en la desigualdad de Chebychev, nos queda:

$$P(|X - \mu| \geq a \cdot \sigma) \leq \frac{\sigma^2}{(a \cdot \sigma)^2} = \frac{1}{a^2},$$

que es otra manera de expresar la desigualdad de Chebychev.

Más formas de la desigualdad de Chebychev

La desigualdad de Chebychev también se puede escribir de al menos dos maneras más:

$$P(\mu - a \leq X \leq \mu + a) \geq 1 - \frac{\sigma^2}{a^2},$$

y tomado como $a = k \cdot \sigma$,

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

La varianza como medida de dispersión

Tomando la segunda expresión que hemos visto para la desigualdad de Chebychev para distintos valores de $k > 0$ obtenemos la siguiente tabla:

k	$P(X - E(X) \geq k \cdot \sigma)$
1	≤ 1
2	≤ 0.25
3	≤ 0.111
4	≤ 0.0025

Interpretación de la desigualdad

- Por ejemplo para $k = 2$, 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 $2 \cdot \sigma$

¡Sea cual sea la distribución de la v.a.!