Estadística I

Profesora: Ana Arribas Gil

Despacho: 7.0.23

Teléfono: 91 624 58 53

e-mail: `aarribas@est-econ.uc3m.es`

Página web (material): `http://www.est.uc3m.es/aarribas` → Docencia

Tutorías: a concretar en cada caso (por e-mail o en clase)

Profesor de prácticas: André Santos (`aalves@est-econ.uc3m.es`)

Página web de la asignatura:

`http://www.est.uc3m.es/esp/nueva_docencia/comp_col_get/lade/estadistica_I/`

Programa de la asignatura

Tema 0.	Introducción	24 sept.
Tema 1.	Distribuciones en el muestreo	24 sept. - 1 oct.
Tema 2.	Estimación puntual	6 oct. - 20 oct.
Tema 3.	Estimadores de máxima verosimilitud	22 oct. - 5 nov.
Tema 4.	Intervalos de confianza	10 nov. - 17 nov.
Tema 5.	Contraste de hipótesis	19 nov. - 1 dic.
Tema 6.	Diagnosis del modelo	1 dic. - 17 dic.

Criterios de evaluación

- Examen final (finales de enero) común a todos los grupos de las distintas titulaciones.

- Hasta un máximo de 0.5 puntos a sumar a la nota del examen:
    - Entrega de los ejercicios propuestos.
    - Asistencia a las prácticas de informática (Statgraphics).

La asistencia a las prácticas de ordenador es obligatoria

Calendario provisional de clases de ejercicios y prácticas

Miércoles	1	Octubre	Práctica 1
Miércoles	15	Octubre	Hoja 1 de ejercicios
Miércoles	29	Octubre	Hoja 2 de ejercicios
Lunes	3	Noviembre	Práctica 2
Miércoles	26	Noviembre	Hoja 3 de ejercicios
Miércoles	3	Diciembre	Hoja 4 de ejercicios
Lunes	15	Diciembre	Prácticas 3 y 4
Miércoles	17	Diciembre	Hoja 4 (bis) de ejercicios

Tema 0: Introducción y repaso

- Objetivos de la asignatura
- Repaso:
  - Variables aleatorias continuas y discretas
  - Función de densidad, probabilidad y distribución
  - Esperanza y varianza de una variable aleatoria
  - Esperanza de sumas de variables y productos de variables independientes
  - Varianza de sumas y diferencias de variables independientes

Objetivos básicos de la asignatura

- Entender los conceptos básicos de parámetro, estimador, distribución muestral, etcétera.
- Estimar parámetros de los modelos probabilísticos discretos y continuos estudiados en *Introducción a la Estadística*.
- Contrastar hipótesis respecto a los parámetros del modelo asumido.
- Evaluar el ajuste del modelo considerado a la realidad experimental u observacional en estudio.

Variables aleatorias

Definición 1.

Una variable aleatoria (v. a.) es una función que asocia un valor numérico a todos los posibles resultados de un experimento aleatorio.

Ejemplo 1.

Consideremos el experimento aleatorio que consiste en lanzar dos veces un dado equilibrado.
El espacio muestral es:

$$\Omega = \{(1,1), \ldots, (1,6), \ldots, (6,1), \ldots, (6,6)\}.$$

$X =$ *"la suma de las dos tiradas" es una variable aleatoria.*

Definición 2.

Una variable aleatoria es discreta si toma un número finito o numerable de valores.
En el Ejemplo 1, X es una v. a. discreta.

Variables aleatorias

Definición 3.

Una variable aleatoria es continua si toma un número infinito no numerable de valores (por ejemplo, en un intervalo de $\mathbb{R}$).

Ejemplo 2.

Consideremos el experimento aleatorio que consiste en medir y pesar a una persona adulta elegida al azar.
El espacio muestral es:

$$\Omega = [1.00, 2.30] \times [20, 200] \quad (m \times kg).$$

$X =$ *"el índice de masa corporal"* $(I.M.C. = \frac{peso\ (kg)}{altura^2 (m)})$ *es una variable aleatoria continua.*

Distribución de una variable aleatoria

La distribución de una v. a. viene dada por los valores que puede tomar y las probabilidades de que la variable tome cada uno de esos valores (o tome valores dentro de un intervalo, en el caso continuo).

La distribución de una variable aleatoria queda caracterizada por:

X v. a. discreta
que toma valores $x_1, x_2, \ldots$

- Función de probabilidad (o de masa)
  $P(X = x_i) \quad \forall x_i, i = 1, 2, \ldots$
- Función de distribución
  $F(x) = P(X \leq x) \quad \forall x \in \mathbb{R}.$

X v. a. continua
que toma valores en $\mathbb{R}$

- Función de densidad
  $f(x) = F'(x) \quad \forall x \in \mathbb{R}.$
- Función de distribución
  $F(x) = P(X \leq x) \quad \forall x \in \mathbb{R}.$

Notación: X denota la v. a. y x los valores que toma.

Distribución de una variable aleatoria discreta

X v. a. discreta que toma los valores $x_1, x_2, \ldots$

Función de probabilidad (o de masa): $P(X = x_i) \quad \forall x_i, i = 1, 2, \ldots$

Propiedades:

$$0 \leq P(X = x_i) \leq 1 \quad \forall x_i, i = 1, 2, \ldots$$

$$\sum_i P(X = x_i) = 1 \quad \text{(suma sobre todos los posibles valores de } X\text{)}$$

Ejemplo 1 (cont.)

La probabilidad de que la suma sea menor o igual que 4:
$P(X \leq 4) = P(X = 2) + P(X = 3) + P(X = 4) = 7/36.$

La probabilidad de que la suma sea mayor que 3:
$P(X > 3) = 1 - P(X \leq 3) = 1 - (P(X = 2) + P(X = 3)) = 33/36.$

Distribución de una variable aleatoria discreta

X v. a. discreta que toma los valores $x_1, x_2, \ldots$

Función de distribución: $F(x) = P(X \leq x) \quad \forall x \in \mathbb{R}$ ¡definida en todo $\mathbb{R}$!

Propiedades:

$$F(-\infty) = 0$$

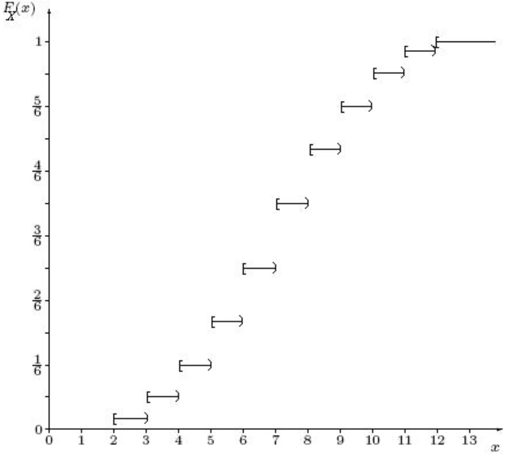
$$F(\infty) = 1$$

$$F(x) \leq F(x+\varepsilon) \quad \forall x \in \mathbb{R}, \forall \varepsilon > 0$$ No decreciente

$$\lim_{\varepsilon \to 0} F(x+\varepsilon) = F(x) \quad \forall x \in \mathbb{R}, \forall \varepsilon > 0$$ Continua por la derecha

Distribución de una variable aleatoria discreta

Ejemplo 1 (cont.) La f. de distribución de la v. a. X



Distribución de una variable aleatoria continua

X v. a. continua que toma valores en $\mathbb{R}$.

RECORDAD: $P(X = x) = 0 \quad \forall x \in \mathbb{R}$

Función de densidad: $f(x) \quad \forall x \in \mathbb{R}$

Propiedades:

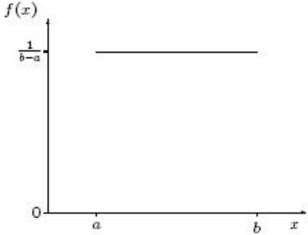
$$f(x) \geq 0 \quad \forall x \in \mathbb{R}$$

$$P(a \leq X \leq b) = \int_a^b f(x)dx \quad \forall a, b \in \mathbb{R}$$

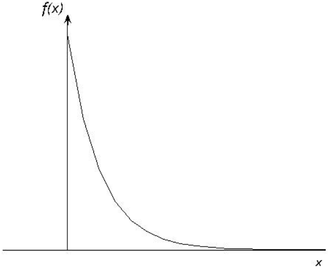
$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)du; \quad f(x) = F'(x) \forall x \in \mathbb{R}$$

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

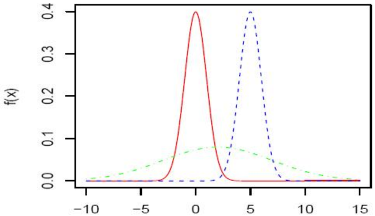
Distribución de una variable aleatoria continua



F. densidad uniforme



F. densidad exponencial



F. densidad normal con varios parámetros

Distribución de una variable aleatoria continua

X v. a. continua que toma valores en $\mathbb{R}$.

Función de distribución: $F(x) = P(X \leq x) \quad \forall x \in \mathbb{R}$.
(Misma definición y propiedades que en el caso discreto)

Propiedades:

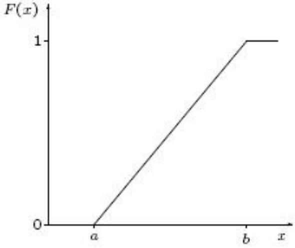
$$F(-\infty) = 0$$

$$F(\infty) = 1$$

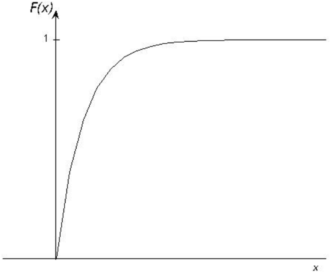
$$F(x) \leq F(x + \varepsilon) \quad \forall x \in \mathbb{R}, \forall \varepsilon > 0$$ No decreciente

$$\lim_{\varepsilon \to 0} F(x + \varepsilon) = F(x) \quad \forall x \in \mathbb{R}, \forall \varepsilon > 0$$ Continua por la derecha

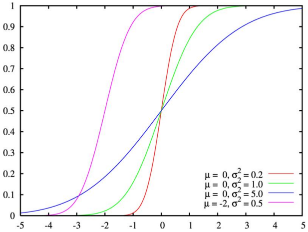
Distribución de una variable aleatoria continua



F. distribución uniforme



F. distribución exponencial



F. distribución normal con varios parámetros

Esperanza y varianza de una variable aleatoria

X v. a. discreta que toma valores $x_1, x_2, \ldots$:

$$\begin{aligned} E[X] &= \sum_i x_i \cdot P(X = x_i) \\ Var[X] &= E[(X - E[X])^2] = \sum_i (x_i - E[X])^2 \cdot P(X = x_i) \\ &= \sum_i x_i^2 \cdot P(X = x_i) - E[X]^2 \end{aligned}$$

X v. a. continua que toma valores en $\mathbb{R}$:

$$\begin{aligned} E[X] &= \int_{\mathbb{R}} x \cdot f(x) dx \\ Var[X] &= E[(X - E[X])^2] = \int_{\mathbb{R}} (x - E[X])^2 \cdot f(x) dx \\ &= \int_{\mathbb{R}} x^2 \cdot f(x) dx - E[X]^2 \end{aligned}$$

EN LOS DOS CASOS: $Var[X] = E[X^2] - E[X]^2$

Variables aleatorias independientes

Recordatorio de Probabilidad:

Ejemplo 1 (cont.) $\Omega = \{(1,1),\ldots,(1,6),\ldots,(6,1),\ldots,(6,6)\}$ $|\Omega| = 36$.

$A =$"la suma de las dos tiradas es 3" es un suceso:

$$A = \{(1,2),(2,1)\} \quad |A| = 2.$$

$B =$"la diferencia de las dos tiradas ($1^a - 2^a$) es 1" es un suceso

$$B = \{(2,1),(3,2),(4,3),(5,4),(6,5)\} \quad |B| = 5.$$

Probabilidad conjunta (que ocurran A y B): $P(A\cap B) = \frac{1}{36}$

Probabilidad condicionada (que ocurra A si ha ocurrido B):

$$P(A|B) = \frac{P(A\cap B)}{P(B)} = \frac{1/36}{5/36} = \frac{1}{5}; \qquad P(B|A) = \frac{P(A\cap B)}{P(A)} = \frac{1/36}{2/36} = \frac{1}{2}$$

Independencia: A y B son independientes si y sólo si $P(A\cap B) = P(A)\cdot P(B)$

En este caso A y B no son independientes:
$P(A)\cdot P(B) = \frac{10}{36^2} \neq \frac{1}{36} = P(A\cap B)$

Variables aleatorias independientes

Ejemplo 1 (cont.) $\Omega = \{(1,1), \ldots, (1,6), \ldots, (6,1), \ldots, (6,6)\} \quad |\Omega| = 36.$

$X =$"la suma de las dos tiradas" es una v. a. discreta
$Y =$"la diferencia de las dos tiradas ($1^a - 2^a$)" es una v. a. discreta

$$P(A \cap B) = P(X = 3, Y = 1) = \frac{1}{36}$$

Definición 4.

Para dos v. a. discretas X e Y se define su función de probabilidad (o de masa) conjunta como:

$$P(X = x, Y = y) \quad \forall x \text{ valor de } X, \forall y \text{ valor de } Y$$

Y se definen las funciones de probabilidad condicionada como:

$\forall y$ *valor de* Y *t.q.* $P(Y = y) > 0$, $\quad P(X = x|Y = y) = \dfrac{P(X = x, Y = y)}{P(Y = y)} \quad \forall x$ *valor de* X

$\forall x$ *valor de* X *t.q.* $P(X = x) > 0$, $\quad P(Y = y|X = x) = \dfrac{P(X = x, Y = y)}{P(X = x)} \quad \forall y$ *valor de* Y

Variables aleatorias independientes

Ejemplo 1 (cont.) Función de probabilidad conjunta de X e Y:

$y \setminus x$	2	3	4	5	6	7	8	9	10	11	12
-5	0	0	0	0	0	$\frac{1}{36}$	0	0	0	0	0
-4	0	0	0	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	0	0	0
-3	0	0	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	0	0
-2	0	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	0
-1	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0
0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$
1	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0
2	0	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	0
3	0	0	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	0	0
4	0	0	0	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	0	0	0
5	0	0	0	0	0	$\frac{1}{36}$	0	0	0	0	0

¡Suma 1!

Función de probabilidad de X condicionada a que Y vale 3:

x	5	7	9
$P(X = x \mid Y = 3)$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

¡Suma 1!

Variables aleatorias independientes

Definición 5.

Dos v. a. discretas X e Y son independientes si y sólo si:

$$P(X = x, Y = y) = P(X = x) \cdot P(Y = y) \quad \forall x \text{ valor de } X, \forall y \text{ valor de } Y$$

o equivalentemente, si y sólo si:

$$P(X = x | Y = y) = P(X = x) \quad \forall x \text{ valor de } X, \forall y \text{ valor de } Y$$

o equivalentemente, si y sólo si:

$$P(Y = y | X = x) = P(Y = y) \quad \forall x \text{ valor de } X, \forall y \text{ valor de } Y$$

Ejemplo 1 (cont.) X e Y no son independientes, ya que, por ejemplo,

$$P(X = 3) \cdot P(Y = 1) = \frac{10}{36^2} \neq \frac{1}{36} = P(X = 3, Y = 1)$$

O también (¡pero con que no se cumpla una vez ya basta!)

$$P(X = 5 | Y = 3) = \frac{1}{3} \neq \frac{1}{9} = P(X = 5)$$

Variables aleatorias independientes

En el caso de las v. a. continuas todo es análogo:

Definición 6.

Para dos v. a. continuas X e Y se define su función de densidad conjunta como una función $f : \mathbb{R}^2 \to [0, \infty)$, verificando $\int_{\mathbb{R}^2} f(x, y)dxdy = 1$ y tal que:

$$P(a \leq X \leq b, c \leq Y \leq d) = \int_c^d \int_a^b f(x, y)dxdy \quad \forall a, b, c, d \in \mathbb{R}$$

Y se definen las funciones de densidad condicionada como:

$$\forall y \in \mathbb{R} \ t.q. \ f(y) > 0, \quad f(x|y) = f(x, y)/f(y) \quad \forall x \in \mathbb{R}$$

$$\forall x \in \mathbb{R} \ t.q. \ f(x) > 0, \quad f(y|x) = f(x, y)/f(x) \quad \forall y \in \mathbb{R}$$

Definición 7.

Dos v. a. continuas X e Y son independientes si y sólo si:

$$f(x, y) = f(x) \cdot f(y) \quad \forall x, y \in \mathbb{R}$$

Variables aleatorias independientes

Caso n-dimensional:

Definición 8.

Las v. a. discretas $X_1, X_2, \ldots, X_n$ *son independientes entre sí, si y sólo si:*

$$P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) = P(X_1 = x_1) \cdot P(X_2 = x_2) \cdot \cdots \cdot P(X_n = x_n)$$

$$\forall x_1, x_2, \ldots, x_n \text{ valores de } X_1, X_2, \ldots, X_n$$

Definición 9.

Las v. a. continuas $X_1, X_2, \ldots, X_n$ *son independientes entre sí , si y sólo si:*

$$f(x_1, x_2, \ldots, x_n) = f(x_1) \cdot f(x_2) \cdot \cdots \cdot f(x_n) \quad \forall x_1, x_2, \ldots, x_n, \in \mathbb{R}$$

Esperanza de sumas de variables y productos de variables independientes

Sean X e Y dos v. a. cualesquiera (discretas o continuas):

$$\begin{aligned} E[X+Y] &= E[X]+E[Y] \\ E[X-Y] &= E[X]-E[Y] \end{aligned}$$

Si además X e Y son independientes, entonces:

$$E[X \cdot Y] = E[X] \cdot E[Y]$$

Caso n-dimensional:

Sean $X_1, X_2, \ldots, X_n$ v. a. cualesquiera (discretas o continuas):

$$E[X_1 + X_2 + \cdots + X_n] = E[X_1] + E[X_2] + \cdots + E[X_n]$$

Si además $X_1, X_2, \ldots, X_n$ son independientes entre sí, entonces:

$$E[X_1 \cdot X_2 \cdot \ldots \cdot X_n] = E[X_1] \cdot E[X_2] \cdot \ldots \cdot E[X_n]$$

Varianzas de sumas y diferencias de variables independientes

Sean X e Y dos v. a. (discretas o continuas) independientes:

$$\begin{aligned} Var[X+Y] &= Var[X] + Var[Y] \\ Var[X-Y] &= Var[X] + Var[Y] \end{aligned}$$

Caso n-dimensional:

Sean $X_1, X_2, \ldots, X_n$ v. a. (discretas o continuas) independientes entre sí:

$$Var[X_1 + X_2 + \cdots + X_n] = Var[X_1] + Var[X_2] + \cdots + Var[X_n]$$