

MÉTODOS ESTADÍSTICOS PARA LA EMPRESA

TEMA 5: ANÁLISIS CONJUNTO DE VARIABLES ALEATORIAS Y DISTRIBUCIÓN DE AGREGADOS

- 5.1.- Distribuciones n-dimensionales. Análisis marginal y condicionado**
- 5.2.- Variables aleatorias independientes. Propiedades**
- 5.3.- Agregación de variables aleatorias**
- 5.4.- Teorema Central del Límite y sus aplicaciones**



Universidad de Oviedo. Facultad de Economía y Empresa. Grado en ADE. Métodos Estadísticos para la Empresa

TEMA 5. COMPETENCIAS

- Saber calcular e interpretar la covarianza y el coeficiente de correlación lineal.
- Conocer las principales propiedades derivadas de la independencia de variables aleatorias.
- Saber aplicar e interpretar el Teorema Central del Límite.



Universidad de Oviedo. Facultad de Economía y Empresa. Grado en ADE. Métodos Estadísticos para la Empresa

MÉTODOS ESTADÍSTICOS PARA LA EMPRESA

TEMA 5: ANÁLISIS CONJUNTO DE VARIABLES ALEATORIAS Y DISTRIBUCIÓN DE AGREGADOS

5.1.- Distribuciones n-dimensionales. Análisis marginal y condicionado



Universidad de Oviedo. Facultad de Economía y Empresa. Grado en ADE. Métodos Estadísticos para la Empresa

Distribución conjunta de dos variables aleatorias

COCHES Se han analizado conjuntamente las variables número de hijos de cada familia (X) y número de coches por familia (Y)

Y \ X	1	2	3
1	0,19	0,16	0,09
2	0,06	0,19	0,31

- ❖ Valores (x_i, y_j)
- ❖ Probabilidades conjuntas p_{ij}

$$P(X=2, Y=1) = 0,16$$



Universidad de Oviedo. Facultad de Economía y Empresa. Grado en ADE. Métodos Estadísticos para la Empresa

Variable aleatoria bidimensional discreta

Observación conjunta de dos variables aleatorias unidimensionales discretas X e Y

(X,Y) v. a. DISCRETA $\{(x_i, y_j)/i, j = 1, 2, \dots\}$

$P:(x,y) \in \mathbb{R}^2 \rightarrow P(x,y) = P(X=x, Y=y) \in [0,1]$

$Y \setminus X$	x_1	...	x_i	...	x_k
y_1	p_{11}		p_{i1}		p_{k1}
...					
y_j	p_{1j}		p_{ij}		p_{kj}
...					
y_l	p_{1l}		p_{il}		p_{kl}

Probabilidad conjunta
 $p_{ij} = P(X=x_i, Y=y_j)$

$$P(X = x_i, Y = y_j) = p_{ij}$$

$$\sum_i \sum_j p_{ij} = 1$$

$$p_{ij} \geq 0$$



Universidad de Oviedo. Facultad de Economía y Empresa. Grado en ADE. Métodos Estadísticos para la Empresa

Distribuciones marginales

$Y \setminus X$	x_1	...	x_i	...	x_k
y_1	p_{11}		p_{i1}		p_{k1}
...					
y_j	p_{1j}		p_{ij}		p_{kj}
...					
y_l	p_{1l}		p_{il}		p_{kl}

Distribución marginal
de Y: $(y_j, p_{.j})$

$$p_{.j} = \sum_i p_{ij}$$

$$p_{.i} = \sum_j p_{ij}$$

Distribución marginal
de X: $(x_i, p_{.i})$



Universidad de Oviedo. Facultad de Economía y Empresa. Grado en ADE. Métodos Estadísticos para la Empresa

Distribuciones marginales

Las **distribuciones marginales** son distribuciones de v.a. unidimensionales, sobre las que:

- ➡ Se pueden calcular probabilidades y
- ➡ Se pueden obtener las correspondientes características marginales de X e Y:

E(X)
Var (X)
E(Y)
Var(Y)



Variables bidimensionales continuas

Observación conjunta de dos variables aleatorias unidimensionales continuas X e Y

(X,Y) v. a. CONTINUA

Función de densidad conjunta

$$f:(x,y) \in \mathbb{R}^2 \rightarrow f(x,y) \in \mathbb{R}$$

$$f(x,y) \geq 0 \quad \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x,y) dx dy = 1$$

Funciones de densidad marginales

$$f_Y(y) = \int_{-\infty}^{+\infty} f(x,y) dx$$

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y) dy$$



Función de distribución conjunta

Dada una v.a. bidimensional (X,Y) definimos la función de distribución conjunta asociada a esta variable como:

**Función de
distribución
conjunta**

$$F:(x,y) \in \mathbb{R}^2 \rightarrow F(x,y) = P(X \leq x, Y \leq y) \in [0,1]$$



Universidad de Oviedo. Facultad de Economía y Empresa. Grado en ADE. Métodos Estadísticos para la Empresa

Relación lineal entre variables aleatorias

Covarianza

$$\sigma_{X,Y} = \text{Cov}(X,Y) = E[(X - \mu_X)(Y - \mu_Y)]$$

$$\sigma_{X,Y} = E(XY) - E(X)E(Y)$$

Coefficiente de correlación lineal

$$\rho_{X,Y} = \frac{\sigma_{X,Y}}{\sigma_X \sigma_Y}$$

$$-1 \leq \rho_{X,Y} \leq 1$$

- El signo de $\rho_{X,Y}$ indica el tipo de dependencia lineal
- El valor absoluto de $\rho_{X,Y}$ indica el grado de relación lineal



Universidad de Oviedo. Facultad de Economía y Empresa. Grado en ADE. Métodos Estadísticos para la Empresa

Vectores aleatorios. Características

Vector aleatorio

Esperanza

Varianzas-Covarianzas

Unidimensional
 X

$$\mu = E(X)$$

$$\sigma^2 = \text{Var}(X)$$

Bidimensional
 (X, Y)

$$\mu = (E(X), E(Y))$$

$$\begin{pmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{YX} & \sigma_Y^2 \end{pmatrix}$$

n-dimensional
 $(X_1, X_2, \dots, X_n)$

$$\mu = (E(X_1), \dots, E(X_n))$$

$$\begin{pmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \dots & \sigma_{2n} \\ \dots & \dots & \dots & \dots \\ \sigma_{n1} & \sigma_{n2} & \dots & \sigma_n^2 \end{pmatrix}$$



Universidad de Oviedo. Facultad de Economía y Empresa. Grado en ADE. Métodos Estadísticos para la Empresa

MÉTODOS ESTADÍSTICOS PARA LA EMPRESA

TEMA 5: ANÁLISIS CONJUNTO DE VARIABLES ALEATORIAS Y DISTRIBUCIÓN DE AGREGADOS

5.2.- Variables aleatorias independientes. Propiedades



Universidad de Oviedo. Facultad de Economía y Empresa. Grado en ADE. Métodos Estadísticos para la Empresa

Variables aleatorias independientes

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

$$F(x,y) = F_X(x) F_Y(y) \quad \forall (x,y) \in \mathbb{R}^2$$

f. de distribución
de (X,Y)

f. de distribución
marginal de X

f. de distribución
marginal de Y

CONDICIÓN DE INDEPENDENCIA

(X,Y) DISCRETA

$$P(x_i, y_j) = P(x_i)P(y_j) \quad \forall (x_i, y_j) \in \mathbb{R}^2$$

(X,Y) CONTINUA

$$f(x,y) = f_X(x)f_Y(y) \quad \forall (x,y) \in \mathbb{R}^2$$



Universidad de Oviedo. Facultad de Economía y Empresa. Grado en ADE. Métodos Estadísticos para la Empresa

Propiedades

Dadas X e Y , v.a. **independientes** se cumple:

☐ $E(XY) = E(X)E(Y)$

☐ $\text{Cov}(X,Y) = 0$ y $\rho_{XY} = 0$

☐ $\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y)$

☐ $\text{Var}(aX+bY) = a^2\text{Var}(X) + b^2\text{Var}(Y) \quad \forall a,b \in \mathbb{R}$

☐ $\text{Var}(X-Y) = \text{Var}(X) + \text{Var}(Y)$

**X e Y v.a. normales e
incorreladas ($\rho_{XY}=0$)**

X e Y independientes



Universidad de Oviedo. Facultad de Economía y Empresa. Grado en ADE. Métodos Estadísticos para la Empresa

Reproductividad

Dada una familia de variables aleatorias ψ , se dice que ésta es reproductiva si y sólo si para todo par de v.a. independientes $X_1, X_2 \in \psi$ se cumple $X_1 + X_2 \in \psi$

V.a. independientes	Suma	Reproductividad
$X \approx B(n_X, p)$ $Y \approx B(n_Y, p)$	$X + Y \approx B(n_X + n_Y, p)$	Modelo Binomial reproductivo respecto a n (p constante)
$X \approx N(\mu_X, \sigma_X)$ $Y \approx N(\mu_Y, \sigma_Y)$	$X + Y \approx N(\mu_X + \mu_Y, \sqrt{\sigma_X^2 + \sigma_Y^2})$	Modelo Normal reproductivo respecto a media y varianza
$X \approx P(\lambda_X)$ $Y \approx P(\lambda_Y)$	$X + Y \approx P(\lambda_X + \lambda_Y)$	Modelo de Poisson reproductivo respecto al parámetro λ



Universidad de Oviedo. Facultad de Economía y Empresa. Grado en ADE. Métodos Estadísticos para la Empresa

MÉTODOS ESTADÍSTICOS PARA LA EMPRESA

TEMA 5: ANÁLISIS CONJUNTO DE VARIABLES ALEATORIAS Y DISTRIBUCIÓN DE AGREGADOS

5.3.- Agregación de variables aleatorias



Universidad de Oviedo. Facultad de Economía y Empresa. Grado en ADE. Métodos Estadísticos para la Empresa

Agregación de variables aleatorias

$$\left\{ \begin{array}{l} X_1 \text{ (Beneficio de la empresa 1)} \\ X_2 \text{ (Beneficio de la empresa 2)} \\ \dots\dots\dots \\ X_n \text{ (Beneficio de la empresa n)} \end{array} \right\}$$

Suma

$$S_n = \sum_{i=1}^n X_i$$

Beneficio total de las empresas

Media

$$\bar{X}_n = \frac{\sum_{i=1}^n X_i}{n}$$

Beneficio medio de las empresas



Universidad de Oviedo. Facultad de Economía y Empresa. Grado en ADE. Métodos Estadísticos para la Empresa

Características de las agregados

Dada la v.a. n-dimensional $(X_1, X_2, \dots, X_n)$ con esperanzas y varianzas finitas:

X_1	X_2	$\dots$	X_n
$E(X_1) = \mu_1$	$E(X_2) = \mu_2$	$\dots$	$E(X_n) = \mu_n$
$\text{Var}(X_1) = \sigma_1^2$	$\text{Var}(X_1) = \sigma_2^2$	$\dots$	$\text{Var}(X_1) = \sigma_n^2$

Esperanz

$$E(S_n) = \sum_{i=1}^n \mu_i \quad E(\bar{X}_n) = \frac{\sum_{i=1}^n \mu_i}{n}$$

Varianza

$$\text{Var}(S_n) = \text{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \sigma_i^2 + \sum_{i \neq j} \sigma_{ij}$$

$X_1, X_2, \dots, X_n$ independientes

$$\text{Var}(S_n) = \sum_{i=1}^n \sigma_i^2$$

$$\text{Var}(\bar{X}_n) = \frac{\sum_{i=1}^n \sigma_i^2}{n^2}$$



Universidad de Oviedo. Facultad de Economía y Empresa. Grado en ADE. Métodos Estadísticos para la Empresa

Desigualdad de Chebyshev

Sea $(X_1, X_2, \dots, X_n)$ v.a. n -dimensional con esperanzas y varianzas finitas

Acotación de Chebyshev para la suma:

$$P(|S_n - E(S_n)| \geq \varepsilon) \leq \frac{\text{Var}(S_n)}{\varepsilon^2}$$

Acotación de Chebyshev para la media:

$$P(|\bar{X}_n - E(\bar{X}_n)| \geq \varepsilon) \leq \frac{\text{Var}(\bar{X}_n)}{\varepsilon^2}$$



Universidad de Oviedo. Facultad de Economía y Empresa. Grado en ADE. Métodos Estadísticos para la Empresa

Variables aleatorias i.i.d.

Dadas $X_1, X_2, \dots, X_n$ v.a. **independientes e idénticamente distribuidas** con:

$$X_i \quad E(X_i) = \mu \quad \text{Var}(X_i) = \sigma^2 \quad \forall i=1, \dots, n$$

MAGNITUD	ESPERANZA	VARIANZA	ACOTACIÓN DE CHEBYSHEV
Suma (S_n)	$E(S_n) = n\mu$	$\text{Var}(S_n) = n\sigma^2$	$P(S_n - n\mu \geq \varepsilon) \leq \frac{n\sigma^2}{\varepsilon^2}$
Media ($\bar{X}_n$)	$E(\bar{X}_n) = \mu$	$\text{Var}(\bar{X}_n) = \frac{\sigma^2}{n}$	$P(\bar{X}_n - \mu \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}$



Universidad de Oviedo. Facultad de Economía y Empresa. Grado en ADE. Métodos Estadísticos para la Empresa

MÉTODOS ESTADÍSTICOS PARA LA EMPRESA

TEMA 5: ANÁLISIS CONJUNTO DE VARIABLES ALEATORIAS Y DISTRIBUCIÓN DE AGREGADOS

5.4.-Teorema Central del Límite y sus aplicaciones



Universidad de Oviedo. Facultad de Economía y Empresa. Grado en ADE. Métodos Estadísticos para la Empresa

Teorema Central del Límite

Enunciado de Levy-Lindeberg:

Dadas $X_1, X_2, \dots, X_n$ v.a. **independientes e idénticamente**

distribuidas con X_i $E(X_i)=\mu$ $Var(X_i)=\sigma^2$ finitas $\forall i=1,\dots,n$

$$S_n = \sum_{i=1}^n X_i \longrightarrow N(n\mu, \sigma\sqrt{n})$$

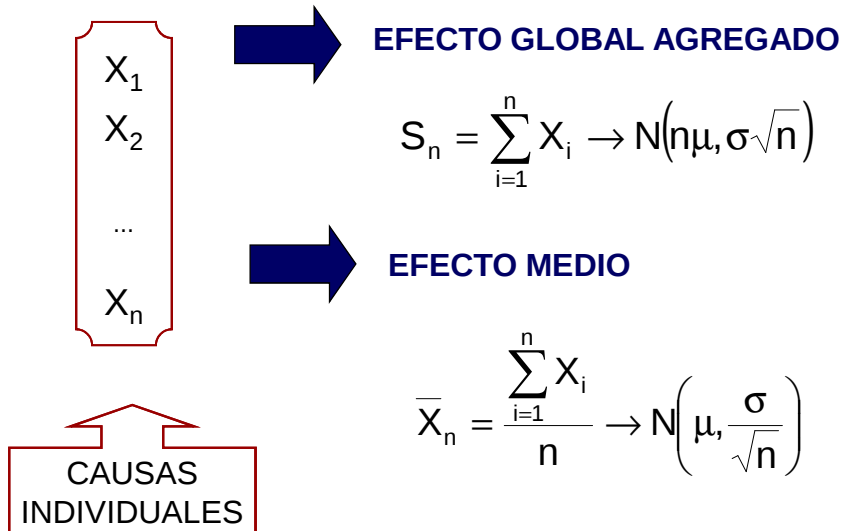
$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \longrightarrow N(0,1)$$

Aproximación $n > 30$



Universidad de Oviedo. Facultad de Economía y Empresa. Grado en ADE. Métodos Estadísticos para la Empresa

Teorema Central del Límite



Universidad de Oviedo. Facultad de Economía y Empresa. Grado en ADE. Métodos Estadísticos para la Empresa

Teorema Central del Límite

Enunciado de De Moivre:

$X_1, X_2, \dots, X_n$ v.a. de **Bernoulli**, de parámetro p , e

Independientes:

$$\frac{\sum_{i=1}^n X_i - np}{\sqrt{npq}} \longrightarrow N(0,1)$$

→ Consecuencia:

$$X_i \approx B(p)$$

$$\sum_{i=1}^n X_i \approx B(n,p) \longrightarrow N(np, \sqrt{npq})$$

Corrección
de
continuidad



Universidad de Oviedo. Facultad de Economía y Empresa. Grado en ADE. Métodos Estadísticos para la Empresa

Aproximación BINOMIAL → NORMAL

