ESTAD _ T14. DISTR. MULTINOMIAL. DISTR. NORMAL MULTIVARIANTE. PROPIEDADES. (incompleto, faltan demostraciones)

1_ DISTRIB. MULTINOMIAL , Multinomial $(n, p_1 - p_k)$

la distrib. multinomial es una distribución discreta multivar.

Se puede considerar una generalización de la Binomial, ya que el nº de resultados posibles (Éxito/Fracaso en la Binomial) es mayor o igual que 3 (pero finito?).

Si consideramos un experimento aleatorio donde los resultados posibles constituyen una partición del espacio muestral:

$A_1, A_2 \dots A_k$ tq $E = A_1 \cup A_2 \cup \dots \cup A_k$

$A_i \cap A_j = \emptyset \ , \forall i \neq j$

y designamos por $p_1 \dots p_k$ las probabilidades de los diferentes sucesos: $P(A_i) = p_i \ , i = 1 \dots k$

bien definidas $\rightarrow \sum_{i=1}^{k} p_i = \sum_{i=1}^{k} P(A_i) = P(\bigcup_{i=1}^{k} A_i) = P(E) = 1$.

El experimento consiste en la realización n veces de la misma prueba de manera independiente, donde las prob. de ocurrencia se mantienen constantes. La variable aleatoria multidimensional que agrupa las v.a. unidim. que miden el nº de veces que aparece cada resultado sigue un distribución multinomial.

$$\xi = (\xi_1, \dots, \xi_k) \longrightarrow \text{Multinomial}(n, p_1, \dots, p_k)$$

donde $\xi_i \equiv$ nº veces que aparece A_i

$\xi_i \rightarrow B(n, p_i)$??

Conceptualmente, un fenómeno se distribuye con arreglo a una distrib. multinomial si:

1_ El experimento consiste en la realización de n ensayos.

2_ Los resultados constituyen una partición del espacio muestral.

Cada resultado lleva asociado una probab. de ocurrencia

3_ las probab. de ocurrencia suman 1

4_ las probab. permanecen constantes a lo largo de todo el experimento.

5_ los ensayos son indep. entre sí.

La distrib. multinomial depende de $k+1$ parámetros, el nº de ensayos y las probab. de ocurrencia de cada suceso. $(n, p_1 - p_k)$

Definición: $\xi \equiv$ nº de veces que aparece cada uno de los posibles resultados A_i, $i = 1 \dots k$, cuando se repite n veces la misma prueba de manera independiente.

Función de cuantía:

$$P(\xi = (x_1, x_2, \dots, x_k)) = \underbrace{PR^{n}_{x_1, x_2 \dots x_k}}_{\text{Combinatoria}} \cdot \underbrace{P(\xi_1 = x_1, \xi_2 = x_2 \dots \xi_k = x_k)}_{\text{1 opción concreta}}$$

$$P(\xi_1 = x_1, \xi_2 = x_2 \dots \xi_k = x_k) = \underbrace{P(\xi_1 = x_1) \cdot P(\xi_2 = x_2) \dots P(\xi_k = x_k)}_{\text{indep}}$$

$$= \frac{n!}{x_1! \dots x_k!} \cdot p_1^{x_1} \cdot p_2^{x_2} \cdot \dots \cdot p_k^{x_k}$$

para $x_i = 0, 1 \dots n \;/\; \sum_{i=1}^{k} x_i = n$

Efectivamente, es una función de probabilidad (está bien definida) porque:

$P(x_1 \dots x_k) \geq 0 \quad , \forall \quad (x_1 \dots x_k)$

$$\sum_{x_1} \dots \sum_{x_k} P(x_1, \dots, x_k) = \sum_{x_1} \dots \sum_{x_k} \frac{n!}{x_1! \dots x_k!} p_1^{x_1} \cdot \dots \cdot p_k^{x_k} =$$

$$= (p_1 + \dots + p_k)^n = 1^n = 1 .$$

Función de distribución: Probabilidad acumulada

$$F(x_1^1, x_2^2, \dots x_k^k) = P(\xi_1 \leq x_1^1, \xi_2 \leq x_2^2, \dots, \xi_k \leq x_k^k) =$$

$$= \sum_{x_1 \leq x_1^1} \sum_{x_2 \leq x_2^2} \dots \sum_{x_k \leq x_k^k} P(x_1, \dots, x_k) =$$

$$= \sum_{x_1 \leq x_1^1} \dots \sum_{x_k \leq x_k^k} \frac{n!}{x_1! \dots x_k!} p_1^{x_1} \dots p_k^{x_k} .$$

para $x_i = 0, 1, \dots n$ / tq $\sum_{i=1}^{k} x_i = n$.

Características $t = (t_1 \dots t_k)$

Función característica, $\varphi(t) = E[e^{it_1\xi_1 + \dots + it_k\xi_k}] =$

$$= \sum_{x_1} \dots \sum_{x_k} e^{it_1x_1 + \dots it_kx_k} \cdot \frac{n!}{x_1! \dots x_k!} p_1^{x_1} \cdot \dots \cdot p_k^{x_k} =$$

$$= \sum_{x_1} \dots \sum_{x_k} \frac{n!}{x_1! \dots x_k!} (e^{it_1}p_1)^{x_1} \cdot (e^{it_2}p_2)^{x_2} \cdot \dots \cdot (e^{it_k}p_k)^{x_k} \overset{\text{fórmula de Leibniz}}{=}$$

$$= (e^{it_1}p_1 + \dots + e^{it_k}p_k)^{x_1 + \dots + x_k} = \sum_{j=1}^{k} (p_j e^{it_j})^n$$

$$\varphi(t) = \sum_{j=1}^{k} (p_j e^{it_j})^n$$

Esperanza $= \alpha_1 = \frac{1}{i} \left| \frac{\partial \varphi(t)}{\partial t} \right|_{t=0}$

Al ser una var. multidimensional, tendrá tantas esperanzas como posibilidades presente el experimento.
La esperanza α_i será el momento respecto al origen de orden 1, marginal de la var. ξ_i.

$\varphi(t_j) = \varphi(0, \dots, t_j \dots 0) = (p_1 + \dots + p_j e^{it_j} + \dots + p_k)^n$

$\frac{\partial \varphi(t_j)}{\partial t_j} = n(p_1 + \dots + p_j e^{it_j} + \dots + p_k)^{n-1} \cdot p_j e^{it_j} \cdot i = n i p_j e^{it_j} (\dots)$

$\left. \frac{\partial \varphi(t_j)}{\partial t_j} \right|_{t_j=0} = n \cdot i \cdot p_j (\underbrace{p_1 + \dots + p_j + \dots + p_k}_{1})^{n-1} = n i \cdot p_j$

$\frac{1}{i} \left| \frac{\partial \varphi(t_j)}{\partial t_j} \right|_{t=0} = n p_j$

como es binomial

Esperanza de $\xi_i \rightarrow E[\xi_i] = n p_i \quad , i = 1 \dots k$

Varianza $= \mu_2 = \alpha_2 - \alpha_1^2$

$\alpha_2 = \frac{1}{i^2} \left| \frac{\partial \varphi^2(t)}{\partial t^2} \right|_{t=0}$

$\frac{\partial^2 \varphi(t_j)}{\partial t_j^2} = n i^2 p_j e^{it_j} (p_1 + \dots + p_j e^{it_j} + \dots + p_k)^{n-1} +$
$+ n i p_j e^{it_j} (n-1)(p_1 + \dots + p_j e^{it_j} + \dots + p_k)^{n-2} \cdot i p_j e^{it_j} =$

$\left. \frac{\partial^2 \varphi(t_j)}{\partial t_j^2} \right|_{t=0} = n i^2 p_j (p_1 + \dots + p_k)^{n-1} + n i p_j (n-1)(p_1 + \dots + p_k)^{n-2} \cdot i \cdot p_j =$

$= n i^2 p_j + n(n-1) i^2 p_j^2$

$\frac{1}{i^2} \left| \frac{\partial^2 \varphi(t_j)}{\partial t_j^2} \right|_{t=0} = n p_j + n(n-1) p_j^2 = n^2 p_j^2 - n p_j^2 + n p_j$

Así, $Var(\xi_j) = \cancel{n^2 p_j^2} - np_j^2 + np_j - \cancel{(np_j)^2} =$

$= np_j - np_j^2 = np_j(1-p_j)$, para $j = 1 \dots K$

<u>Covarianza</u>, $\mu_{11} = \alpha_{11} - \alpha_{10}\alpha_{01}$

$$Cov(\xi_i, \xi_j) = E(\xi_i \cdot \xi_j) - E(\xi_i)\cdot E(\xi_j) = \boxed{-np_ip_j}$$

$$E(\xi_i \cdot \xi_j) = \frac{1}{i^2}\left|\frac{\partial^2 \varphi(t_i,t_j)}{\partial t_i \partial t_j}\right|_{t_i=t_j=0} \overset{\text{Utilizando}}{=} n^2 p_i p_j - np_ip_j$$

$$Cov(\xi_i,\xi_j) = \cancel{n^2 p_ip_j} - np_ip_j - \cancel{n^2 p_ip_j} = -np_ip_j$$

$$\frac{\partial \varphi(t_i,t_j)}{\partial t_i} = ni\,p_ie^{it_i}(p_1 + \dots + p_ie^{it_i} + \dots + p_je^{it_j} + \dots + p_K)^{n-1}$$

$$\frac{\partial^2 \varphi(t_i,t_j)}{\partial t_i \partial t_j} = nip_ie^{it_i}(n-1)(p_1 + \dots + p_ie^{it_i} + \dots + p_je^{it_j} + \dots + p_K)^{n-2} \cdot$$

$$\cdot\, p_j i e^{it_j}$$

$$\frac{1}{i^2}\left|\frac{\partial^2 \varphi(t_i,t_j)}{\partial t_i \partial t_j}\right|_{t_i=t_j=0} = \frac{1}{\cancel{i^2}} n\cancel{i^2} p_i(n-1)p_j = n^2 p_ip_j - np_ip_j$$

<u>PROPIEDADES</u>

P1 → Propiedad reproductiva: La suma de dos multinomiales <u>indep.</u> con las mismas probabilidades tb. es multinomial.

$$\left.\begin{array}{l} \xi \rightarrow M(n, p_1 \dots p_K) \\ \eta \rightarrow M(m, p_1 \dots p_K) \end{array}\right\} \gamma = \xi + \eta \rightarrow M(n+m, p_1 \dots p_K)$$

Dem: Al ser distribuciones independientes, la función característica de la suma es el producto de las funciones caract. individuales que, al ser un producto de potencias con la misma base es la base con la suma de los exponentes.

$$\varphi_{\xi+\eta}(t_1 \dots t_k) = E\left[e^{it_1(x_1+y_1)+\dots+it_k(x_k+y_k)}\right] =$$
$$= E\left[e^{it_1x_1+\dots+it_kx_k}\right] \cdot E\left[e^{it_1y_1+\dots+it_ky_k}\right] =$$
$$= \left(\sum_{i=1}^{k} p_i e^{it_i}\right)^n \cdot \left(\sum_{i=1}^{k} p_i e^{it_i}\right)^m = \left(\sum_{i=1}^{k} p_i e^{it_i}\right)^{n+m}$$

P2 ⟹ La distribución multinomial es la distribución límite de la distrib. hipergeométrica multivariante.

$$HM(N; n, p_1 \dots p_k) \xrightarrow[N \to +\infty]{} M(n, p_1 \dots p_k)$$

N = tamaño poblacional = nº individuos antes de realizar la 1ª extracción sin reemplaz.

n = nº extracciones

2_ DISTRIB. NORMAL MULTIVARIANTE.

a) Distrib. normal bidimensional

Definición: Una variable aleatoria bidimensional (ξ_1, ξ_2) de tipo continuo sigue una distrib. normal bivariante de parámetros:

μ_1 y μ_2 $\quad -\infty < \mu_i < +\infty$

σ_1 y σ_2 $\quad \cancel{\sigma_i < +\infty}\; 0 < \sigma_i$

ρ $\quad -1 < \rho < 1$ ← IMPORTANTE ($\rho \neq \pm 1$)

si su función de densidad es:

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \cdot e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\cdot\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]}$$

Una expresión equivalente es:

$$f(x_1, x_2) = \frac{1}{2\pi\sqrt{|\Sigma|}} \cdot e^{-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)}$$, donde:

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \quad \text{y} \quad \Sigma = E\left[(x-\mu)_{2\times1}(x-\mu)'_{1\times2}\right] =$$

esperanzas marginales

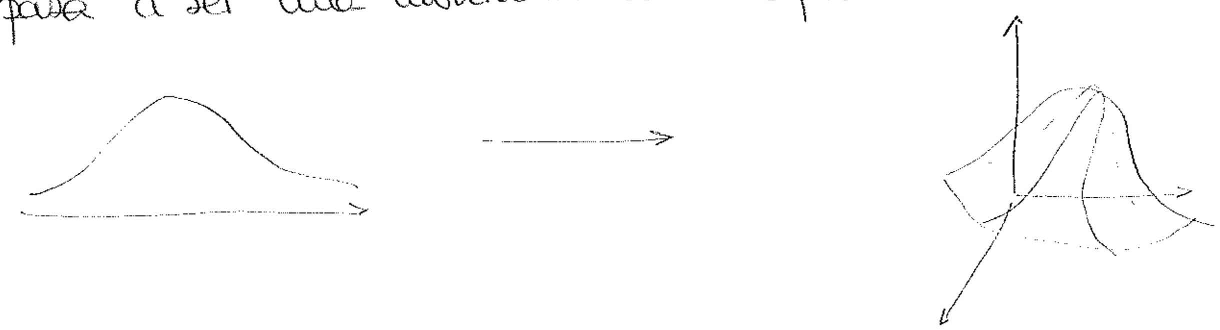
$$= \underbrace{\begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}}_{\text{matriz de var-cov.}}$$

De manera abreviada se puede expresar la normal bivariante como $N_2(\mu, \Sigma)$ donde $\mu \equiv$ vector medias

$\Sigma \equiv$ matriz var-cov.

$|\Sigma| \neq 0$

Gráficamente, la campana de Gauss en el plano pasa a ser una montaña en el espacio.



la probabilidad conjunta de que la v.a. bidimensional (ξ_1, ξ_2) tome valores en una región R (R = intervalo $\mathbb{R}^2$) viene dado por:

$$R = \{ (x_1, x_2) \in \mathbb{R}^2 / \ a_1 \leq x_1 \leq b_1 , \ a_2 \leq x_2 \leq b_2 \}$$

$$P((\xi_1, \xi_2) \in R) = P(a_1 \leq x_1 \leq b_1, a_2 \leq x_2 \leq b_2) =$$

$$= \int_{a_1}^{b_1} \int_{a_2}^{b_2} f(x_1 x_2) \, dx_1 \, dx_2 \leftarrow$$ volumen comprendido entre la superficie y R.

Función característica conjunta

$$\varphi(t_1, t_2) = e^{i(\mu_1 t_1 + \mu_2 t_2) - \frac{1}{2}(t_1^2 \sigma_1^2 + t_2^2 \sigma_2^2 - 2\rho \sigma_1 \sigma_2 t_1 t_2)}$$

Por las propiedades de f. característica sabemos que:

$$\left.\begin{aligned} \varphi(t_1, 0) &= \varphi_{\xi_1}(t) = e^{i\mu_1 t_1 - \frac{1}{2} t_1^2 \sigma_1^2} \\ \varphi(0, t_2) &= \varphi_{\xi_2}(t) = e^{i\mu_2 t_2 - \frac{1}{2} t_2^2 \sigma_2^2} \end{aligned}\right\}$$ f. características de las distrib. marginales

las f. características de las distrib. marginales corresponden a la f. característica de una Normal unidimensional, por lo que los parámetros de la distrib. conjunta cobran significado:

$\mu_i \rightarrow$ media marginal de ξ_i
$\sigma_i \rightarrow$ desviación típica de ξ_i
$\rho \rightarrow$ coef. correlación de ξ_1 y ξ_2

$\rho = \frac{COV}{\sigma_1 \sigma_2}$
$COV = \mu_{11} = \alpha_{11} - \alpha_{10}\alpha_{01}$
Hay que calcular $\alpha_{11} = \left.\frac{\partial^2 \varphi}{\partial t_1 \partial t_2}\right|_0 / i^2$

MIRAR ATRÁS $\Longrightarrow$

Funciones de densidad marginales

$$f_1(x_1) = \frac{1}{\sigma_1 \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x_1 - \mu_1}{\sigma_1}\right)^2} \qquad -\infty < x_1 < +\infty$$

$$f_2(x_2) = \frac{1}{\sigma_2 \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x_2 - \mu_2}{\sigma_2}\right)^2} \qquad -\infty < x_2 < +\infty$$

ESTAD_ T14

Distribuciones condicionales

$$f(x_2/x_1) = \frac{f(x_1,x_2)}{f_1(x_1)} = \frac{\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \cdot e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - \frac{2\rho(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]}}{\frac{1}{\sigma_1\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2}}$$

$$= \frac{1}{\sigma_2\sqrt{1-\rho^2}\sqrt{2\pi}} \cdot e^{-\frac{1}{2(1-\rho^2)\sigma_2^2}\left[x_2 - \left(\mu_2 + \rho\, \sigma_2/\sigma_1 (x_1-\mu_1)\right)\right]^2}$$

que es la función de densidad de una distrib. normal

de media $\rightarrow \mu_2 + \rho\, \sigma_2/\sigma_1\, (x_1-\mu_1)$

de desv. típica $\rightarrow \sigma_2\sqrt{1-\rho^2}$

Observación $\rightarrow$ Normal bivariantes $\Rightarrow$ $\left\{\begin{matrix}\text{marginales} \\ \text{condicionadas}\end{matrix}\right\}$ normales univariantes

$\not\Leftarrow$

PROPIEDADES

P1 $\rightarrow$ Independencia: Si ξ_1 y ξ_2 son indep, entonces el coeficiente de correlación lineal es nulo, $\rho = 0$, y la función de densidad conjunta queda

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{1}{2}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x-\mu_2}{\sigma_2}\right)^2\right]}$$

que verifica $f(x,y) = f_1(x) \cdot f_2(y)$

OJO: (ξ_1,ξ_2) indep $\xRightarrow{(\sigma_{12}=0)} \rho = 0$

$\Leftarrow$ Sólo en el caso Normal

Obs $\rightarrow$ Partiendo de dos distrib. normales no independientes, siempre es posible obtener independencia rotando los ejes.

P2 $\rightarrow$ Dada (ξ_1, ξ_2) una v.a. bidimensional, la combinación lineal de sus componentes es una normal unidim.

P3 $\rightarrow$ Sean $\xi_1, \dots \xi_n$ v.a.i.d $N(\mu_i, \sigma_i)$.
Definimos dos variables combinaciones lineales de las primeras (con $\neq$s coef, no todos nulos)
$\eta = \sum a_i \xi_i \quad ; \gamma = \sum b_i \xi_i$
La función de densidad conjunta de (η, γ) es la de una normal bidimensional.

Consideraciones sobre ρ

$-1 < \rho < 1 \Rightarrow \Sigma$ no singular $\Rightarrow \exists$ f. densidad conjunta
$\exists$ f. característica conjunta

$|\rho| = 1 \Rightarrow \rho = \pm 1 \Rightarrow \Sigma$ singular $\Rightarrow \nexists$ f. densidad conjunta
$\exists$ f. característica cjk (siempre)
$\downarrow$
distrib. degenerada

b) Distribución normal n-dimensional

Podemos generalizar lo visto anteriormente con dos variables al caso de más de dos.

$(\xi_1 \dots \xi_n)$ es una v.a. normal multivariante si tiene por función de densidad conjunta:

$$f(x_1 \dots x_n) = \frac{1}{\sqrt{(2\pi)^n \cdot |\Sigma|}} \cdot e^{-\frac{1}{2}(X-\mu)^t \Sigma^{-1} (X-\mu)}$$

donde: $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$; $\mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}$, $\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1n} \\ \vdots & \sigma_2^2 & & \\ \vdots & & \ddots & \\ \sigma_{1n} & & \dots & \sigma_n^2 \end{pmatrix}$