

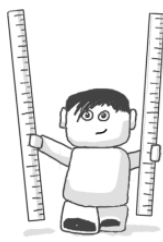


Teoría de probabilidades IV

Curso 2023-24

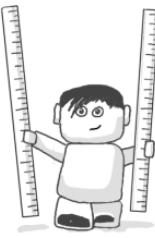
Facultad de Física

Técnicas Experimentales II



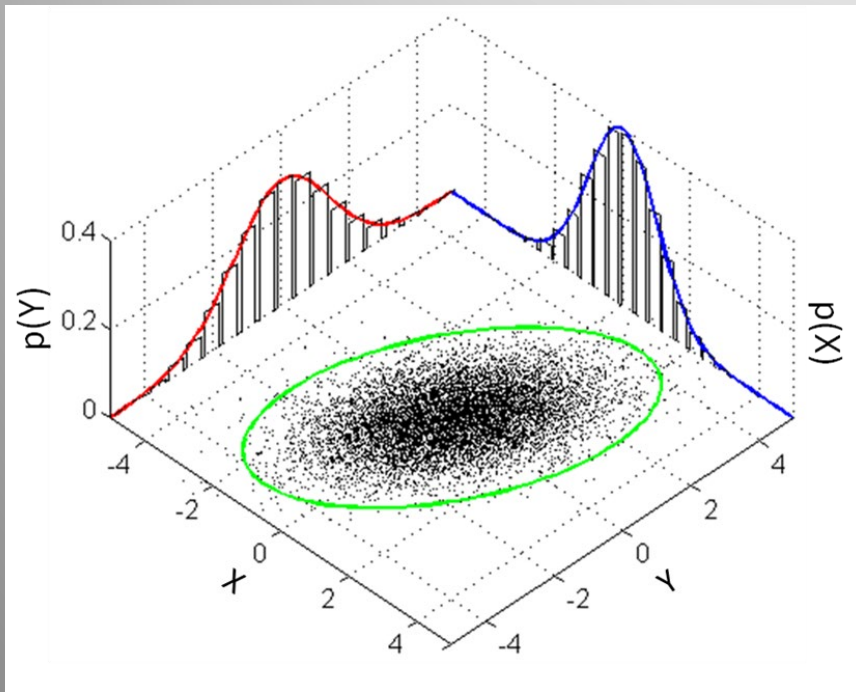
Objetivos tema:

- **Distribuciones de probabilidad multidimensionales. Probabilidades marginales. Probabilidad condicional. Teorema de Bayes. Medidas características.**
- **Suma de variables aleatorias.**
- **Teorema del límite central.**
- **Distribución multinomial.**
- **Rotación del espacio bivariante.**
- **Distribución normal bivariante.**



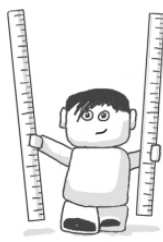
Distribuciones multidimensionales

En muchos casos, un suceso aleatorio puede implicar la determinación de varias magnitudes simultáneamente. En estos casos la variable aleatoria se denomina variable aleatoria vectorial o multidimensional.



$$\begin{array}{ccc} \Omega & \xrightarrow{X} & \mathbb{R}^n \\ A & \hookrightarrow & X(A) = (x_1, x_2, \dots, x_n) \end{array}$$

Las componentes x_i de los valores de la variable aleatoria (x_1, x_2, \dots, x_n) pueden tomar valores discretos o continuos. Si existen componentes de ambos tipos diremos que es una variable aleatoria mixta.



Distribuciones multidimensionales

Para las variables aleatorias multidimensionales definiremos una función de probabilidad o densidad de probabilidad conjunta.

$$\begin{array}{ccc} \Omega & \xrightarrow{X} & \mathbb{R}^2 \\ A & \hookrightarrow & X(A) = (x, y) \end{array}$$

Variable discreta $p(x_i, y_j)$

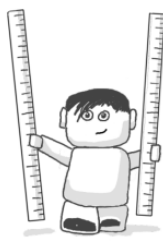
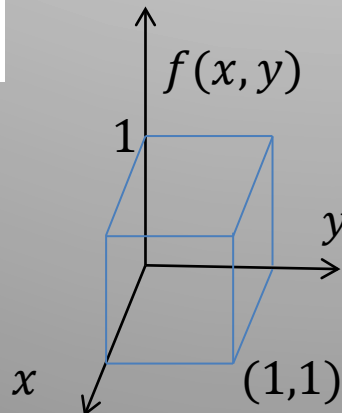
$$0 \leq p(x_i, y_j) \leq 1 \quad \forall i, j$$

$$\sum_i \sum_j p(x_i, y_j) = 1$$

Variable continua $f(x, y)$

$$0 \leq f(x, y) \quad x, y \in \mathbb{R}$$

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



Distribuciones marginales

Son aquellas distribuciones de probabilidad para un cierto valor x_i de una componente de la variable aleatoria con independencia de los valores que puedan adoptar el resto de las componentes:

Variable discreta $p(x_i, y_j)$

$$p(x_i) = \sum_j p(x_i, y_j)$$

$$p(y_j) = \sum_i p(x_i, y_j)$$

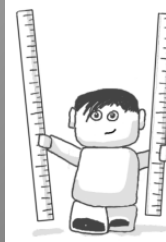
$$\sum_i p(x_i) = \sum_j p(y_j) = 1$$

Variable continua $f(x, y)$

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

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

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



Probabilidad condicionada

Análogamente a la definición realizada al principio del capítulo, podremos calcular la probabilidad condicionada a partir de la probabilidad conjunta y las marginales de la variable aleatoria multidimensional.

Variable discreta $p(x_i, y_j)$

$$p(x_i|y_j) = \frac{p(x_i, y_j)}{p(y_j)}$$

$$p(y_j|x_i) = \frac{p(x_i, y_j)}{p(x_i)}$$

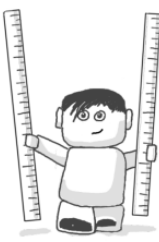
$$\begin{aligned} p(x_i, y_j) &= p(x_i|y_j) p(y_j) = \\ &= p(y_j|x_i) p(x_i) \end{aligned}$$

Variable continua $f(x, y)$

$$f(x|y) = \frac{f(x, y)}{f(y)}$$

$$f(y|x) = \frac{f(x, y)}{f(x)}$$

$$\begin{aligned} f(x, y) &= f(x|y) f(y) = \\ &= f(y|x) f(x) \end{aligned}$$



Independencia estadística

Dos variables x e y son estadísticamente independientes cuando los valores que adopta una de ellas no tienen ninguna influencia en aquellos que adopta la otra, de tal modo que:

Variable discreta $p(x_i, y_j)$

$$p(x_i|y_j) = p(x_i)$$



$$p(x_i) p(y_j) = p(x_i, y_j)$$



$$p(y_j|x_i) = p(y_j)$$

Variable continua $f(x, y)$

$$f(x|y) = f(x)$$

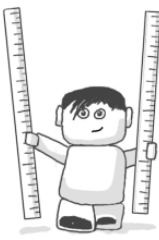


$$f(x) f(y) = f(x, y)$$



$$f(y|x) = f(y)$$

Para variables independientes la función (densidad) de probabilidad será el producto de las funciones (densidad) de probabilidad marginales.



Teorema de Bayes

En este caso de variables multidimensionales también podemos escribir el teorema de Bayes

Variable discreta $p(x_i, y_j)$

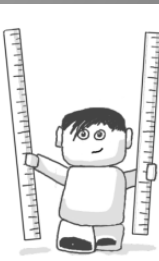
$$p(y_k|x_i) = \frac{p(x_i|y_k) p(y_k)}{\sum_j p(x_i|y_j) p(y_j)}$$

$$p(x_l|y_j) = \frac{p(y_j|x_l) p(x_l)}{\sum_i p(y_j|x_i) p(x_i)}$$

Variable continua $f(x, y)$

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

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



Medidas características: media

Definimos como la media de una variable aleatoria bidimensional

$$\begin{array}{ccc} \Omega & \xrightarrow{X} & \mathbb{R}^2 \\ A & \hookrightarrow & X(A) = (x, y) \end{array}$$

Al vector 2-dimensional formado por los valores esperados μ_i de las distribuciones marginales de probabilidad:

$$\vec{\mu} = (\mathbb{E}\{x\}, \mathbb{E}\{y\}) = (\mu, \nu)$$

Siendo:

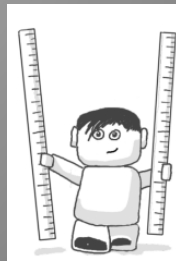
Discreto

$$\mu = \mathbb{E}\{x\} = \sum_i x_i p(x_i) = \sum_i x_i \left(\sum_j p(x_i, y_j) \right)$$

Continuo

$$\mu = \mathbb{E}\{x\} = \int_{-\infty}^{+\infty} f(x) x \, dx = \int_{-\infty}^{+\infty} x \, dx \left(\int_{-\infty}^{+\infty} f(x, y) \, dy \right)$$

Recíprocamente para la variable y



Medidas características: varianza

También podemos definir las varianzas de las distribuciones marginales de la variable aleatoria como

$$\sigma_x^2 = \mathbb{E}\{(x - \mu)^2\}$$

$$\sigma_y^2 = \mathbb{E}\{(y - \nu)^2\}$$

Discreto

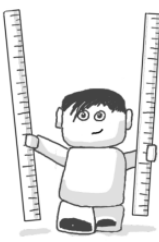
$$\sigma_x^2 = \sum_i (x_i - \mu)^2 p(x_i)$$

$$\sigma_y^2 = \sum_j (y_j - \nu)^2 p(y_j)$$

Continuo

$$\sigma_x^2 = \int_{-\infty}^{+\infty} f(x) (x - \mu)^2 dx$$

$$\sigma_y^2 = \int_{-\infty}^{+\infty} f(y) (y - \nu)^2 dy$$



Medidas características: covarianza

Considerando una variable aleatoria bidimensional (x, y) , la covarianza es una medida de la correlación estadística entre dos variables:

$$\mathbb{E}\{x\} = \mu$$

$$\mathbb{E}\{y\} = \nu$$

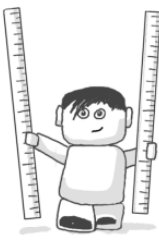
$$\text{cov}(x, y) = \mathbb{E}\{(x - \mu)(y - \nu)\} = \mathbb{E}\{xy\} - \mu\nu$$

Discreto

$$\text{cov}(x, y) = \sum_{i,j} (x_i - \mu) (y_j - \nu) p(x_i, y_j)$$

Continuo

$$\text{cov}(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \mu)(y - \nu) f(x, y) dx dy$$



Medidas características: covarianza

En el caso de variables estadísticamente independientes se cumple que:

$$p(x_i, y_j) = p(x_i) p(y_j)$$

$$f(x, y) = f(x) f(y)$$

Por lo tanto, podremos escribir que:

$$\text{cov}(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \mu)(y - \nu) f(x, y) dx dy$$

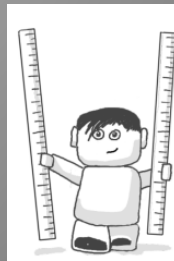
$$\text{cov}(x, y) = \int_{-\infty}^{+\infty} dx (x - \mu) f(x) \int_{-\infty}^{+\infty} dy (y - \nu) f(y) = 0$$

N.B.: El recíproco no es cierto

$$\text{cov}(x, y) = 0$$



$$f(x, y) = f(x)f(y)$$



Medidas características: coeficiente de correlación

La covarianza de dos magnitudes tiene las unidades de éstas y es susceptible de variar si realizamos una transformación de escala en x o y

$$\tilde{x} = a x$$

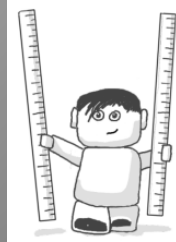
$$\tilde{y} = b y$$

$$\text{cov}(\tilde{x}, \tilde{y}) = \mathbb{E}\{(a x - a \mu)(b y - b \nu)\} = a b \text{cov}(x, y)$$

Para obtener una medida característica adimensional del grado de correlación estadística se suele usar el coeficiente de correlación definido como

$$\rho(x, y) = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y}$$

$$0 \leq |\rho(x, y)| \leq 1$$



Medidas características: matriz de varianzas/covarianzas

Si consideramos dos variables aleatorias x e y denominaremos la matriz de covarianzas como

$$M = \begin{pmatrix} \sigma^2(x) & \text{cov}(x, y) \\ \text{cov}(x, y) & \sigma^2(y) \end{pmatrix}$$

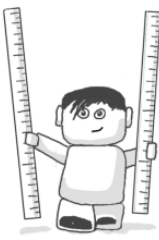
Es una matriz simétrica que cumple que:

$$\det(M) = \sigma^2(x) \sigma^2(y) - \text{cov}^2(x, y)$$

Por lo tanto:

$$\frac{\det(M)}{\sigma^2(x) \sigma^2(y)} = 1 - \rho^2 \geq 0$$

Si $\det(M) = 0$ entonces $\rho = \pm 1$ por lo que las variables siguen una asociación lineal.



Distribución uniforme bivalente

Consideremos una distribución de probabilidad de la forma:

$$f(x, y) = \begin{cases} \frac{1}{(b-a)(d-c)}; & a \leq x \leq b; \quad c \leq y \leq d \\ 0 & \text{en otro caso} \end{cases}$$

$$f(x) = \int_c^d f(x, y) dy = \int_c^d \frac{1}{(b-a)(d-c)} dy = \frac{1}{(b-a)}; \quad a \leq x \leq b$$

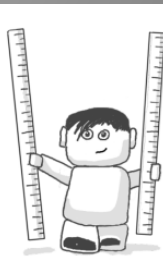
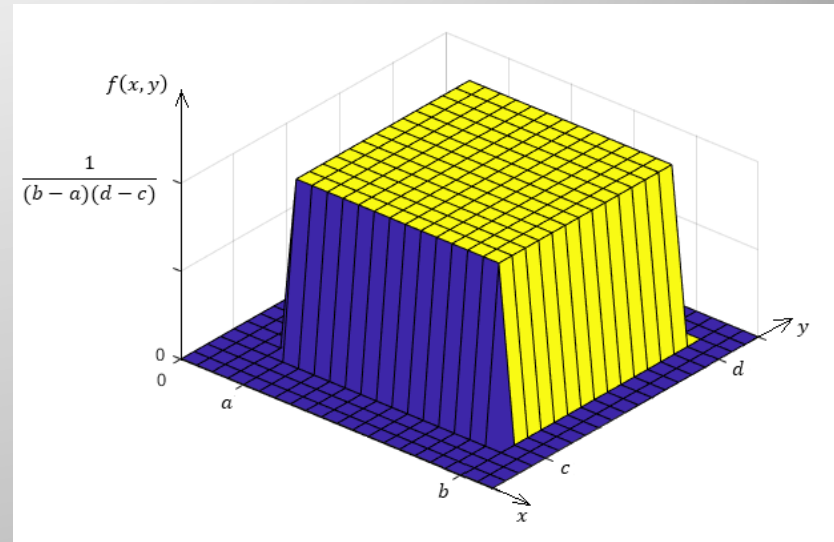
$$f(y) = \int_a^b f(x, y) dx = \int_a^b \frac{1}{(b-a)(d-c)} dx = \frac{1}{(d-c)}; \quad c \leq y \leq d$$

$$\mathbb{E}\{x\} = \frac{a+b}{2}$$

$$\mathbb{E}\{y\} = \frac{c+d}{2}$$

$$\sigma^2(x) = \frac{(b-a)^2}{12}$$

$$\sigma^2(y) = \frac{(d-c)^2}{12}$$



Distribución uniforme bivalente

Podremos también calcular la covarianza:

$$f(x, y) = \begin{cases} \frac{1}{(b-a)(d-c)}; & a \leq x \leq b; \quad c \leq y \leq d \\ 0 & \text{en otro caso} \end{cases}$$

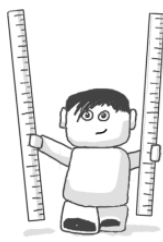
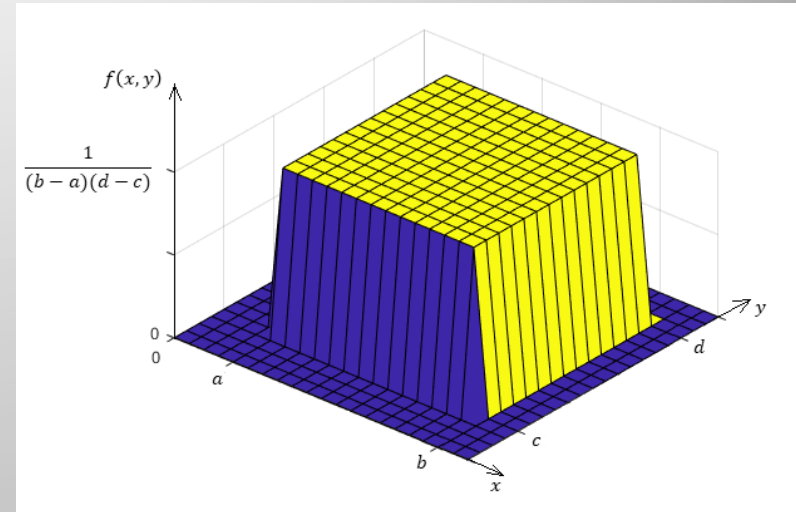
$$\text{cov}(x, y) = \int_a^b dx \int_c^d dy \frac{1}{(b-a)(d-c)} \left(x - \frac{a+b}{2} \right) \left(y - \frac{c+d}{2} \right)$$

En este caso se cumple:

$$\text{cov}(x, y) = 0$$

Ya que las variables aleatorias son estadísticamente independientes pues se verifica:

$$f(x, y) = \frac{1}{(b-a)(d-c)} = f(x)f(y) = \frac{1}{(b-a)} \frac{1}{(d-c)}$$



Distribución lineal bivalente

Consideremos la función de densidad:

$$f(x, y) = \begin{cases} x + y; & 0 \leq x \leq 1; \quad 0 \leq y \leq 1 \\ 0 & \text{en otro caso} \end{cases}$$

$$\int_0^1 dx \int_0^1 dy f(x, y) = \int_0^1 dx \int_0^1 dy (x + y) = 1$$

$$f(x) = \int_0^1 f(x, y) dy = \int_0^1 (x + y) dy = \left(x + \frac{1}{2}\right); \quad 0 \leq x \leq 1$$

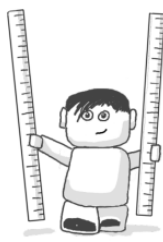
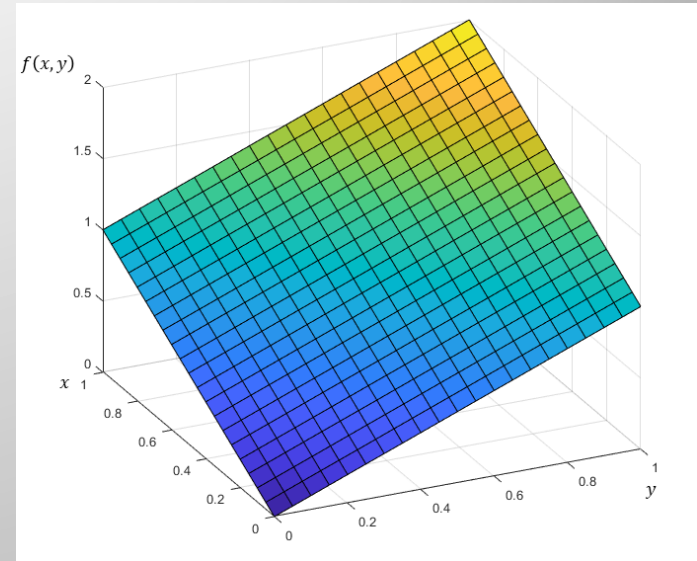
$$f(y) = \int_0^1 f(x, y) dx = \int_0^1 (x + y) dx = \left(y + \frac{1}{2}\right); \quad 0 \leq y \leq 1$$

$$\mathbb{E}\{x\} = \frac{7}{12}$$

$$\mathbb{E}\{y\} = \frac{7}{12}$$

$$\sigma^2(x) = \frac{11}{(12)^2}$$

$$\sigma^2(y) = \frac{11}{(12)^2}$$



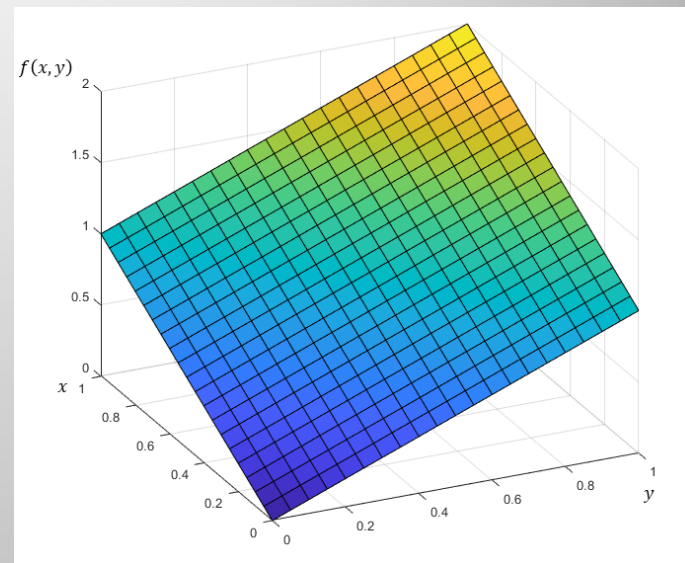
Distribución lineal bivalente

Podremos calcular las probabilidades condicionadas y aplicar el teorema de Bayes:

$$f(x|y) = \frac{x+y}{y+1/2}$$

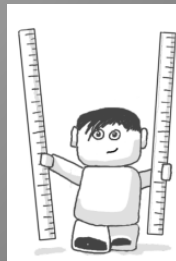
$$f(y|x) = \frac{x+y}{x+1/2}$$

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



Podremos entonces escribir:

$$f(y|x) = \frac{\frac{x+y}{y+1/2} (y+1/2)}{\int_0^1 \frac{x+y}{y+1/2} (y+1/2) dy} = \frac{x+y}{\int_0^1 (x+y) dy} = \frac{x+y}{x+1/2}$$



Distribución lineal bivalente

Podremos calcular la covarianza:

$$\text{cov}(x, y) = \int_0^1 dx \int_0^1 dy \left(x - \frac{7}{12}\right) \left(y - \frac{7}{12}\right) (x + y) = -\frac{1}{(12)^2}$$

Como vemos se trata de variable con covarianza negativa, no son independientes ya que:

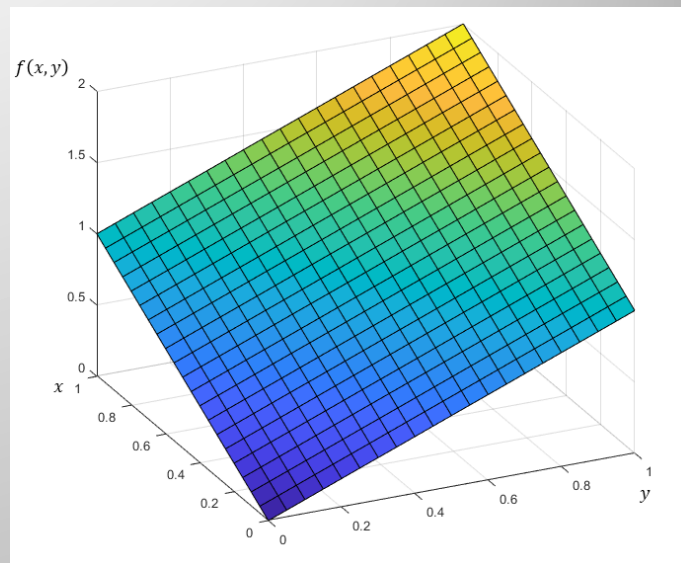
$$f(x, y) \neq f(x)f(y)$$

$$M = \begin{pmatrix} \sigma^2(x) & \text{cov}(x, y) \\ \text{cov}(x, y) & \sigma^2(y) \end{pmatrix} = \frac{1}{(12)^2} \begin{pmatrix} 11 & -1 \\ -1 & 11 \end{pmatrix}$$

Podremos calcular el determinante de la matriz de covarianza y el coeficiente de correlación:

$$\det(M) = \frac{120}{(12)^4} = \frac{5}{3} \frac{1}{(12)^2}$$

$$\rho = -\frac{1}{11}$$

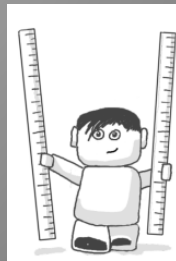


N.B.: Tomando una densidad de probabilidad definida entre -1 y 1:

$$f(x, y) = \begin{cases} \frac{3}{8}(x^2 + y^2); & -1 \leq x \leq 1; -1 \leq y \leq 1 \\ 0 & \text{en otro caso} \end{cases}$$

$$\text{cov}(x, y) = 0$$

Sin embargo las variables aleatorias no son independientes.



Momentos de la distribución bivalente

Definimos el momento de una distribución de probabilidad de órdenes (r, s) respecto al punto (a, b) como el valor esperado:

$$M_{r,s}(a, b) = \mathbb{E}\{(x - a)^r (y - b)^s\}$$

Discreto

$$M_{r,s}(a, b) = \sum_i \sum_j (x_i - a)^r (y_j - b)^s p(x_i, y_j)$$

Continuo

$$M_{r,s}(a, b) = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dy (x - a)^r (y - b)^s f(x, y)$$

Algunas de las propiedades de estos momentos son:

$$M_{0,0}(a, b) = 1$$

$$M_{1,0}(0, b) = \mu = \mathbb{E}\{x\}$$

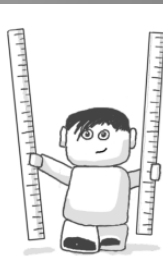
$$M_{0,1}(a, 0) = \nu = \mathbb{E}\{y\}$$

$$M_{2,0}(\mu, b) = \sigma^2(x)$$

$$M_{0,2}(a, \nu) = \sigma^2(y)$$

Incluyendo la covarianza:

$$M_{1,1}(\mu, \nu) = \text{cov}(x, y)$$



Sumando variables aleatorias

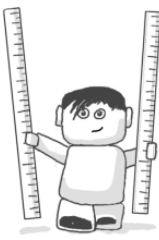
Consideremos dos variables aleatorias X e Y . Si construimos una nueva variable aleatoria mediante su suma $Z = X + Y$, en general las función de probabilidad, media y varianza de Z se modifican respecto a X e Y .

$$E\{z\} = E\{x + y\}$$

$$E\{z\} = \sum_k p(z_k) z_k = \sum_i \sum_j (x_i + y_j) p(x_i, y_j)$$

El valor esperado de la nueva variable Z será:

$$E\{z\} = \sum_i x_i \sum_j p(x_i, y_j) + \sum_j y_j \sum_i p(x_i, y_j) = \sum_i x_i p(x_i) + \sum_j y_j p(y_j) = E\{x\} + E\{y\}$$



Sumando variables aleatorias

Si consideramos ahora la varianza de $Z = X + Y$,

$$\sigma^2(z) = E\{(z - E\{z\})^2\} = E\{(x + y - E\{x + y\})^2\}$$

$$\sigma^2(z) = \sum_i \sum_j (x_i + y_j - E\{x + y\})^2 p(x_i, y_j)$$

Si las variables aleatorias son estadísticamente independientes

$$p(x_i, y_j) = p(x_i)p(y_j)$$

$$E\{x + y\} = E\{x\} + E\{y\}$$

$$\sigma^2(z) = \sum_i \sum_j (x_i + y_j - E\{x\} - E\{y\})^2 p(x_i) p(y_j)$$

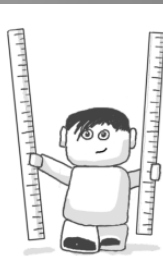
$$\sigma^2(z) = \sum_i \sum_j (x_i - E\{x\})^2 p(x_i) p(y_j) + \sum_i \sum_j (y_j - E\{y\})^2 p(x_i) p(y_j) + 2 \sum_i \sum_j (x_i - E\{x\})(y_j - E\{y\}) p(x_i) p(y_j)$$

$$\sigma^2(x)$$

$$\sigma^2(y)$$

$$\text{cov}(x, y) = 0$$

$$\sigma^2(z) = \sigma^2(x) + \sigma^2(y)$$

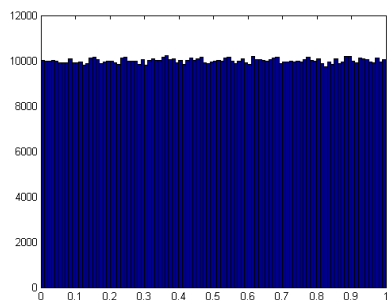


Sumando variables aleatorias

$$f(x_i) = \begin{cases} 1 & 0 \leq x_i \leq 1 \\ 0 & \text{otro caso} \end{cases}$$

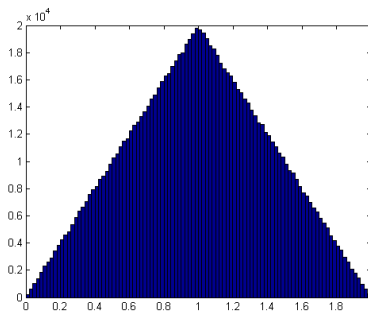
x_1

$$\mu = 0,5 \\ \sigma^2 = 1/12$$



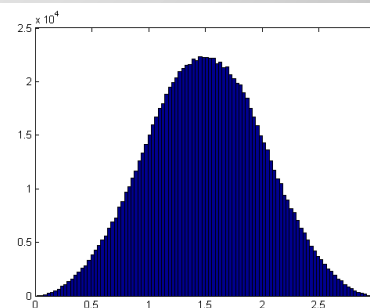
$x_1 + x_2$

$$\mu = 1 \\ \sigma^2 = 1/6$$



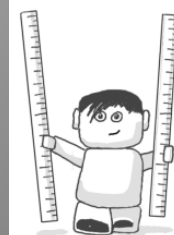
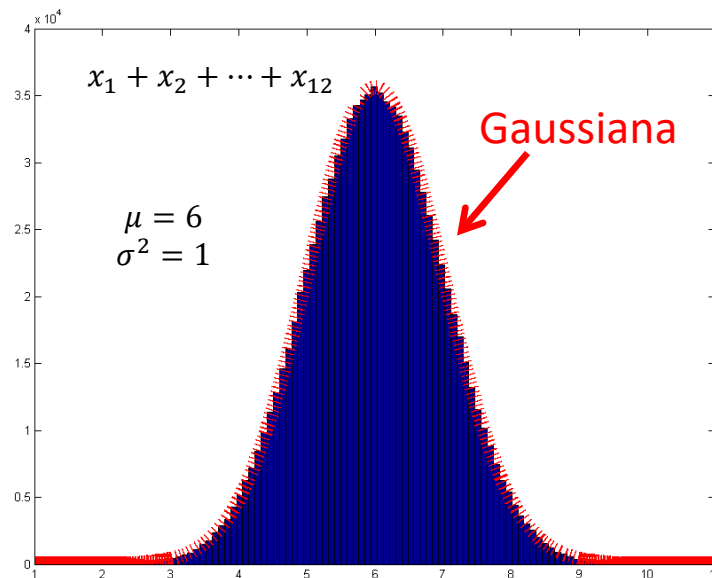
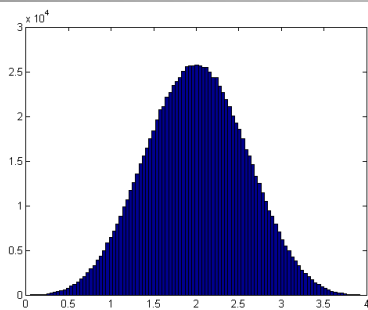
$x_1 + x_2 + x_3$

$$\mu = 3/2 \\ \sigma^2 = 1/4$$



$x_1 + x_2 + x_3 + x_4$

$$\mu = 2 \\ \sigma^2 = 1/3$$



Teorema del límite central

Sean n variables aleatorias $X_1, X_2, X_3, \dots, X_n$ estadísticamente independientes con la misma distribución de probabilidad. Cada una de las variables aleatorias X_i ; $i = 1, \dots, n$ tiene valor esperado μ y varianza σ^2 .

Consideremos la variable aleatoria:

$$\bar{X} = \frac{1}{n}(X_1 + X_2 + \dots + X_n)$$

$$\mathbb{E}\{\bar{X}\} = \mu$$

$$\sigma^2(\bar{X}) = \frac{\sigma^2}{n}$$

Cuando $n \rightarrow \infty$ la distribución de probabilidad de \bar{X} tiende a una distribución normal $N(\mu, \frac{\sigma}{\sqrt{n}})$.

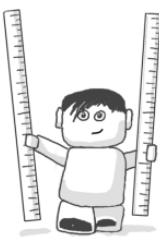
Equivalentemente, considerando la variable aleatoria

$$\bar{Y} = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

$$\mathbb{E}\{\bar{Y}\} = 0$$

$$\sigma^2(\bar{Y}) = 1$$

Cuando $n \rightarrow \infty$ la distribución de probabilidad de \bar{Y} tiende a una distribución normal estándar $N(0, 1)$.



Teorema del límite central

Podemos demostrar el teorema usando la función generatriz de momentos. La idea es demostrar que:

$$\mathcal{M}_{\bar{Y}}(t) \xrightarrow{n \rightarrow \infty} \exp\left(\frac{t^2}{2}\right)$$

Lo cual es equivalente a demostrar que $f(\bar{Y}) \rightarrow N(0, 1)$

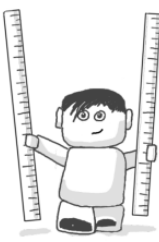
$$\mathcal{M}_{\bar{Y}}(t) = \mathbb{E}\{\exp(t\bar{Y})\} = \mathbb{E}\left\{\exp\left(t \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{X_i - \mu}{\sigma}\right)\right\}$$

$$z_i = \frac{X_i - \mu}{\sigma}$$

$$\mathbb{E}\{z_i\} = 0$$

$$\sigma^2(z_i) = 1$$

$$\mathcal{M}_{\bar{Y}}(t) = \mathbb{E}\{\exp(t\bar{Y})\} = \mathbb{E}\left\{\exp\left(t \frac{1}{\sqrt{n}} \sum_{i=1}^n z_i\right)\right\}$$



Teorema del límite central

Podemos entonces escribir:

$$\mathcal{M}_{\bar{Y}}(t) = \mathbb{E}\{\exp(t\bar{Y})\} = \mathbb{E}\left\{\prod_{i=1}^n \exp\left(t \frac{1}{\sqrt{n}} z_i\right)\right\}$$

Al tratarse de variables estadísticamente independientes, tendremos que:

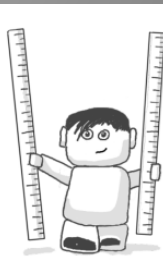
$$\mathcal{M}_{\bar{Y}}(t) = \prod_{i=1}^n \mathbb{E}\left\{\exp\left(t \frac{1}{\sqrt{n}} z_i\right)\right\}$$

$$\mathbb{E}\left\{\exp\left(\frac{t}{\sqrt{n}} z_i\right)\right\} = 1 + \frac{t}{\sqrt{n}} \mathbb{E}\{z_i\} + \frac{1}{2!} \left(\frac{t}{\sqrt{n}}\right)^2 \mathbb{E}\{z_i^2\} + \frac{1}{3!} \left(\frac{t}{\sqrt{n}}\right)^3 \mathbb{E}\{z_i^3\} + \frac{1}{4!} \left(\frac{t}{\sqrt{n}}\right)^4 \mathbb{E}\{z_i^4\} + \dots$$

$$\mathbb{E}\left\{\exp\left(\frac{t}{\sqrt{n}} z_i\right)\right\} = 1 + \frac{t^2}{2n} + \frac{t^3}{6 n^{3/2}} \mathbb{E}\{z_i^3\} + \frac{t^4}{24 n^2} \mathbb{E}\{z_i^4\} + \dots$$

Como son variables idénticamente distribuidas:

$$\mathbb{E}\{z_i^3\} = \mathbb{E}\{z^3\}; \quad \mathbb{E}\{z_i^4\} = \mathbb{E}\{z^4\}; \quad \dots \quad \mathbb{E}\{z_i^n\} = \mathbb{E}\{z^n\}$$



Teorema del límite central

Por lo tanto:

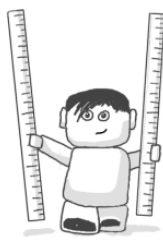
$$\mathcal{M}_{\bar{Y}}(t) = \prod_{i=1}^n \left[1 + \frac{t^2}{2n} + \frac{t^3}{6n^{3/2}} \mathbb{E}\{z^3\} + \frac{t^4}{24n^2} \mathbb{E}\{z^4\} + \dots \right]$$

$$\mathcal{M}_{\bar{Y}}(t) = \left[1 + \frac{1}{n} \left(\frac{t^2}{2} + \frac{t^3}{6n^{1/2}} \mathbb{E}\{z^3\} + \frac{t^4}{24n} \mathbb{E}\{z^4\} + \dots \right) \right]^n$$

Observemos que:

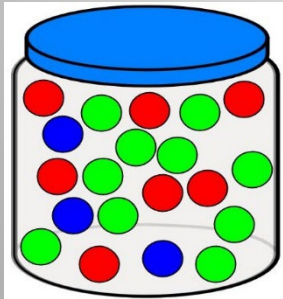
$$\lim_{n \rightarrow \infty} \left(\frac{t^2}{2} + \frac{t^3}{6n^{1/2}} \mathbb{E}\{z^3\} + \frac{t^4}{24n} \mathbb{E}\{z^4\} + \dots \right) = \frac{t^2}{2}$$

$$\lim_{n \rightarrow \infty} \mathcal{M}_{\bar{Y}}(t) = \lim_{n \rightarrow \infty} \left[1 + \frac{1}{n} \left(\frac{t^2}{2} + \frac{t^3}{6n^{1/2}} \mathbb{E}\{z^3\} + \frac{t^4}{24n} \mathbb{E}\{z^4\} + \dots \right) \right]^n = \exp\left(\frac{t^2}{2}\right)$$



Distribución multinomial

Consideremos un proceso aleatorio que pueda arrojar tres posibles resultados mutuamente excluyentes (A,B,C).



Si realizamos n veces el experimento aleatorio, obtendremos n_1 veces el resultado A, n_2 veces el resultado B y $n - n_1 - n_2$ veces el resultado C.

La probabilidad individual asociada a obtener el resultado A será p_1 , para el resultado B será p_2 y $p_3 = 1 - p_1 - p_2$ para C.

Un resultado específico $n_1, n_2, n - n_1 - n_2$ tiene probabilidad

$$p_1^{n_1} p_2^{n_2} (1 - p_1 - p_2)^{n - n_1 - n_2}$$

Pero hay un conjunto de combinaciones equiprobables que dan el mismo resultado

$$\frac{n!}{n_1! n_2! (n - n_1 - n_2)!}$$

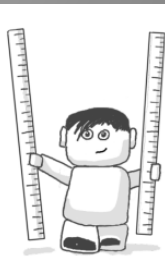
Por lo que:

$$p(x_1 = n_1; x_2 = n_2; x_3 = n_3) = \frac{n!}{n_1! n_2! n_3!} (p_1^{n_1} p_2^{n_2} p_3^{n_3})$$

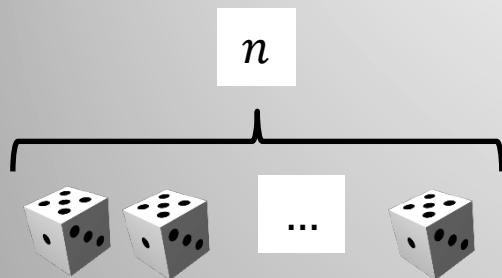
$$n_1, n_2 = 0, 1, 2, \dots, n$$

$$n_3 = n - n_1 - n_2$$

$$p_3 = 1 - p_1 - p_2$$



Distribución multinomial



Supongamos un proceso aleatorio cuyo resultado puede ser asignado a k clases distintas. Si realizamos n experimentos aleatorios tendremos

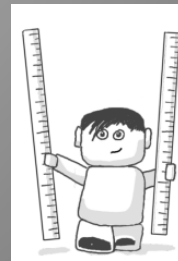
$$n = n_1 + n_2 + \dots + n_k$$

La probabilidad de que un suceso individual pertenezca a la clase j -ésima será p_j

$$p_1 + p_2 + \dots + p_k = 1$$

La probabilidad asociada a obtener (n_1, n_2, \dots, n_k)

$$p(x_1 = n_1; x_2 = n_2; \dots; x_k = n_k) = \frac{n!}{n_1! n_2! \dots n_k!} (p_1^{n_1} p_2^{n_2} \dots p_k^{n_k})$$



Distribución multinomial

Las distribuciones marginales de la distribución binomial pueden obtenerse a partir de la probabilidad multinomial como

$$p(x_1 = n_1) = \frac{p_1^{n_1} n!}{n_1!} \sum_{n_2=0}^{n-n_1} \sum_{n_3=0}^{n-n_1-n_2} \dots \sum_{n_{k-1}=0}^{n-n_1-n_2-\dots-n_{k-2}} \frac{p_2^{n_2} p_3^{n_3} \dots p_{k-1}^{n_{k-1}} p_k^{n-n_1-\dots-n_{k-1}}}{n_2! n_3! \dots n_{k-1}! (n - n_1 - \dots - n_{k-1})!}$$

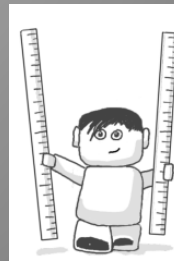
Usando la fórmula del binomio de Newton reiteradamente podemos obtener

$$p(x_1 = n_1) = \frac{n!}{n_1! (n - n_1)!} p_1^{n_1} (1 - p_1)^{n-n_1}$$

Que coincide con la distribución de probabilidad binomial, y por tanto las medias y varianzas marginales son:

$$\mu(x_i) = E\{x_i\} = n p_i$$

$$\sigma^2(x_i) = E\{(x_i - \mu(x_i))^2\} = n p_i (1 - p_i)$$



Distribución multinomial

Para el cálculo de las covarianzas entre los valores de las variables x_1 y x_2 (extensible a cualquier par) observemos que si en el cálculo anterior no sumamos sobre el índice n_2

$$p(x_1 = n_1; x_2 = n_2) = p_1^{n_1} p_2^{n_2} (1 - p_1 - p_2)^{n-n_1-n_2} \frac{n!}{n_1! n_2! (n - n_1 - n_2)!}$$

La covarianza de estas variables puede escribirse como

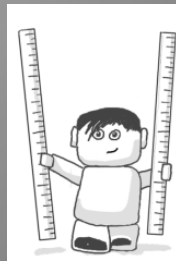
$$\text{cov}(x_1, x_2) = E\{(x_1 - \mu_1)(x_2 - \mu_2)\} = E\{x_1 x_2\} - \mu_1 \mu_2$$

Podemos calcular $E\{x_1 x_2\}$ mediante

$$E\{x_1 x_2\} = \sum_{n_1=0}^{n-1} \sum_{n_2=0}^{n-n_1} n_1 n_2 p_1^{n_1} p_2^{n_2} (1 - p_1 - p_2)^{n-n_1-n_2} \frac{n!}{n_1! n_2! (n - n_1 - n_2)!} = n(n-1)p_1 p_2$$

$$\text{cov}(x_1, x_2) = n(n-1)p_1 p_2 - n^2 p_1 p_2 = -n p_1 p_2$$

Las variables x_1, x_2 no son por tanto estadísticamente independientes.



Distribución multinomial

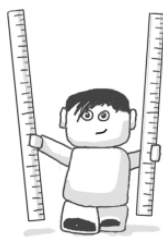
Se realiza una encuesta entre 100 personas sobre su preferencia sobre tres partidos políticos A, B y C, obteniéndose los resultados 17, 48 y 35 respectivamente. Si no existen otras opciones políticas y se considera que la preferencia entre estas opciones es equiprobable, ¿Cuál es la probabilidad de este resultado?

$$p(x_1 = n_1; x_2 = n_2) = p_1^{n_1} p_2^{n_2} (1 - p_1 - p_2)^{n - n_1 - n_2} \frac{n!}{n_1! n_2! (n - n_1 - n_2)!}$$

$$p(x_1 = 17; x_2 = 48) = (1/3)^{17} (1/3)^{48} (1/3)^{35} \frac{100!}{17! 48! 35!}$$

$$p(x_1 = 17; x_2 = 48) = 3.97 \times 10^{-6}$$

Bajo estas hipótesis el resultado es poco probable (i.e. el resultado 33,33,34 bajo estas hipótesis tiene una probabilidad de 8.1×10^{-3})



Distribución normal bidimensional

Si consideramos dos variables aleatorias x , y con distribución gaussiana de la forma:

$$f(x) = \frac{1}{\sqrt{2\pi} \sigma_x} \exp \left[-\frac{1}{2} \left(\frac{(x - \mu_x)^2}{\sigma_x^2} \right) \right]$$

$$f(y) = \frac{1}{\sqrt{2\pi} \sigma_y} \exp \left[-\frac{1}{2} \left(\frac{(y - \mu_y)^2}{\sigma_y^2} \right) \right]$$

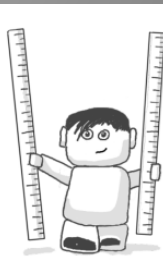
Podremos construir una densidad de probabilidad bivalente normal:

$$f(x, y) = \frac{1}{2\pi \sigma_x \sigma_y} \exp \left[-\frac{1}{2} \left(\frac{(x - \mu_x)^2}{\sigma_x^2} \right) - \frac{1}{2} \left(\frac{(y - \mu_y)^2}{\sigma_y^2} \right) \right] = f(x) f(y)$$

Esta distribución de probabilidad representa a dos variables con distribuciones marginales gaussianas y estadísticamente independientes que verifican:

$$\text{cov}(x, y) = 0$$

¿Es posible una distribución bivalente con distribuciones marginales gaussianas con covarianza no nula?



Rotación del espacio bivariente

Si consideramos dos variables aleatorias x, y tales que:

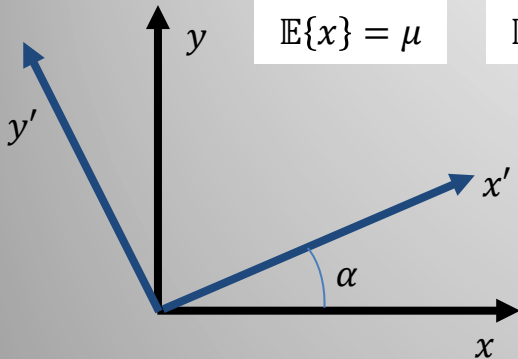
$$\mathbb{E}\{x\} = \mu$$

$$\mathbb{E}\{y\} = \nu$$

$$\mathbb{E}\{(x - \mu)^2\} = \sigma^2(x)$$

$$\mathbb{E}\{(y - \nu)^2\} = \sigma^2(y)$$

$$\text{cov}(x, y) = 0$$



Realizando una rotación del espacio bivariente:

$$\begin{aligned} x &= \cos \alpha \, x' - \sin \alpha \, y' \\ y &= \sin \alpha \, x' + \cos \alpha \, y' \end{aligned}$$

$$\begin{aligned} x' &= \cos \alpha \, x + \sin \alpha \, y \\ y' &= -\sin \alpha \, x + \cos \alpha \, y \end{aligned}$$

$$\mathbb{E}\{x'\} = \mu' = \cos \alpha \, \mu + \sin \alpha \, \nu$$

$$\mathbb{E}\{y'\} = \nu' = -\sin \alpha \, \mu + \cos \alpha \, \nu$$

$$\sigma^2(x') = \cos^2 \alpha \, \sigma^2(x) + \sin^2 \alpha \, \sigma^2(y)$$

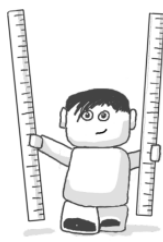
$$\sigma^2(y') = \sin^2 \alpha \, \sigma^2(x) + \cos^2 \alpha \, \sigma^2(y)$$

Podemos calcular la covarianza de las nuevas variables:

$$\text{cov}(x', y') = \mathbb{E}\{x' y'\} - \mathbb{E}\{x'\} \mathbb{E}\{y'\}$$

Desarrollando esta expresión obtenemos:

$$\text{cov}(x', y') = -\cos \alpha \, \sin \alpha \, [\sigma^2(x) - \sigma^2(y)]$$



Rotación del espacio bivariente

A su vez es posible demostrar que:

$$\sigma^2(x') - \sigma^2(y') = \cos 2\alpha \{ \sigma^2(x) - \sigma^2(y) \}$$

Por lo tanto se verifica que:

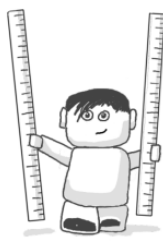
$$\text{cov}(x', y') = -\frac{1}{2} \tan 2\alpha \{ \sigma^2(x') - \sigma^2(y') \}$$

Esta expresión nos permite escribir el coeficiente de correlación de las nuevas variables:

$$\rho = \frac{\text{cov}(x', y')}{\sigma(x') \sigma(y')} = -\frac{1}{2} \tan 2\alpha \frac{\sigma^2(x') - \sigma^2(y')}{\sigma(x') \sigma(y')}$$

Utilizando las expresiones anteriores (o bien mediante las propiedades algebraicas de las transformaciones ortogonales) podemos encontrar que:

$$\sigma^2(x) \sigma^2(y) = (1 - \rho^2) \sigma^2(x') \sigma^2(y')$$



Rotación del espacio bivariente

La distribución normal bivariente se transforma mediante:

$$f(x, y) = f(x', y') = \frac{1}{2\pi \sigma(x) \sigma(y)} \exp \left[-\frac{1}{2} \left(\frac{(x - \mu)^2}{\sigma^2(x)} \right) - \frac{1}{2} \left(\frac{(y - v)^2}{\sigma^2(y)} \right) \right]$$

Que podemos reescribir:

$$f(x', y') = \frac{1}{2\pi \sigma(x') \sigma(y') \sqrt{1 - \rho^2}} \exp \left\{ -\frac{[\cos \alpha (x' - \mu') - \sin \alpha (y' - v')]^2}{2 \sigma^2(x)} - \frac{[\sin \alpha (x' - \mu') + \cos \alpha (y' - v')]^2}{2 \sigma^2(y)} \right\}$$

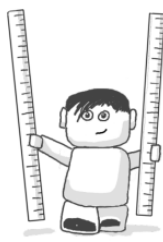
$$f(x', y') = \frac{1}{2\pi \sigma(x') \sigma(y') \sqrt{1 - \rho^2}} \exp \left\{ -\frac{1}{2} \left(\frac{\cos^2 \alpha}{\sigma^2(x)} + \frac{\sin^2 \alpha}{\sigma^2(y)} \right) (x' - \mu')^2 - \frac{1}{2} \left(\frac{\sin^2 \alpha}{\sigma^2(x)} + \frac{\cos^2 \alpha}{\sigma^2(y)} \right) (y' - v')^2 - \frac{1}{2} (2 \cos \alpha \sin \alpha) (x' - \mu') (y' - v') \left(\frac{1}{\sigma^2(y)} - \frac{1}{\sigma^2(x)} \right) \right\}$$

Observemos que:

$$\left(\frac{\cos^2 \alpha}{\sigma^2(x)} + \frac{\sin^2 \alpha}{\sigma^2(y)} \right) = \frac{1}{\sigma^2(x')(1 - \rho^2)}$$

$$\left(\frac{\sin^2 \alpha}{\sigma^2(x)} + \frac{\cos^2 \alpha}{\sigma^2(y)} \right) = \frac{1}{\sigma^2(y')(1 - \rho^2)}$$

$$(2 \cos \alpha \sin \alpha) \left(\frac{1}{\sigma^2(y)} - \frac{1}{\sigma^2(x)} \right) = -\frac{2\rho}{1 - \rho^2} \frac{1}{\sigma(x') \sigma(y')}$$



Distribución normal bidimensional

Podemos extender la distribución de gauss al caso multidimensional. Consideremos dos variables aleatorias x, y

$$f(x, y) = \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1 - \rho^2}} \exp \left[-\frac{1}{2(1 - \rho^2)} \left(\frac{(x - \mu_x)^2}{\sigma_x^2} - \frac{2\rho(x - \mu_x)(y - \mu_y)}{\sigma_x \sigma_y} + \frac{(y - \mu_y)^2}{\sigma_y^2} \right) \right]$$

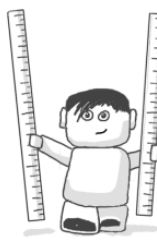
Siendo

$$\rho = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y}$$

Las distribuciones marginales son gaussianas univariantes

$$f(x) = \int_{-\infty}^{+\infty} dy f(x, y) = \frac{1}{\sqrt{2\pi} \sigma_x} \exp \left[-\frac{(x - \mu_x)^2}{2 \sigma_x^2} \right]$$

$$f(y) = \int_{-\infty}^{+\infty} dx f(x, y) = \frac{1}{\sqrt{2\pi} \sigma_y} \exp \left[-\frac{(y - \mu_y)^2}{2 \sigma_y^2} \right]$$



Distribución normal bidimensional

Otra forma de expresar la distribución normal bivalente que se extiende de modo general al caso multidimensional es mediante notación matricial.

La matriz M de covarianza y su inversa vienen dadas por:

$$M = \begin{pmatrix} \sigma_x^2 & \text{cov}(x, y) \\ \text{cov}(x, y) & \sigma_y^2 \end{pmatrix} = \begin{pmatrix} \sigma_x^2 & \sigma_x \sigma_y \rho \\ \sigma_x \sigma_y \rho & \sigma_y^2 \end{pmatrix}$$

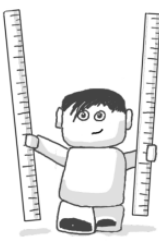
$$M^{-1} = \frac{1}{1 - \rho^2} \begin{pmatrix} 1/\sigma_x^2 & -\rho/\sigma_x \sigma_y \\ -\rho/\sigma_x \sigma_y & 1/\sigma_y^2 \end{pmatrix}$$

Usando notación vectorial:

$$(\vec{x} - \vec{\mu}) = \begin{pmatrix} x - \mu_x \\ y - \mu_y \end{pmatrix}$$

Podemos escribir la densidad de probabilidad normal bidimensional como:

$$f(x, y) = \frac{1}{(\sqrt{2\pi})^2 \sqrt{\det(M)}} \exp \left[-\frac{1}{2} (\vec{x} - \vec{\mu})^T M^{-1} (\vec{x} - \vec{\mu}) \right]$$



Distribución normal bidimensional

Podemos establecer en la gaussiana bidimensional los contornos de equiprobabilidad que vienen dados por la condición

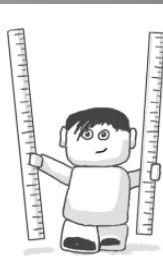
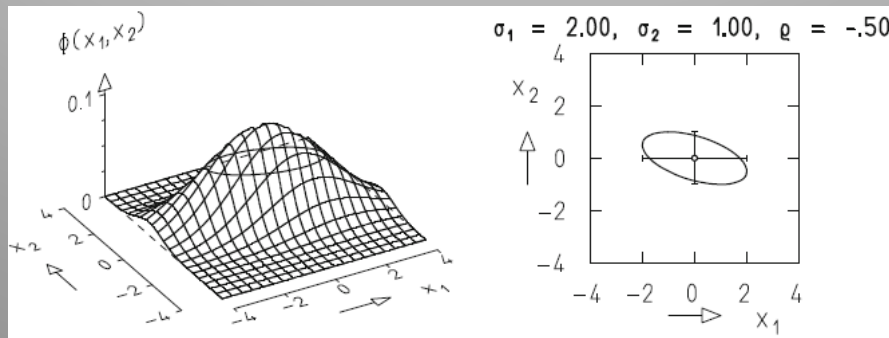
$$-\frac{1}{2(1-\rho^2)} \left(\frac{(x-\mu_x)^2}{\sigma_x^2} - \frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} + \frac{(y-\mu_y)^2}{\sigma_y^2} \right) = -\frac{1}{2} \Delta$$

$$\Delta > 0$$

Cuyo resultado serán cónicas en el plano. En particular para $\Delta=1$

$$\left(\frac{(x-\mu_x)^2}{\sigma_x^2} - \frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} + \frac{(y-\mu_y)^2}{\sigma_y^2} \right) = (1-\rho^2)$$

En general esto corresponde a una elipse que se conoce como *elipse de covarianza*



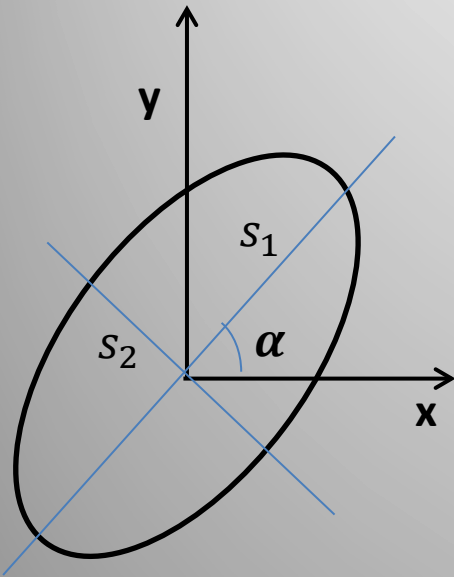
Distribución normal bidimensional

Los semiejes de la elipse s_1, s_2 y el ángulo α que forman con los ejes x, y vienen dados por

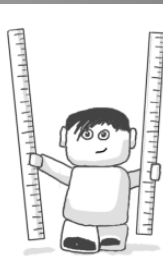
$$\tan(2\alpha) = \frac{2\rho\sigma_x\sigma_y}{\sigma_x^2 - \sigma_y^2} = \frac{2 \operatorname{cov}(x, y)}{\sigma_x^2 - \sigma_y^2}$$

$$s_1^2 = \frac{(1 - \rho^2)\sigma_x^2\sigma_y^2}{\sigma_y^2\cos^2(\alpha) - 2\rho\sigma_x\sigma_y\sin(\alpha)\cos(\alpha) + \sigma_x^2\sin^2(\alpha)}$$

$$s_2^2 = \frac{(1 - \rho^2)\sigma_x^2\sigma_y^2}{\sigma_y^2\sin^2(\alpha) + 2\rho\sigma_x\sigma_y\sin(\alpha)\cos(\alpha) + \sigma_x^2\cos^2(\alpha)}$$

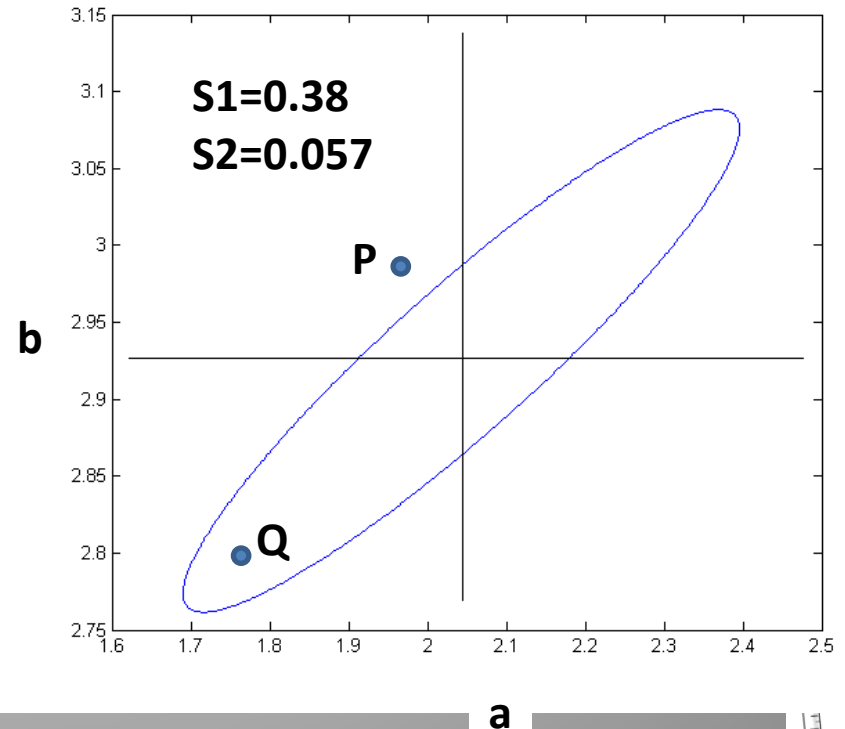
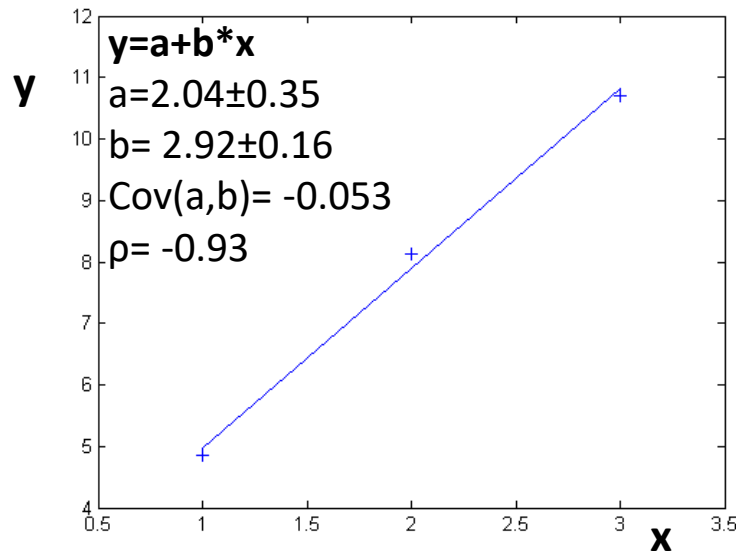


La probabilidad de que la variable (x, y) esté dentro de la elipse de covarianza es $1 - \frac{1}{\sqrt{e}} = 0.393$

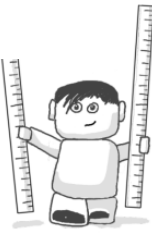


Ajuste a una recta

Consideremos los datos $x=(1,2,3)$ $y=(4.85,8.12,10.70)$ siendo la incertidumbre de y igual a $u_y=0.23$. Se nos pide realizar un ajuste lineal $y=a+b*x$ y evaluar la incertidumbre de a y b .



Una curva de isoprobabilidad del punto P es menos probable que la de Q aunque Q diste más de los valores medios.



Bibliografía:

Capítulo 6. “Probabilidad y estadística” George C. Canavos, Ed. Mc Graw-Hill

Capítulo 6. “Fundamentos de estadística” Daniel Peña, Alianza Editorial

Capítulo 2. “Tratamiento de datos físicos” Faustino Gómez, Luis M Varela, USC

**Capítulo 5. “Statistical and Computational Methods for Scientists and Engineers”
Siegmund Brandt, Springer**

