MATEMATICAS

A. Matemáticas básicas

I. Cálculo

1. Diferencial en una variable

• Derivada de una función de una variable

Definición.- Sea f(x) una función continua en el punto $x=a$. La derivada de $f(x)$, respecto a x, en el punto $x=a$, que representaremos con el símbolo D$f(a)$, es el

$$\lim_{\Delta x \to 0} \frac{f(a+\Delta x)-f(a)}{\Delta x}$$

Definición.- Sea f una función definida en un intervalo abierto que contiene a *a.* Entonces la pendiente *m* de la recta tangente a la gráfica de *f* en el punto *P(a,f(a))* está dada por

$$m = \lim_{h \to 0} \frac{f(a+h)-f(a)}{h}$$

siempre y cuando este limite exista.

Definición.- Sea f una función definida en un intervalo abierto que contiene a *a*. Entonces la derivada de *f* en α, denotada por $f'(a)$, está dada por

$$f'(a) = \lim_{h \to o} \frac{f(a+h)-f(a)}{h}$$

Reglas para encontrar derivadas

1. $\frac{d}{dx}(c) = 0$

2. $\frac{d}{dx}(x) = 1$

3. $\frac{d}{dx}(u+v+...) = \frac{d}{dx}(u) + \frac{d}{dx}(v) + ...$

4. $\frac{d}{dx}(cu) = c\frac{d}{dx}(u)$

5. $\frac{d}{dx}(uv) = u\frac{d}{dx}(v) + v\frac{d}{dx}(u)$

6. $\frac{d}{dx}(uvw) = uv\frac{d}{dx}(w) + uw\frac{d}{dx}(v) + vw\frac{d}{dx}(u)$

7. $\frac{d}{dx}\left[\frac{u}{c}\right] = \frac{1 \bullet d}{c\, dx}(u)$, $c \neq 0$

8. $\frac{d}{dx}\left(\frac{c}{u}\right) = c\,\frac{d}{dx}\left(\frac{1}{u}\right) = -\frac{c}{u^2} \bullet \frac{d}{dx}(u)$, $u \neq 0$

9. $\frac{d}{dx}\left(\frac{u}{v}\right) = \frac{v\frac{d}{dx}(u) - u\frac{d}{dx}(v)}{v^2}$, $v \neq 0$

10. $\frac{d}{dx}(x^m) = mx^{m-1}$

11. $\frac{d}{dx}(u^m) = mu^{m-1}\frac{d}{dx}(u)$

Ejemplos:
Encuentre la pendiente de la recta tangente a la gráfica de *f* en el punto *P(a, f(a))*.

1. $f(x)=2-x^3$

$=\frac{d}{dx}(2) - \frac{d}{dx}(x^3)$

$= 0 - 3x^2 = -3x^2$

2. $f(x)=3x-5$

$=3\frac{d}{dx}(x) - \frac{d}{dx}(5)$

$= 3(1) – 0 = 3$

3. $f(x)= \sqrt{x} +1$

$= \frac{1}{2}(x)^{-1/2}\frac{d}{dx}(x) + \frac{d}{dx}(1)$

$= \frac{1}{2}(x)^{-1/2}$

$= \frac{1}{2\sqrt{x}}$

4. $f(x)= \frac{1}{x} - 1$

$$= \frac{d}{dx}(1/x) - \frac{d}{dx}(1)$$

$$= \frac{x \frac{d}{dx}(1) - (1)\frac{d}{dx}(x)}{x^2} - 0$$

$$= -\frac{1}{x^2}$$

La posición de un punto *P* moviéndose sobre una recta coordenada *l* está dada por *f(t)* donde t está medido en segundos y *f(t)* en centímetros.
Encuentre la velocidad media de *P* en el siguiente intervalo de tiempo:

$f(t)= 4t^2 + 3t$ en el intervalo (1,1.2)

$$= 4 \frac{d}{dx}(t^2) + 3 \frac{d}{dx}(t)$$

$$= 4(2t) + 3(1)$$
$$= 8t + 3$$

Sumando los puntos del intervalo y dividiendo entre 2

$$\frac{1+1.2}{2} = 1.1$$

Sustituyendo en la derivada resultante

$$=8(1.1) + 3$$
$$=11.8$$

Hallar *f'(x)*

$f(x)=\sqrt{3x+1} = (3x+1)^{1/2}$

$$f'(x)= \frac{1}{2}(3x+1)^{-1/2} \frac{d}{dx}(3x+1)$$

$$= \frac{1}{2}(3x+1)^{-1/2}(3)$$

$$= \frac{3}{2\sqrt{3x+1}}$$

Derivar las siguientes funciones:

1. $g(w)=\dfrac{1}{w^4}$

$$=\frac{w^4\,(0)-1(4w^3)}{(w^4)^2}$$

$$=\frac{-4w^3}{w^8} = -4w^{3-8} = -4w^{-5}$$

2. $g(x) = (x^3-7)(2x^2+3)$

$$= (x^3-7)\frac{d}{dx}(2x^2+3)+(2x^2+3)\frac{d}{dx}(x^3-7)$$

$$= (x^3-7)(4x)+(2x^2+3)(3x^2)$$

$$= 4x^4-28x+6x^4+9x^2$$

$$= 10x^4+9x^2-28x$$

3. $f(x)=\dfrac{4x-5}{3x+2}$

$$=\frac{(3x+2)\frac{d}{dx}(4x-5)-(4x-5)\frac{d}{dx}(3x+2)}{(3x+2)^2}$$

$$=\frac{(3x+2)(4)-(4x-5)(3)}{(3x+2)^2}$$

$$=\frac{12x+8-12x+15}{(3x+2)^2} = \frac{23}{(3x+2)^2}$$

4. $h(z) = \dfrac{8-z+3z^2}{2-9z}$

$$=\frac{(2-9z)\frac{d}{dx}(8-z+3z^2)-(8-z+3z^2)\frac{d}{dx}(2-9z)}{(2-9z)^2}$$

$$=\frac{(2-9z)(-1+6z)-(8-z+3z^2)(-9)}{(2-9z)^2}$$

$$=\frac{(-2+12z+9z-54z^2)-(-72+9z-27z^2)}{(2-9z)^2}$$

$$= \frac{-2+12z+9z-54z^2+72-9z+27z^2}{(2-9z)^2}$$

$$= \frac{-27z^2+12z+70}{(2-9z)^2}$$

Usar $\Delta y = f(x_2)-f(x_1) = f(x_1+\Delta x)-f(x_1)$ para encontrar Δy usando los valores iniciales de x y Δx indicados.

1. $y = 2x^2-4x+5$, $x=2$, $\Delta x=-0.2$

$= f(1.8)-f(2)$
$= \{2(1.8)^2-4(1.8)+5\} – \{2(2)^2-4(2)+5\}$
$= (6.48-7.2+5)-(8-8+5)$
$= 4.28-5$
$= -0.72$

2. $y = 1/x^2$, $x=3$, $\Delta x=0.3$

$= f(3.3)-f(3)$

$$= \left\{\frac{1}{(3.3)^2}\right\} - \left\{\frac{1}{(3)^2}\right\}$$

$$= \frac{1}{10.89} - \frac{1}{9}$$

$= -0.0192837$

Derivar las funciones definidas.

1. $f(x) = (x^2-3x+8)^3$

$$= 3(x^2-3x+8)^2 \frac{d}{dx}(x^2-3x+8)$$

$$= 3(x^2-3x+8)^2(2x-3)$$

2. $g(x) = (8x-7)^{-5}$

$$= -5(8x-7)^{-6} \frac{d}{dx}(8x-7)$$

$$= -5(8x-7)^{-6}\,(8)$$

$$= -40(8x-7)^{-6}$$

3. $f(x) = \dfrac{x}{(x^2-1)^4}$

$$= \frac{(x^2-1)^4 \dfrac{d}{dx}(x) - x \dfrac{d}{dx}(x^2-1)^4}{((x^2-1)^4)^2}$$

$$= \frac{((x^2-1)^4\,(1)) - ((x)\,((4\,(x^2-1)^3\,(2x)))}{((x^2-1)^4)^2}$$

$$= \frac{((x^2-1)^4\,(1)) - (x\,(8x\,(x^2-1)^3\,)}{((x^2-1)^4)^2}$$

$$= \frac{(x^2-1)^4 - 8x^2\,(x^2-1)^3\,)}{(x^2-1)^8}$$

$$= \frac{-7x^2(x^2-1)^3}{(x^2-1)^8}$$

4. $g(z) = \left(z^2 - \dfrac{1}{z^2}\right)^6$

$$= 6\left(z^2 - \frac{1}{z^2}\right)^5 \frac{d}{dx}\left(z^2 - \frac{1}{z^2}\right)$$

$$= 6\left(z^2 - \frac{1}{z^2}\right)^5 \left(2z - \frac{((z^2)(0)-(1)(2z))}{(z^2)^2}\right)$$

$$= 6\left(z^2 - \frac{1}{z^2}\right)^5 \left(2z + \frac{2z}{z^4}\right)$$

$$= 6\left(z^2 - \frac{1}{z^2}\right)^5 \left(2z + \frac{2}{z^3}\right)$$

4. $K(u) = \dfrac{(u^2+1)^3}{(4u-5)^5}$

$$= \left(u^2+1\right)^3 \left(-\frac{20}{(4u-5)^6}\right) + (4u-5)^{-5}\,(6u(u^2+1)^2)$$

$$= \frac{-20(u^2+1)^3}{(4u-5)^6} + 6u\,(4u-5)^{-5}\;(u^2+1)^2$$

$$= \frac{-20(u^2+1)^3 + 6u(4u-5)^{-5}\,(u^2+1)^2\,(4u-5)^6}{(4u-5)^6}$$

$$= \frac{-20(u^2+1)^3 + 6u(4u-5)\,(u^2+1)^2}{(4u-5)^6}$$

$$= \frac{-20(u^2+1)^3}{(4u-5)^6} + \frac{6u\,(4u-5)(u^2+1)^2}{(4u-5)^6}$$

$$= \frac{-20(u^2+1)^3}{(4u-5)^6} + \frac{6u\,(u^2+1)^2}{(4u-5)^6}$$

$$= \frac{-20(u^2+1)^3 + 6u\,(u^2+1)^2\;(4u-5)}{(4u-5)^6}$$

$$= \frac{(u^2+1)^2[-20(u^2+1)+6u\,(4u-5)\,]}{(4u-5)^6}$$

$$= \frac{(u^2+1)^2\,[-20u^2-20+24u^2-30u]}{(4u-5)^6}$$

$$= \frac{(u^2+1)^2\,[4u^2-30u-20]}{(4u-5)^6}$$

Encontrar al menos una función implícita *f* determinada por la ecuación dada.

$3x-2y+4 = 2x^2+3y-7x$

$-2y-3y = 2x^2-7x-3x-4$
$-5y = 2x^2 – 10x –4$

$$y= \frac{2x^2}{-5} - \frac{10x}{-5} - \frac{4}{-5}$$

$$y= \frac{2x^2}{-5} + 2x + \frac{4}{-5}$$

Derive la función definida

1. $f(x) = {}^3\sqrt{x^2} + 4\sqrt{x^3}$

$= x^{2/3} + 4x^{3/2}$

$$– \left(\frac{2}{3}\right) X^{-1/3} + \left(\frac{3}{2}\right) (4x)^{1/2}$$

$$– \left(\frac{2}{3}\right) X^{-1/3} + \frac{12x^{1/2}}{2}$$

$$– \left(\frac{2}{3}\right) X^{-1/3} + 6x^{1/2}$$

2. $k(r) = \sqrt[3]{8r^3 + 27}$

$$= (8r^3 + 27)^{1/3} \bullet 3(8r^2)$$

$$= \frac{1}{3} (8r^3 + 27)^{-2/3} \bullet 24r^2$$

$$= \frac{24}{3} r^2 (8r^3 + 27)^{-2/3}$$

$$= 8r^2 (8r^3 + 27)^{-2/3}$$

- **Máximos y mínimos de una función de una variable**

Si la función es creciente a la izquierda del punto, y de creciente a la derecha, lo llamaremos máximo. Si a la izquierda del punto la función es decreciente y a la derecha creciente, diremos que se trata de un mínimo.

Definición.- Sea $f(x)$ con primera derivada y segunda derivada alrededor de un punto x_0. Diremos que hay un máximo en x_0, si $f'(x_0) = 0$ y si $f''(x_0) < 0$.

Definición.- Sea $f(x)$ con primera y segunda derivada alrededor de un punto x_0, se dice que en x_0 hay un mínimo si $f'(x) = 0$ y si $f''(x) > 0$.

Ejemplos:

1. $f(x) = 3x^3 + 2x – 1$

   $f'(x) = 9x^2 + 2$

   $f''(x) = 18x$

- **Problemas que requieren el concepto de la diferencial.**

Definición.- Sea $y = f(x)$ donde f es derivable y sea Δx un incremento de x. Entonces

(i) la diferencial dy de la variable dependiente y está dada por $dy = f'(x)\, \Delta x$.

(ii) la diferencial dx de la variable independiente x está dada por $dx = \Delta x$.

Definición.- Sea $w = f(x, y)$. Las diferenciales dx y dy de las variables independientes x y y se definen como

$$dx = \Delta x \text{ y } dy = \Delta y,$$

donde Δx y Δy son incrementos de x y y. La diferencial dw de la variable dependiente w se define por medio de

$$dw = f_x(x, y)\, dx + f_y(x, y)\, dy = \frac{\partial w}{\partial x} dx + \frac{\partial w}{\partial y} dy$$

Definición.- Sea w = f(x, y). Decimos que f es diferenciable o que tiene una diferencial en (x0, y0) si Δw se puede expresar en la forma

$$\Delta w = f_x(x_0, y_0)\, \Delta x + f_y(x_0, y_0)\, \Delta y + \eta_1 \Delta x + \eta_2 \Delta y$$

donde η_1 y η_2 tienden a cero cuando $(\Delta x, \Delta y) \rightarrow (0, 0)$.

Use diferenciales para estimar el cambio en $f(x, y, z) = x^2z^3 - 3yz^2 + x^{-3} + 2y^{1/2}z$ cuando (x, y, z) cambia de (1, 4, 2) a (1.02, 3.97, 1.96).

$= (2xz^3 + (-3x^{-4}))\ dx + (-3z^2 + y^{-1/2}z)\ dy + (x^2 3z^2 + (-3y2z) + 2y^{1/2})\ dz$

$x_1 = 1$	$y_1 = 4$	$z_1 = 2$
$x_2 = 1.02$	$y_2 = 3.97$	$z_2 = 1.96$
$\Delta x = 1.02 - 1$	$\Delta y = 3.97 - 4$	$\Delta z = 1.96 - 2$
$dx = \Delta x = 0.02$	$dy = \Delta y = -0.03$	$dz = \Delta z = -0.04$

$= (2(1)(2)^3+(-3(1)^{-4}))\ (.02) + (-3(2)^2 + 4^{-1/2}\ (2))\ (-0.03) + ((1)^2\ 3(2)^2 + (-3\ (4)(2(2)) + 2\ (4)^{1/2})\ (-0.04)$

$= (2(8)+ (-3))\ (0.02) + (-3(4) + (.5)(2))\ (-0.03) + ((1)(12)+(-12)(4)+2(2))\ (-0.04)$

$= (16-3)(0.02)+ (-12+1)\ (-0.03) + (12+(-48)+4)\ (-0.04)$

$= (13)(0.02)+(-11)(-0.03) + (-32)\ (-0.04)$

$= 0.26 + 0.33 + 1.28$

$= 1.87$

Use diferenciales para estimar el cambio en $f(x, y) = x^2 - 3x^3y^2 + 4x - 2y^3 + 6$ cuando (x, y) cambia de (-2, 3) a (-2.02, 3.01).

$= (2x + 9x^2y^2 +4)\ dx + (-3x^3\ 2y - 6y^2)\ dy$

$= (2x + 9x^2y^2 +4)\ dx - (6x^3y - 6y^2)\ dy$

$x_1 = -2$	$y_1 = 3$
$x_2 = -2.02$	$y_2 = 3.01$
$\Delta x = -2.02 - (-2)$	$\Delta y = 3.01-3$
$\Delta x = -2.02 +2$	$\Delta y = 0.01$
$dx = \Delta x = -0.02$	$dy = \Delta y = 0.01$

$= (2(-2) + 9(-2)^2\ (3)^2 + 4)\ (-4.02) + (-6(-2)^3\ (3) - 6(3)^2)\ (0.01)$

$= (-4 - 324 + 4\ )\ (-0.02) + (144 - 54)(0.01)$
$= (-324)\ (-0.02) + (90)\ (0.01)$

$= 6.48 + 0.9$

$= 7.38$

Encuentre *dw*.

1. $w = x^3 – x^2y + 3y^2$

$$dw = \frac{\partial w}{\partial x}\, dx + \frac{\partial w}{\partial y}\, dy$$
$$= (3x^2 – 2xy)\, dx + (x^2 + 6y)\, dy$$

2. w = x2 sen y + 2y3/2

$$dw = \frac{\partial w}{\partial x}\, dx + \frac{\partial w}{\partial y}\, dy$$

$$= (2x \text{ sen } y)\, dx + x^2 \frac{\partial}{\partial y} (\text{sen } y) + \text{sen } y \frac{\partial}{\partial y} (x^2) + 2 \left(\frac{3}{2}\ y^{-1/2} \right)$$

$$= (2x \text{ sen } y)\, dx + (x^2 \cos y + 3y^{1/2})\, dy$$

3. $w = x^2\, e^{xy} + (1/y^2)$

$$= x^2 \frac{\partial}{\partial x} (e^{xy}) + e^{xy} \frac{\partial}{\partial x} (x^2) + \frac{\partial}{\partial x} (1/y^2) + x^2 \frac{\partial}{\partial y} (e^{xy}) + e^{xy} \frac{\partial}{\partial y} (x^2) +$$

$$\frac{\partial}{\partial y} (1/y^2)$$

$$= x^2(xe^{xy}) + e^{xy}\, (2x) + 0 + x^2\, (ye^{xy}) + e^{xy}(0) + y^2 \frac{\frac{\partial}{\partial y} (– 1) \quad \frac{\partial}{\partial y} (y^2)}{y^4}$$

$$= x^3e^{xy} + 2xe^{xy} + x^2ye^{xy} + \left(\frac{-2y}{y^4} \right)$$

$$= x^3e^{xy} + 2xe^{xy} + x^2ye^{xy} – \frac{2}{y^3}$$

$$= x^3e^{xy} + 2xe^{xy} + x^2ye^{xy} - 2y^{-3}$$

$$= e^{xy}(x^2y + 2x)\,dx + (x^3 e^{xy} - 2y^{-3})\,dy$$

4. $w = x^2 \ln (y^2 + z^2)$

$$= x^2 \frac{\partial}{\partial x} \ln (y^2+z^2) + \ln (y^2+z^2) \frac{\partial}{\partial x} (x^2) + x^2 \frac{\partial}{\partial y} \ln (y^2+z^2) + \ln (y^2+z^2) \frac{\partial}{\partial y} (x^2)$$

$$+ x^2 \frac{\partial}{\partial z} \ln (y^2+z^2) + \ln (y^2 + z^2) \frac{\partial}{\partial z} (x^2)$$

$$= \ln (y^2 + z^2)(2x) + x^2 \left(\ln \frac{1}{y^2+z^2}\right)\left(\frac{\partial}{\partial y}\ y^2 + z^2\right) + x^2 \left(\ln \frac{1}{y^2 + z^2}\right)\left(\frac{\partial}{\partial z}\ y^2+z^2\right)$$

$$= 2x \ln (y^2 + z^2) + x^2 \left(\frac{1}{y^2+z^2}\right)(2y)\,dy + x^2\left(\frac{1}{y^2 + z^2}\right)(2z)\,dz$$

$$= 2x \ln (y^2 + z^2)dx + x^2 \left(\frac{2y}{y^2+z^2}\right) dy + x^2 \left(\frac{2z}{y^2 + z^2}\right) dz$$

$$= 2x \ln (y^2 + z^2)dx + \frac{2x^2y}{y^2+z^2}\,dy + \frac{2x^2z}{y^2 + z^2}\,dz$$

2. Cálculo integral en una variable

• Problemas aplicando el concepto de integral

Definición.- Sea *f* una función definida en un intervalo cerrado [a, b]. La integral definida de *f* desde a hasta b denotada por $\delta_a^b f(x)\,dx$, está dada por

$$\delta_a^b f(x)\,dx = \lim_{\|p\| \to 0} \sum_i f(w_i)\,\Delta x_i$$

siempre que el límite exista.

Ejemplos:

1. δ_{-3}^{2} (2x + 6) dx

$= \delta_{-3}^{2}$ x dx= $2x + 6 \Big|_{-3}^{2}$

$= \delta_{-3}^{2}$ 2x dx + δ_{-3}^{2} 6dx = $2\delta_{-3}^{2}$ x dx + $6\delta_{-3}^{2}$ dx

$$= 2\left(\frac{x^{1+1}}{1+1}\Big|_{-3}^{2}\right) + 6x\Big|_{-3}^{2}$$

$$= 2\left(\frac{x^{2}}{2}\Big|_{-3}^{2}\right) + 6x\Big|_{-3}^{2}$$

$$= x2\Big|_{-3}^{2} + 6x\Big|_{-3}^{2}$$

$= \{(2)^2 – (-3)^2\} + \{(6)2 – 6(-3)\}$

$= 4-9+12+18 = 25$

2. $\delta_{0}^{3} \sqrt{9-x^2}$ dx

$$= \frac{1}{2} \; x\sqrt{9-x^2} + \quad \frac{1}{2} \; - \; 9 \text{ arc sen*} \quad \frac{x}{3}$$

$$= \frac{1}{2} \; x\sqrt{9-x^2} + \frac{9}{2} \text{ arc sen } \frac{x}{3}$$

$$= \frac{1}{2} (3) \sqrt{9-x^2} + \frac{9}{2} \text{ arc sen } \frac{3}{3} - \frac{1}{2} (0) \sqrt{9-(0)^2} + \frac{9}{2} \text{ arc sen } \frac{0}{3}$$

$$= \frac{9}{2} \; (90)$$

*NOTA: Ejemplo
sen^{-1} = arc sen
$sen\theta = 10$
$\theta = sen^{-1} (10)$
θ = arc sen (10)

$$= \frac{9}{2} \quad \frac{\pi}{2}$$

$$= \frac{9\pi}{4}$$

3. Cálculo de varias variables

- **Gradiante de una función**

Definición.- Si f es una función de dos variables, entonces el gradiante de f se define como

$$\nabla f(x, y) = \langle f_x(x, y), f_y(x, y)\rangle = f_x(x, y)\mathbf{i} + f_y(x, y)\mathbf{j}.$$

- **Derivada direccional de una función**

Definición.- Si *f* es una función de *x* y *y* y $\boldsymbol{u} = \langle \cos\theta, \operatorname{sen}\theta\rangle$ es un vector unitario, entonces la **derivada direccional de *f* en la dirección de** $\vec{\mathbf{u}}$, denotada por $D_u f(x, y)$, está dada por

$$D_u f(x, y) = \lim_{t \to 0} \frac{f(x+t\cos\theta, y+t\operatorname{sen}\theta) - f(x, y)}{t}$$

- **Derivadas de funciones vectoriales**

Definición.- Una función vectorial *r* es continua en *a* si $\lim_{t \to a} r(t) = r(a)$

Definición.- Si r es una función vectorial, entonces la derivada de *r* es la función vectorial *r'* definida por

$$r'(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \quad r(t + \Delta t) - r(t)]$$

para todo t para el cual el límite existe.

4. Ecuaciones diferenciales

Si una ecuación contiene las derivadas diferenciales de una o más variables dependientes con respecto a una o más variables independientes, se dice que es una ecuación diferencial. Las ecuaciones diferenciales se clasifican de acuerdo con las propiedades siguientes:

Clasificación según el tipo

Si una ecuación contiene solo derivadas ordinarias de una o más variables dependientes con respecto a una sola variable independiente, se dice entonces que una ecuación diferencial ordinaria.

Ejemplo:

Las ecuaciones $\frac{dy}{dx} - 5y = 1,$

$$(x+y)\,dx - 4y\,dy = 0,$$

$$\frac{du}{dx} - \frac{du}{dx} = x,$$

$$\frac{d^2y}{dx^2} - 2\frac{dy}{dx} + 6y = 0$$

son ecuaciones diferenciales ordinarias

Una ecuación que contiene las derivadas parciales de una o más variables dependientes de dos o más variables independientes se llama ecuación diferencial parcial.

Ejemplo:

Las ecuaciones

$$\frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x},$$

$$x\frac{\partial u}{\partial y} + y\frac{\partial v}{\partial x} = u,$$

$$\frac{\partial^2 u}{\partial x \partial y} = x + y,$$

$$a^2\frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 u}{\partial t^2} - 2k\frac{\partial u}{\partial t}$$

son ecuaciones diferenciales parciales.

<u>Clasificación según el orden</u>

El orden de la derivada más alta en una ecuación diferencial se llama orden de la ecuación.

Ejemplo:

La ecuación $\frac{d^2y}{dx^2} \cdot 5 \left(\frac{dy}{dx}\right)^3 - 4y = x$ es una ecuación diferencial ordinaria de segundo orden. Puesto que la ecuación diferencial

$$x^2dy + y\,dx = 0$$

puede llevarse a la forma

$$x^2\,\frac{dy}{} + y = 0$$

dx

dividiendo entre la diferencial dx, es un ejemplo de ecuación diferencial ordinaria de primer orden. La ecuación

$$c^2 \frac{\partial^4 u}{\partial x^4} + \frac{\partial^2 u}{\partial t^2} = 0$$

es una ecuación diferencial parcial de cuarto orden.

Una ecuación diferencial ordinaria general de orden *n* se representa con frecuencia mediante la expresión simbólica

$$F\left(x, y, \frac{dy}{dx}, \dots, \frac{d^n y}{dx^n} \right) = 0$$

<u>Clasificación según la linealidad o no linealidad</u>

Se dice que una ecuación diferencial es lineal si tiene la forma

$$a_n(x) \frac{d^n y}{dx^n} + a_{n-1}(x) \frac{d^{n-1} y}{dx^{n-1}} + \dots + a_1(x) \frac{dy}{dx} + a_0(x)\, y = g(x)$$

Las ecuaciones diferenciales lineales se caracterizan por

(a) la variable dependiente *y* junto con todas sus derivadas son de primer grado; esto es, la potencia de cada término en *y* es 1.

(b) Cada coeficiente depende sólo de la variable independiente *x*.

Si no se cumple lo anterior la ecuación es no lineal.

Ejemplo:

Las ecuaciones

$$x\,dy + y\,dx = 0$$

$$y^n - 2y' + y = 0$$

$$x^3 \frac{d^3 y}{dx^3} - x^2 \frac{d^2 y}{dx^2} + 3x \frac{dy}{dx} + 5y = e^x$$

son ecuaciones diferenciales ordinarias lineales de primero, segundo y tercer orden, respectivamente.

$$\frac{dy}{dx} = xy^{1/2}$$

$$yy'' - 2y' = x + 1$$

$$\frac{d^3y}{dx^3} + y^2 = 0$$

son ecuaciones diferenciales ordinarias no lineales de primero, segundo y tercer orden, respectivamente.

5. Series de Fourier

•**Funciones periódicas**

Una función periódica se puede definir como una función para la cual $f(t) = f(t + T)^{(1.1)}$ para todo valor de t. La constante mínima T que satisface la relación (1.1), se obtiene

$$f(t) = f(t + nT), \qquad n = 0, \pm 1, \pm 2, \ldots$$

II. Álgebra

1. **Clásica**

• Funciones, tipos y propiedades

Función.- Es una relación que asigna a cada elemento del dominio uno y sólo un elemento del contradominio. Este último se llama el valor de la función para el elemento dado del dominio.

Una función f de A en B, se escribe f: A→B es la relación de A en B.

Propiedades.
(a) Dom (f) = A
(b) Si (a, b) y (a, c) pertenecen a f, entonces b = c.

La propiedad (b) dice que, si (a, b) $\in f$, entonces b está determinada únicamente por a. Por esta razón, también se escribe b = f (a) y se enlista la relación f como $\{(a, f(a)) | a \in A\}$. Las funciones son también llamadas **aplicaciones** o **transformaciones**.

Ejemplos:
1. Sean A = {1, 2, 3, 4} y B = {a, b, c, d} y sea

$$f = \{(1, a), (2, a), (3, d), (4, c)\} \text{ **}$$

Entonces f es una función, ya que ningún elemento de A aparece como primer elemento de dos pares ordenados diferentes. Aquí se tiene

$f(1) = a$
$f(2) = a$
$f(3) = d$
$f(4) = c$

El codominio de f, Cod (f) = {a, d, c}.

Una función puede tomar el mismo valor en dos elementos diferentes de A.

2. Sean A = {1, 2, 3} y B = {x, y, z}

R = {(1, x), (2, x)} y S = {(1, x), (2, z), (3, y)}

Ninguna de estas relaciones es una función de A en B, por diferentes razones. La relación S no es una función ya que contiene los pares ordenados (1, x)(1, y) lo que viola la propiedad (b) de la definición de una función.

La relación R no es una función de A en B, ya que el Dom (R) $\neq$ A.

<u>Tipos</u>

Una función f: A$\rightarrow$B se llama **inyectiva**, o uno a uno si para toda a, a' en A,

$$a \neq a' \text{ implica que } f(a) \neq f(a')$$

La función f definida en el ejemplo 1 (**) no es inyectiva ya que

$$f(1) = f(2) = a$$

♦Sea A = B = Z y sea f: A$\rightarrow$B definida por

$$f(a) = a + 1 \text{ para } a \in A$$

f consta de todos los pares ordenados (a, a+1) para a $\in$ Z. Entonces cada a $\in$ A aparece como el primer elemento de algún par, por lo cual Dom (f) = A. También, si (a, b) $\in f$ y (a, c) $\in f$, de modo que

$$b = f(a) = a+1$$

y

$$c = f(a) = a+1$$

entonces

$$b = c$$

Por consiguiente, f es una función. Supóngase que

$$f(a) = f(a')$$

para a y a' en A. Entonces

$$a+1 = a'+1$$

por lo cual

$$a = a'$$

De aquí que f sea inyectiva.

A una función *f*: A $\rightarrow$ B se le llama **suprayectiva** si *f*(A) = B, esto es, si el Cod(*f*)=B. *f* es suprayectiva si todo elemento b $\in$ B es el segundo elemento en algún par ordenado (a, b) $\in$ *f*.
f es suprayectiva si para cada b $\in$ B se puede encontrar alguna *a* $\in$ A tal que b=*f*(*a*).
Tomando el ejemplo con ♦ referencia . Sea b un elemento arbitrario de B.
Es posible encontrar un elemento *a* $\in$ A tal que

$$f(a) = b$$

ya que

$$f(a) = a+1$$

es necesario un elemento *a* en A tal que

$$a+1 = b$$

Por supuesto,

$$a = b-1$$

lo que satisface la ecuación deseada ya que b –1 está en A. De aquí que f sea suprayectiva.

Cuando una función es inyectiva y suprayectiva, se dice que f es una **biyección** o una correspondencia uno a uno.

A una función A $\rightarrow$ B se le llama **invertible** si su relación inversa, f^{-1} es también una función. Solo si f es inyectiva y suprayectiva (biyección) entonces es invertible.

• Números primos

Un número primo es un entero mayor que la unidad, que no tiene más factores enteros positivos que él mismo y la unidad. Los primeros números primos son: 1, 3, 5, 7, 11, 13, 17, 19, 23, ...

Todo entero positivo n>1 puede ser escrito en una sola forma así:

$$n = p_1^{k1} \; p_2^{k2} \ldots p_s^{ks}$$

donde $p_1 < p_2 < \ldots < p_s$ son los distintos números primos que dividen a *n* y las *k* son los enteros positivos que dan el número de veces en cada número primo ocurre como un factor de *n*.

Ejemplo:

$$9 = 3\bullet 3 = 3^2$$

$$24 = 12\bullet 2 = 2\bullet 2\bullet 2\bullet 3$$
$$= 2^3\bullet 3$$

$$30 = 2 \bullet 3 \bullet 5$$

- **Números complejos**

Los números complejos son pares ordenados de nuevos objetos para los que las nociones de igualdad, adición y multiplicación no están definidas inicialmente. Todo número complejo tiene raíces cuadradas.

2. Lineal

- **Sistemas de ecuaciones lineales**

Un par de ecuaciones lineales se puede resolver trazando la gráfica de ambas sobre los mismos ejes y determinando las coordenadas del punto de intersección.

Cualquier sucesión de valores x1 = s1 y x2 = s2 tales que

$$a_{11}s_1 + a_{12}s_2 = b_1$$
$$a_{21}s_1 + a_{22}s_2 = b_2$$

le llamamos una solución del sistema de ecuaciones lineales.

Si el sistema de ecuaciones lineales tiene solución se le llama compatible o consistente. Si no tiene solución le llamamos incompatible o inconsistente.

Sistema con solución única.

Considérese el sistema

$$x - y = 7$$
$$x + y = 5$$

Al sumar las dos ecuaciones se obtiene, por el resultado A, la ecuación siguiente: 2x = 12 (es decir, x = 6). Entonces, de la segunda ecuación, y = 5 – x = 5 – 6 = -1. Por lo tanto, el par (6, -1) satisface el sistema. Por la forma en que se encontró la solución, se ve que no existe ningún otro par que satisfaga ambas ecuaciones. Por tanto, el sistema tiene una **solución única**.

Sistema con un número infinito de soluciones

Considérese el sistema

$$x - y = 7$$
$$2x - 2y = 14$$

Es obvio que estas dos ecuaciones son equivalentes. A fin de comprobar esto, multiplíquese la primera por 2. x - y = 7 o y = x –7. Por tanto, el par (x, x-7) es una solución del sistema para todo número real x. El sistema tiene un número infinito de soluciones. Por ejemplo, los pares siguientes son soluciones: (7, 0), (0, -7), (8, 1), (1, -6), (3, -4) y (-2, -9).

Sistema sin solución

Considérese el sistema

$$x - y = 7$$
$$2x - 2y = 13$$

Multiplicando la primera ecuación por 2, se obtiene $2x - 2y = 14$. Esto contradice a la segunda ecuación.
Entonces el sistema no tiene solución.

Ejemplo:
Resolver el sistema

$$\begin{aligned} 2x_1 + 4x_2 + 6x_3 &= 18 \\ 4x_1 + 5x_2 + 6x_3 &= 24 \\ 3x_1 + x_2 - 2x_3 &= 4 \end{aligned}$$

Dividiendo la primera ecuación entre 2.

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 9 \\ 4x_1 + 5x_2 + 6x_3 &= 24 \\ 3x_1 + x_2 - 2x_3 &= 4 \end{aligned}$$

Multiplicando por –4 ambos lados de la primera ecuación y sumando esta nueva ecuación a la segunda. Se obtiene entonces

$$\begin{array}{r} -4x_1 - 8x_2 - 12x_3 = -36 \\ 4x_1 + 5x_2 + 6x_3 = 24 \\ \hline -3x_2 - 6x_3 = -12 \end{array}$$

El sistema ahora es

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 9 \\ -3x_2 - 6x_3 &= -12 \\ 3x_1 + x_2 - 2x_3 &= 4 \end{aligned}$$

La primera ecuación se multiplica por –3 y el resultado se suma a la tercera ecuación:

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 9 \\ -3x_2 - 6x_3 &= -12 \\ -5x_2 - 11x_3 &= -23 \end{aligned}$$

La segunda ecuación se divide entre –3:

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 9 \\ x_2 + 2x_3 &= 4 \\ -5x_2 - 11x_3 &= -23 \end{aligned}$$

La segunda ecuación se multiplica por –2 y el resultado se suma a la primera, y luego la segunda ecuación se multiplica por 5 y el resultado se suma a la tercera:

$$\begin{aligned} x_1 \quad - x_3 &= 1 \\ x_2 + 2x_3 &= 4 \\ - x_3 &= -3 \end{aligned}$$

La tercera ecuación se multiplica por –1:

$$\begin{aligned} x_1 \quad - x_3 &= 1 \\ x_2 + 2x_3 &= 4 \end{aligned}$$

$$x_3 = 3$$

Por último, la tercera ecuación se suma a la primera y luego la tercera Ecuación se multiplica por –2 y el resultado se suma a la segunda, obteniéndose el sistema siguiente [el cual es equivalente al primer sistema]:

$$\begin{aligned} x_1 \quad\quad &= 4 \\ x_2 \quad &= -2 \\ x_3 &= 3 \end{aligned}$$

- **Triangulación y diagonalización**

Se dice que una matriz cuadrada es **triangular superior** si todos los elementos situados debajo de su diagonal principal son cero.

Se dice que una matriz cuadrada es **triangular inferior** si todos los elementos situados arriba de su diagonal principal son cero.

Ejemplo:

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & -5 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ -5 & 1 & 0 \\ 2 & 3 & 1 \end{pmatrix}$$

triangular superior　　　　triangular inferior

- **Norma**

Longitud o norma de un vector.- Si $v \in \Re^n$, entonces la longitud o norma de v, denotada por $|v|$, está dada por

$$|v| = \sqrt{v \bullet v}$$

Ejemplo:

Norma de un vector en $\Re^2$ Sea $v = (x, y) \in \Re^2$. Entonces $|v| = \sqrt{x^2+y^2}$ es la definición ordinaria de la longitud de un vector en el plano.

Norma de un vector en $\Re^3$ Si $v = (x, y, z) \in \Re3$, entonces $|v| = \sqrt{x^2 + y^2 + z^2}$

Norma de un vector en $\Re^5$ Si $v = (2, -1, 3, 4, -6) \in \Re^5$, entonces $|v| = \sqrt{4 + 1 + 9 + 16 +36} = \sqrt{66}$.

- **Proyecciones**

Definición.- Sean u y v en $\Re^2$ diferentes de cero. La proyección de u en (sobre) v es el vector, denotado por $proy_v^u$, definido por:

$$proy_v^u = \frac{u \bullet v}{||v||^2}\ v$$

Al escalar $\frac{u \bullet v}{||v||^2}$ le llamamos la componente de u en la dirección de v.

Ejemplos:

u = (2, 5)
v = (7, 3)

Encontrar $proy_v^u$

Encontramos la componente de u en la dirección de v

$u \bullet v = (2)(7) + (5)(3) = 29$

$||v||^2 = \sqrt{v_1^2 + v_2^2}$

$||v||^2 = 7^2 + 3^2 = 58$

$$\frac{u \bullet v}{||v||^2} = \frac{29}{58} = \frac{1}{2}$$

entonces $proy_v^u = \frac{1}{2}\ (7, 3)$

$$= \left(\frac{7}{2}, \frac{3}{2}\right) = (3.5,\ 1.5)$$

¿Cuál es la distancia de u a $proy_v^u$?

$$d(u, proy_v^u) = ||u - proy_v^u||$$

$= ||(-1.5, 3.5)||$
$= \sqrt{(-1.5)^2 + (3.5)^2}$
$= \sqrt{14.5}$
$= 3.8$

- **Bases ortogonales y ortonormales**

Definición.- Conjunto ortonormal en $\Re^n$ El conjunto de vectores $S = \{u_1, u_2, ..., u_k\}$ en $\Re^n$ se llama conjunto ortonormal si

$$u_i \bullet u_j = 0 \qquad \text{si } i \neq j$$
$$u_i \bullet u_j = 1$$

Si sólo se satisface la ecuación (1), se dice que el conjunto es ortogonal.

Un conjunto de vectores es ortonormal si un par cualquiera de ellos es ortogonal y si cada uno tiene longitud 1.

3. Teoría de grupos

- **Grupos**

Un grupo (G, *) es un monoide, con idéntico e, que tiene la propiedad adicional de que, para cualquier elemento $a \in G$, existe un elemento $a' \in G$ tal que $a * a' = a' * a = e$. Por consiguiente, un grupo es un conjunto G con una operación binaria * en G tal que

1. $(a * b) * c = a * (b * c)$ para elementos cualquiera a, b, y c en G.
2. Existe un elemento único e en G tal que
$$a * e = e * a \qquad \text{para cualquier } a \in G$$
3. Para cada $a \in G$ existe un elemento $a' \in G$, al que se le llama inverso de a, tal que
$$a * a' = a' * a = e$$

Se dice que un grupo G es **abeliano** o conmutativo si $ab = ba$ para todos los elementos a y b en G.

Ejemplos:
1. El conjunto de todos los enteros Z con la operación de suma ordinaria es un grupo abeliano. Si $a \in Z$, entonces el inverso de a es el negativo $-a$.
2. El conjunto Z^+ bajo la operación de multiplicación ordinaria no es un grupo ya que el elemento 2 en Z^+ no tiene inverso. Sin embargo, este conjunto con la operación dada es un monoide.

3. El conjunto de los números reales sin el cero bajo la operación de multiplicación ordinaria es un grupo. Un inverso de $a \neq 0$ es $1/a$.

4. Sea G el conjunto de los números reales sin el cero y sea
$$a * b = \frac{ab}{2}$$
Demuestre que (G, *) es un grupo abeliano.

Solución. Primero, se verificara que * es una operación binaria. Si a y b son elementos de G, entonces a * b (= ab/2) es un número real diferente de cero y, por tanto, está en G. En seguida se verificara su propiedad asociativa. Como

$$(a * b) * c = \left(\frac{ab}{2}\right) * c = \frac{(ab)\,c}{4}$$

y

$$a * (a * b) = a * \left(\frac{ab}{2}\right) = \frac{a\,(ab)}{4}$$

la operación * es asociativa.
El número 2 es el idéntico en G, si $a \in G$, entonces,

$$a * 2 = \frac{(a)(2)}{2} = a = \frac{(2)(a)}{2} = 2 * a$$

Por último, si $a \in G$, entonces $a' = 4/a$ es un inverso de a ya que,

$$a * a = a * \frac{4}{a} = \frac{a(4/a)}{2} = 2 = \frac{(4/a)(a)}{2} = \frac{4}{a} * a = a' * a$$

Como $a * b = b * a$ para todas las a y b en G, se concluye que G es un grupo abeliano.

Propiedades que son satisfechas por cualquier grupo *G*.

1. Sea G un grupo. Cada elemento a en G tiene un inverso único en G.

   $$aa^{-1} = a^{-1}a = e$$

2. Sea G un grupo y sean a, b y c elementos en G. Entonces,
   (a) $ab = ac$ implica que $b = c$ (propiedad de cancelación izquierda).
   (b) $ba = ca$ implica que $b = c$ (propiedad de cancelación derecha).

3. Sea G un grupo y sean a y b elementos en G. Entonces,
   (a) $(a^{-1})^{-1} = a$
   (b) $(ab)^{-1} = b^{-1}a^{-1}$

4. Sea G un grupo y sean a y b elementos de G. Entonces,
   (a) La ecuación $ax = b$ tiene una solución única en G.
   (b) La ecuación $ya = b$ tiene una solución única en G.

Si un grupo G tiene un número finito de elementos, su operación binaria puede darse por una tabla de multiplicación. La tabla de multiplicación del grupo $G = \{a_1, a_2, ..., a_n\}$ bajo la operación binaria * deberá satisfacer las siguientes propiedades:

1. El renglón etiquetado por e deberá de contener los elementos

$$a_1, a_2, ..., a_n$$

y la columna etiquetada por *e* deberá contener los elementos

$$\begin{array}{c} a_1 \\ a_2 \\ . \\ . \\ . \\ a_n \end{array}$$

2. Cada elemento b en el grupo deberá aparecer exactamente una vez en cada renglón y en cada columna de la tabla. Por tanto, cada columna y cada renglón es una permutación de los elementos $a_1, a_2, ... , a_n$ de *G* y cada renglón (y cada columna) determina una permutación diferente.

Si *G* es un grupo que tiene un número finito de elementos, se dice que *G* es un grupo finito y el orden de *G* es el número de elementos $|G|$ en *G*. Las tablas de multiplicación de todos los grupos de órdenes 1, 2, 3 y 4 son:

Si *G* es un grupo de orden 1, entonces $G = \{e\}$, y se tiene $ee = e$. Ahora sea $G = \{e, a\}$ un grupo de orden 2. Entonces se obtendrá la tabla de multiplicación

	e	*a*
e	*e*	*a*
a	*a*	

donde es necesario llenar el espacio en blanco. Puede llenarse con *a* o *e*. Como no es posible repetir elementos en un mismo renglón o columna, se deberá escribir *e* en el espacio en blanco.

	e	*a*
e	*e*	*a*
a	*a*	

Esta tabla es de orden 2.

Sea $G = \{e, a, b\}$ un grupo de orden 3. Se tiene la tabla de multiplicación donde es necesario llenar los cuatro espacios en blanco.

	e	*a*	*b*
e	*e*	*a*	*b*
a	*a*		
b	*b*		

	e	*a*	*b*
e	*e*	*a*	*b*
a	*a*	*b*	*e*
b	*b*	*e*	*a*

Sea $G = \{e, a, b, c\}$ de orden 4. La tabla de multiplicación es:

	e	*a*	*b*	*c*

	e	*a*	*b*	*c*

	e	*a*	*b*	*c*

	e	*a*	*b*	*c*

e	*e*	*a*	*b*	*c*
a	*a*	*e*	*c*	*b*
b	*b*	*c*	*e*	*a*
c	*c*	*b*	*a*	*e*

e	*e*	*a*	*b*	*c*
a	*a*	*e*	*c*	*b*
b	*b*	*c*	*a*	*e*
c	*c*	*b*	*e*	*a*

e	*e*	*a*	*b*	*c*
a	*a*	*b*	*c*	*e*
b	*b*	*c*	*e*	*a*
c	*c*	*e*	*a*	*b*

e	*e*	*a*	*b*	*c*
a	*a*	*c*	*e*	*b*
b	*b*	*e*	*c*	*a*
c	*c*	*b*	*a*	*e*

Subconjuntos de un grupo G

Sea H un subconjunto de un grupo G tal que:
(a) El idéntico e de G pertenece a H.
(b) Si a y b pertenecen a H, entonces $ab \in H$.
(c) Si $a \in H$, entonces $a^{-1} \in H$.

Entonces, a H se le llama **subgrupo** de G. La parte (b) anterior, dice que H es un subsemigrupo de G. Un subgrupo de G puede verse como un subsemigrupo que tiene las propiedades (a) y (c) anteriores.

Ejemplo:
Sea G un grupo. Entonces G y $H = \{e\}$ son subgrupos de G, a estos se les llama subgrupos triviales de G.

5. sean (G, *) y (G', *') dos grupos y sea $f : G \to G'$ un homomorfismo de G en G'.
   (a) Si e es idéntico en G y e' es el idéntico en G', entonces $f(e) = e'$.
   (b) Si $a \in G$, entonces $f(a^{-1}) = (f(a))^{-1}$.
   (c) Si H es un subgrupo de G, entonces
   $$f(H) = \{f(h) | h \in H\}$$
   es un subgrupo de G'.

Productos y cocientes de los grupos

Si G_1 y G_2 son grupos, entonces $G = G_1 \times G_2$ es un grupo con la operación definida por
$$(a_1, b_1)(a_2, b_2) = (a_1a_2, b_1b_2)$$

Sea R una relación de congruencia en el grupo (G, *). Entonces el semigrupo (G/R, $\otimes$) es un grupo, donde la operación $\otimes$ se define en G/R por
$$[a] \otimes [b] = [a * b]$$

(a) Si R es una relación de congruencia en un grupo G, entonces la función $f_R : G \to G/R$, dada por $f_R(a) = [a]$, es un homomorfismo de grupos.
(b) Si $f : G \to G'$ es un homomorfismo del grupo (G', *) en el grupo (G', *'), y si R es la relación definida en G por $a\ R\ b$ si y sólo si $f(a) = f(b)$, para las a y b en G, entonces:

(1) R es una relación de congruencia.
(2) La función $\overline{f} : G/R \to G'$, dada por $\overline{f}$ ([a]) = $f(a)$, es un isomorfismo del grupo (G/R, $\otimes$) en el grupo (G', *').

Sea H un subgrupo de un grupo G y sea $a \in G$. La clase lateral izquierda de H en G determinada por a es el conjunto $aH = \{ah|h \in H\}$. La clase lateral Derecha de H en G determinada por a es el conjunto $Ha = \{ha|h \in H\}$. Un Subgrupo H de G es normal si $aH = Ha$ para todas las a en G.

Sea R una relación de congruencia en el grupo G y sea H = [e], la clase de equivalencia que contiene al idéntico. Entonces H es un subgrupo normal de G para cada $a \in G$, [a] = $aH = Ha$.

Sea N un subgrupo normal de un grupo G y sea R la siguiente relación en G:

$$a\ R\ b \quad \text{si y sólo si} \quad a^{-1}\ b \in N.$$

Entonces:
(a) R es una relación de congruencia en G.
(b) N es la clase de equivalencia [e] respecto a R, donde e es el identico de G.

III. Geometría analítica

Lineal

GEOMETRÍA ANALÍTICA DE LA RECTA

Los primeros métodos de la geometría analítica se deben a Menaecmo (aprox. 350 a.C.), quien llega a plantearse problemas de intersección de superficies, aplicando técnicas que, si bien no incluyen todavía las coordenadas, las llevan ocultas en su tratamiento conceptual.

Algo parecido ocurre con Apolonio de Perga (250 a.C.-190 a.C. aprox.), el cual demostró diversos resultados relacionados con rectas y circunferencias empleando técnicas similares a las de Menaecmo.

El matemático parisino Nicole Oresme (1321-1382), obispo de Lisieux, hizo algunos trabajos haciendo uso de la longitud y la latitud, equivalentes a las actuales abscisa y ordenada.

La geometría analítica propiamente dicha comienza con los matemáticos René Descartes (1596-1650) y Pierre de Fermat (1601-1665), quienes en sus trabajos llegan a considerar sistemas de coordenadas, aunque sólo admitían coordenadas positivas. El principal logro de la misma es la transformación mutua entre enunciados de tipo geométrico y enunciados de tipo algebraico.

Fermat, en su *Introduction to Loci* , estudia ya algunas ecuaciones de primero y segundo grado, con lo que consigue clasificar las rectas y algunas de las cónicas, siempre con la limitación que le imponía el no admitir coordenadas negativas.

A principios del siglo XIX, con la construcción de la geometría proyectiva, se dio un fuerte avance a la geometría analítica.

SISTEMAS DE REFERENCIA - COORDENADAS

Un *sistema de referencia* en el plano es un par formado por un punto, llamado *origen*, y una base de vectores, $R = \{O, \vec{u}1, \vec{u}2\}$.

Si la base es ortonormal, el *sistema de referencia* se dice *ortonormal* .

Dado un sistema de referencia $R = \{O, \vec{u}1, \vec{u}2\}$ y un punto P del plano, las coordenadas de P respecto a R son las coordenadas del vector $\overrightarrow{OP}$ respecto a la base $\{\vec{u}1, \vec{u}2\}$

Si $\overrightarrow{OP} = x\vec{u}11 + y\vec{u}2$, las coordenadas de P son x e y y se escribe $P = (x, y)$ ó $P(x, y)$.

Vector que une dos puntos

Dados los puntos $P(x0, y0)$ y $Q(x1, y1)$, el vector que los une viene dado por la expresión $\overrightarrow{PQ} = (x1 - x0)\vec{u}1 + (y1 - y0)\vec{u}2$.

Demostración:

Por definición de coordenadas se tiene que

$\overrightarrow{OP} = x0\,\vec{u}1 + y0\,\vec{u}2$

$\overrightarrow{OQ} = x1\,\vec{u}1 + y1\,\vec{u}2$

Como $\overrightarrow{OP} + \overrightarrow{PQ} = \overrightarrow{OQ}$, despejando:

$\overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP} = (x1\,\vec{u}1 + y1\,\vec{u}2) - (x0\,\vec{u}1 + y0\,\vec{u}2) =$

$= (x1 - x0)\vec{u}1 + (y1 - y0)\vec{u}2$

Sistema de referencia estándar del plano

El sistema de referencia del plano más utilizado es $\{O, \vec{u}1, \vec{u}2\}$, donde:

- O es el origen de coordenadas, después de fijar los ejes de abscisas y ordenadas.
- $\vec{u}1$ es el vector que tiene por origen el punto (0,0) y por extremo el (1,0).
- $\vec{u}2$ es el vector que tiene por origen el punto (0,0) y por extremo el (0,1).

Además es un sistema de referencia ortonormal:

$$|\vec{u}_1| = \sqrt{1^2 + 0^2} = \sqrt{1} = 1 \quad |\vec{u}_2| = \sqrt{0^2 + 1^2} = 1$$

$$y\ \vec{u}_1 \bullet \vec{u}_1 = |\vec{u}_1| \cdot |\vec{u}_1| \cdot \cos 0^o = 1 \cdot 1 \cdot 1 = 1; \quad \vec{u}_2 \bullet \vec{u}_2 = |\vec{u}_2| \cdot |\vec{u}_2| \cdot \cos 0^o = 1 \cdot 1 \cdot 1 = 1$$

$$\vec{u}_1 \bullet \vec{u}_2 = |\vec{u}_1||\vec{u}_2| \cdot \cos 90^o = 1 \cdot 1 \cdot 0 = 0$$

Vectores en el plano

- Todo punto P del plano determina un vector cuyo origen sea el origen de coordenadas y el extremo sea el punto P. Es decir, un punto P determina el vector $\overrightarrow{OP}$.
- Recíprocamente, dados dos puntos del plano S y Q, de coordenadas $S(x0, y0)$ y

$Q(x1, y1)$, existe un vector equipolente al vector $\overrightarrow{SQ}$ cuyo origen coincide con el origen de coordenadas. Este vector es el que tiene por extremo ($x1 - x0$, $y1 - y0$).

A partir de ahora se trabajará con vectores que tienen por origen el origen de coordenadas y por extremo cualquier otro punto del plano, *P*. En este caso las coordenadas del vector coincidirán con las coordenadas del punto *P*.

Por tanto, al hablar del vector (4, -3), por ejemplo, ha de interpretarse el vector que tiene por origen *O* y por extremo (4, -3).

Punto medio de un segmento

Las coordenadas del punto medio, *M*, de un segmento cuyos extremos son $P(x0, y0)$ y $Q(x1, y1)$ vienen dadas por:

$$M\left(\frac{x_0 + x_1}{2}, \frac{y_0 + y_1}{2}\right)$$

Demostración:

$$\overrightarrow{PM} = \frac{1}{2} \cdot \overrightarrow{PQ}$$

Sumando $\overrightarrow{OP}$ a ambos miembros se tiene:

$$\overrightarrow{OM} = \overrightarrow{OP} + \overrightarrow{PM} = \overrightarrow{OP} + \frac{\overrightarrow{PQ}}{2} = \frac{2 \cdot \overrightarrow{OP} + \overrightarrow{PQ}}{2} = \frac{\overrightarrow{OP} + (\overrightarrow{OP} + \overrightarrow{PQ})}{2} = \frac{\overrightarrow{OP} + \overrightarrow{OQ}}{2}$$

$$\overrightarrow{OM} = \frac{\overrightarrow{OP} + \overrightarrow{OQ}}{2} = \frac{x_0\vec{u}_1 + y_0\vec{u}_2 + x_1\vec{u}_1 + y_1\vec{u}_2}{2} = \frac{x_0 + x_1}{2}\vec{u}_1 + \frac{y_0 + y_1}{2}\vec{u}_2$$

Luego M tiene por coordenadas $\left(\frac{x_0 + x_1}{2}, \frac{y_0 + y_1}{2}\right)$.

ECUACION PARAMÉTRICA DE LA RECTA

Una recta queda perfectamente determinada si:

A) Se conoce un punto y su dirección (vector paralelo a la recta).

B) Se conocen dos puntos de ella.

A) *Ecuación paramétrica de la recta conocido un punto y su dirección*

Sea el punto $P(x0, y0)$ de la recta y el vector $\vec{u}$, de coordenadas (*a*, *b*), que determina la dirección de la recta, ambos conocidos. Sea *Q* un punto genérico de la recta, cuyas coordenadas (*x, y*) no se conocen. Hay que estudiar la condición que han de cumplir

x e *y*.

Suponiendo que *Q* pertenece a la recta, el vector $\overrightarrow{PQ}$ es paralelo al vector $\vec{u}$, luego existe un número real *t* tal que $\overrightarrow{PQ} = t \cdot \vec{u}$.

Escribiendo esta expresión en forma de coordenadas,

$(x - x_0, y - y_0) = t(a, b)$ □ $(x - x_0, y - y_0) = (ta, tb)$

Por tanto,

$x - x_0 = ta$ $x = x_0 + ta$

□

$y - y_0 = tb$ $y = y_0 + tb$

Estas ecuaciones se llaman *ecuaciones paramétricas* de la recta. Dando valores a t, llamado *parámetro*, se pueden obtener tantos puntos de la recta como se desee.

B) *Ecuaciones paramétricas de la recta conocidos dos puntos*

Si se conocen dos puntos de la recta, $P(x_0, y_0)$, y $Q(x_1, y_1)$, es claro que el vector de dirección es el vector $\overrightarrow{PQ}$ de coordenadas $(x_1 - x_0, y_1 - y_0)$.

Por tanto, el problema consiste en encontrar la ecuación de la recta que pasa por

el punto $P(x_0, y_0)$ y tiene por vector de dirección $(x_1 - x_0, y_1 - y_0)$.

Si $R(x, y)$ es un punto genérico de la recta,

$x = x_0 + t(x_1 - x_0)$

$y = y_0 + t(y_1 - y_0)$

ECUACIÓN CONTINUA DE LA RECTA

Dados un punto $P(x_0, y_0)$ de una recta y su vector direccional $\vec{u}(a, b)$, las ecuaciones paramétricas de la recta son:

$x = x_0 + ta$

$y = y_0 + tb$

Despejando t en ambas ecuaciones:

$$t = \frac{x - x_0}{a}, \quad t = \frac{y - y_0}{b}$$

Puesto que el valor de t es común para las dos ecuaciones:

$$\frac{x - x_0}{a} = \frac{y - y_0}{b}$$

Observación:

Puesto que se ha dividido entre a y b, se ha de suponer que a □ 0 y b □ 0.

En el caso en que una de las dos coordenadas del vector sea cero [no puede ser

$a = 0$ y $b = 0$ pues el vector $\vec{u}(a, b)$ sería $\vec{u}(0, 0)$, un punto], por ejemplo $a = 0$, las

ecuaciones paramétricas son:

$x = x0$

$y = y0 + tb$

Por tanto, un punto (x, y) pertenece a esta recta siempre que $x = x0$ e y tome cualquier valor. En consecuencia, se admite la anterior ecuación aun cuando uno de los números a, ó b sean cero, siempre que se interprete que si el denominador de una fracción es 0, debe ser 0 su numerador correspondiente.

PENDIENTE DE UNA RECTA

Se llama *pendiente* de una recta a la tangente del ángulo que forma la recta con el semieje positivo de abscisas, medido siempre en sentido contrario al de las agujas de un reloj.

La pendiente de la recta es *tg* □.

Al ángulo □ se le llama *inclinación* de la recta.

<u>**Interpretación de la pendiente de una recta**</u>

Si $\vec{u}$ es el vector direccional de una recta, de coordenadas $\vec{u}(a, b)$, que forma un ángulo □ con el semieje positivo de abscisas, trazando una circunferencia de centro O y radio el módulo de $\vec{u}$, $|\vec{u}|$, se sabe por trigonometría que:

$$sen\ \alpha = \frac{b}{|\vec{u}|},\quad cos\ \alpha = \frac{a}{|\vec{u}|},\quad tg\ \alpha = \frac{sen\ \alpha}{cos\ \alpha} = \frac{b}{a}$$

Esta igualdad proporciona un método sencillo para calcular la pendiente de una recta.

Si no se conoce el vector de dirección pero sí dos puntos de la recta, el vector de dirección es $\overrightarrow{PQ}(x_1 - x_0,\ y_1 - y_0)$, y la pendiente es $tg\ \alpha = \dfrac{y_1 - y_0}{x_1 - x_0}$.

La pendiente de una recta se la suele denotar con la letra m.

Por tanto, $m = tg$ □.

ECUACION DE LA RECTA (PUNTO-PENDIENTE)

Sea una recta que pasa por el punto $P(x0,\ y0)$ y tiene por vector direccional $\vec{u}(a, b)$. Su ecuación continua es, como se sabe:

$\dfrac{x - x_0}{a} = \dfrac{y - y_0}{b}$. A partir de esta igualdad se obtiene:

$$a(y - y_0) = b(x - x_0) \Rightarrow y - y_0 = \frac{b}{a}(x - x_0)$$

Pero $\dfrac{b}{a} = m$; por tanto:

$y - y_0 = m(x - x_0)$

ECUACIÓN EXPLÍCITA DE LA RECTA

Despejando *y* en la ecuación forma punto-pendiente de la recta:

$y - y_0 = m(x - x_0) \Rightarrow y = mx - mx_0 + y_0$

Llamando $p = -mx_0 + y_0$, que es un número conocido, resulta:

$y = mx + p$

Ésta es la llamada *ecuación explícita* de la recta.

Al número *p* se le llama *ordenada en el origen de la recta*. En la ecuación

$y = mx + p$, haciendo $x = 0$, resulta $y = m \cdot 0 + p = p$, por tanto, la recta pasa por el punto $(0, p)$, de aquí el nombre que se le da a *p*.

ECUACIÓN IMPLÍCITA DE LA RECTA

Si a partir de cualquiera de las ecuaciones de la recta se trasladan todos los términos al primer miembro, se obtiene una ecuación de la forma $Ax + By + C = 0$, donde *A, B* y *C* son números conocidos y *x* e *y* son las incógnitas.

En esta ecuación conviene considerar dos casos:

a) Si $B = 0$, la ecuación es $Ax + C = 0$, y despejando x, $x = -\frac{C}{A}$.

La recta es paralela al eje de ordenadas y su inclinación es, en consecuencia, de 90º.

b) Si $B \neq 0$, se despeja y en la ecuación $Ax + By + C = 0$.

$By = -Ax - C$, y dividiendo entre *B,*

$$y = -\frac{A}{B}x - \frac{C}{B}$$

Comparando esta igualdad con la ecuación explícita ya obtenida $y = mx + p$, se deduce que:

$$m = -\frac{A}{B} \quad \text{y} \quad p = -\frac{C}{B}$$

ECUACIÓN CANÓNICA DE UNA RECTA

La *ordenada en el origen* de una recta es la ordenada del punto de la recta cuya abscisa es cero. Es de la forma $(0, p)$.

La *abscisa en el origen* de una recta es la abscisa del punto cuya ordenada es cero. Es de la forma $(l, 0)$.

Suponiendo que una recta tiene ordenada en el origen y abscisa en el origen, la ecuación de esta recta [pasa por $(0, p)$ y $(l, 0)$] será:

$$\frac{x - 0}{l} = \frac{y - p}{-p} \Rightarrow \frac{x}{l} = -\frac{y}{p} + \frac{p}{p} \Rightarrow$$

$$\Rightarrow \frac{x}{l} - \frac{y}{p} = 1$$

REPRESENTACIÓN DE UNA RECTA

Una recta queda determinada por dos puntos. Esto es, conocidos dos puntos, basta unirlos con una regla para tener trazada la recta.

POSICION RELATIVA DE DOS RECTAS (I)

Dadas dos rectas *r* y *s*, las posiciones que pueden tener una respecto a otra son:

- *Coincidentes*. Las rectas son iguales.
- *Secantes*. Las rectas se cortan en un punto.
- *Paralelas.* Las rectas son distintas y no se cortan en ningún punto.

Para averiguar, a partir de las ecuaciones de las rectas, cuál es su posición relativa, se distinguirán varios casos:

A) Ecuaciones dadas en forma explícita

Sean *r*: $y = mx + p$

s: $y = m'x + p'$

r y *s* son coincidentes si $m = m''$ y $p = p'$

□ *r* y *s* son paralelas y no coincidentes si $m = m'$ y *p* □ *p'*

□ *r* y *s* son secantes si *m* □ *m'*

B) Ecuaciones dadas en forma general

Sean $r : Ax + By + C = 0$

$s : A' x + B' y + C' = 0$

- *r* y *s* son coincidentes si $\frac{A}{A'} = \frac{B}{B'} = \frac{C}{C'}$
- *r* y *s* son paralelas si $\frac{A}{A'} = \frac{B}{B'} \neq \frac{C}{C'}$
- *r* y *s* se cortan en un punto si $\frac{A}{A'} \neq \frac{B}{B'}$

<u>Cálculo del punto intersección de dos rectas</u>

Si dos rectas *r* y *s* se cortan en un punto, este punto ha de verificar las ecuaciones

de *r* y *s.*

Por tanto, calcular el punto de intersección de *r* y *s* consiste en resolver el sistema formado por las dos ecuaciones de las rectas.

<u>Ángulo de dos rectas</u>

Sea *r* una recta de pendiente *m* e inclinación □.

Sea *s* otra recta de pendiente *m*' e inclinación □.

Se trata de calcular el ángulo □ que forman las dos rectas.

Se considera el triángulo formado por *r, s* y el eje de abscisas.

El ángulo □ es un ángulo exterior de dicho triángulo, por lo que es igual a la suma de los dos ángulos interiores no adyacentes; es decir,

□ = □□; de donde □□□ - □

$$\left.\begin{array}{l}\text{Como } tg\,\alpha = m \\ \text{y } tg\,\beta = m'\end{array}\right\}\ tg\,\varnothing = tg(\alpha \quad \beta) = \frac{tg\,\alpha - tg\,\beta}{1 + tg\,\alpha \cdot tg\,\beta} = \frac{m - m'}{1 + mm'}$$

Así pues,

$$tg\,\varnothing = \frac{m - m'}{1 + mm'}$$

□ *Observación:*

Como al cruzarse dos rectas se forman dos ángulos distintos, se considera que el ángulo que forman las rectas es el agudo, por lo que si saliese negativa la tangente, habría que cambiarla de signo (recuérdese que las tangentes de dos ángulos suplementarios son opuestas). Así, la fórmula correcta será:

$$tg\,\varnothing = \left|\frac{m - m'}{1 + m \cdot m'}\right|$$

POSICION RELATIVA DE DOS RECTAS (II)

<u>Perpendicularidad de rectas</u>

La condición para que dos rectas sean perpendiculares es que el producto de sus pendientes sea -1.

Demostración:

Un ángulo recto no tiene tangente, lo cual, traducido a la fórmula anterior, ocurre únicamente si el denominador es cero. Así, las dos rectas son perpendiculares si

$1 + m \cdot m' = 0$, o lo que es lo mismo si $m \cdot m' = -1$, que era la condición anunciada.

<u>Distancia entre dos puntos</u>

La distancia entre los puntos $P(x0, y0)$ y $Q(x1, y1)$ viene expresada por la fórmula

$$d(P,\ Q) - \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2}$$

Demostración:

La distancia entre *P* y *Q* es el módulo del vector $\overrightarrow{PQ}$= (x1 - x0)$\vec{u}$1 + (y1 - y0)$\vec{u}$2

$$\left|\overrightarrow{PQ}\right| = \sqrt{\overrightarrow{PQ} \bullet \overrightarrow{PQ}} = \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2}$$

<u>Distancia de un punto a una recta</u>

La distancia de un punto $P(x0, y0)$ a la recta que tiene por ecuación general

r: $Ax + By + C = 0$ es:

$$d = \frac{|Ax_0 + By_0 + C|}{\sqrt{A^2 + B^2}}$$

Demostración:

Si en la ecuación de la recta se pasa *y* al segundo miembro, *Ax* + *C* = -*By*; y dividiendo ambos miembros por -*AB*:

$$\frac{Ax + C}{-AB} = \frac{-By}{-AB}, \text{ o bien, } \frac{x + C/A}{-B} = \frac{y - 0}{A}$$

Ésta es la ecuación continua de la recta, cuyo vector de dirección es el vector

(- *B, A*).

Elegido el sistema de referencia del plano, se observa que si se llama $\vec{a}$ al vector $\vec{a}$(*A, B*), el producto escalar de por $\vec{a}$ es:

$\vec{u}$ □ $\vec{a}$= (-*B,A*) □ (*A, B*) = -*BA* + *AB* = 0, lo que indica que los vectores $\vec{u}$ y $\vec{a}$ son

perpendiculares.

Se elige un punto cualquiera de la recta, *E*(*x*1, *y*1); se une *E* con *P* y se traza la perpendicular desde P hasta la recta *r* , prolongando el vector $\vec{a}$ hasta encontrar el punto de intersección *H*.

El triángulo *EPH* es rectángulo, con lo que

$$\cos\alpha = \frac{\overline{PH}}{\overline{PE}} \Rightarrow \overline{PH} = \overline{PE} \cdot \cos\alpha$$

Se halla el producto escalar de los vectores $\overrightarrow{PE}$ y $\vec{a}$.

$$\left|\overrightarrow{PE} \bullet \vec{a}\right| = \overline{PE} \bullet |\vec{a}| \cdot \cos\alpha = |\vec{a}| \cdot \overline{PE} \cdot \cos\alpha = |\vec{a}| \cdot \overline{PH}; \text{ y despejando } \overline{PH}:$$

$$d = \overline{PH} = \frac{\left|\overrightarrow{PE} \bullet \vec{a}\right|}{|\vec{a}|}$$

Por otro lado,

$\overrightarrow{PE} \bullet \vec{a}$ = (*x*1 - *x*0, *y*1 - *y*0)□(*A, B*) = *A*(*x*1 - *x*0) + *B*(*y*1 - *y*0) = *Ax*1 + *By*1 - *Ax*0 - *By*0, y puesto que *E*(*x*1, *y*1) es un punto de la recta, verifica la ecuación de la misma:

*Ax*1 + *By*1 + *C* = 0 □ *Ax*1 + *By*1 = - *C*

Llevando esta igualdad a la expresión de $\overrightarrow{PE} \bullet \vec{a}$.

$\overrightarrow{PE} \bullet \vec{a}$ = - *C* - *AX*0 - *By*0 = - (*AX*0 + *By*0 + *C*), de donde se deduce que

| $\overrightarrow{PE} \bullet \vec{a}$| = | - (*AX*0 + *By*0 + *C*) | = | *AX*0 + *By*0 + *C* |

Además, el módulo de $\vec{a}$ es $|\vec{a}| = \sqrt{A^2 + B^2}$. Por último, sustituyendo en la

expresión de *d*:

$$d = \overline{PH} = \frac{\left|\overrightarrow{PE} \bullet \vec{a}\right|}{\left|\vec{a}\right|} = \frac{\left|Ax_0 + By_0 + C\right|}{\sqrt{A^2 + B^2}}$$

LUGARES GEOMÉTRICOS

Se llama *lugar geométrico* a cualquier conjunto de puntos que vienen caracterizados por una cierta propiedad.

Por ejemplo, el lugar geométrico de los puntos del plano que se encuentran a distancia fija *r* de un punto señalado, *O*, es la circunferencia centrada en *O* y con radio *r* .

El lugar geométrico de los puntos del plano que se encuentran a igual distancia de dos puntos dados, es la mediatriz del segmento que los une.

El lugar geométrico de los puntos del plano que se encuentran a distancia fija de una recta, es un conjunto formado por dos rectas paralelas a la recta dada.

Ecuación de un lugar geométrico

Para hallar la ecuación de un lugar geométrico se toma un punto genérico *X* de coordenadas (*x, y*) y se intenta escribir en forma de ecuación la condición que define al lugar.

Bisectriz de un ángulo

Los puntos de la bisectriz de un ángulo equidistan de los lados del ángulo; por lo tanto, la bisectriz está contenida en el lugar geométrico de los puntos que equidistan de los lados.

Este lugar geométrico está constituido por las bisectrices de los cuatro ángulos que se forman al cortar las dos rectas. Dichas bisectrices coinciden dos a dos.

Así pues, el lugar geométrico de los puntos que equidistan de dos rectas está constituido por dos rectas que son las bisectrices de los ángulos que forman.

ECUACIONES IMPLICITAS

A.-DOS PLANOS

$\prod$ Ax+By+Cz+D=0
$\prod'$ A'x+B'y+C'z+D'=0

Ambas ecuaciones forman un sistema que podemos analizar a partir del Teorema de Rouchè.

a) h=h'=2<n Sistema compatible indeterminado , infinitas soluciones , existen infinitos puntos en común luego los **planos se cortan en una recta.**

b) h=1 h'=2 Sistema incompatible. No tiene solución. No hay puntos en común. Los **planos son paralelos.**

Condición de paralelismo→ = =
c)h=h'=1<n Sistema compatible indeterminado. Infinitas soluciones. **Planos coincidentes.**

B.-TRES PLANOS

$\prod$ Ax+By+Cz+D=0
$\prod'$ A'x+B'y+C'z+D'=0
$\prod''$ A''x+B''y+C''z+D''=0

a) h=h'=3=n Sistema compatible determinado. Única solución. **Se cortan en un punto.**
b)h=2 h'=3 Sistema incompatible. No tiene puntos en común. **Los planos se cortan 2 a 2 formando un prisma triangular, o bien 2 son paralelos y el tercero les corta.**
c)h=h'=2 <n Sistema compatible indeterminado. Infinitas soluciones . Luego **se cortan en una recta.**
d) h=1 h'=2 Sistema incompatible. No tiene solución. Los tres planos son **paralelos.**

e) h=h'=1<n Sisatema compatible indeterminado. Infinitas soluciones. Los planos **son coincidentes.**

C.-RECTA Y PLANO

$\prod$ Ax+By+Cz+D=0

a)h=h'=3=n Sistema compatible determinado. Solución única. Se **cortan en un punto.**
b)h=2 h'=3 Sistema incompatible. No tiene solución. Luego la **recta es perpendicular al plano.**
c)h=h'=2<n Sistema compatible indeterminado. Infinitas soluciones. La**recta esta contenida en el plano.**

D.-DOS RECTAS

a)h=3 h'=4 Sistema incompatible. No hay solución. Las rectas **se cruzan.**
b)h=h'=3=n Sistema compatible determinado. Solución unica. Un solo punto en común. **Se cortan.**

c)h=2 h'=3 Sistema incompatible. No hay puntos en común. Estan en el mismo plano. Luego **son paralelas.**
d)h=h'=2<n Sistema compatible indeterminado. Infinitos puntos en comun. Las rectas son **coincidentes.**

ECUACIONES VECTORIALES

A.-DOS PLANOS:
Formamos la matriz con los cuatro vectores directores y analizamos el rango.
a) r[v,w,v',w']=3 Los planos se cortan en una recta.
b) r[v,w,v',w']=2

b.1)r[a-a_1,v,w]=3 Los planos son paralelos.
b.2) r[a-a_1,v,w]=2 Los planos son coincidentes.

B.-RECTA Y PLANO
a) r [u,v,w]=3 Se cortan

b) r [u,v,w]=2

b.1) r[a-a_1,v,w]=3 La recta y el plano son paralelos.
b.2) r[a-a_1,v,w]=2 La recta esta incluida en el plano.

C.-DOS RECTAS
a) r[u,u']=2 Tienen dos posibilidades (cortarse o cruzarse)

a.1) r[a-a',u,u'] = 3 Se cruzan.
a.2) r[a-a',u,u'] = 2 Se cortan.

b) r [u,u']=1 Otras 2 posibilidades.

b.1) r[a-a',u] = 2 Las rectas son paralelas.
b.2) r[a-a',u] = 1 Las rectas son coincidentes.

ÁNGULOS

A.-FORMADO POR DOS RECTAS
B.-FORMADO POR DOS PLANOS

$\prod$ Ax+By+Cz+D=0
$\prod$' A'x+B'y+C'z+D'=0

C.-ANGULO ENTRE RECTA Y PLANO
$\prod$ Ax+By+Cz+D=0

DISTANCIAS

2. No lineal

- **Ecuaciones paramétricas y polares de las cónicas**

$x = \cos\theta$, $y = \text{sen}\,\theta$ — Ecuaciones paramétricas o representación paramétrica de la circunferencia.

$x = f(t)$, $y = g(t)$ — Ecuaciones paramétricas de la curva *C*
$y = \pm\sqrt{1 - t^2}$

Las ecuaciones paramétricas de un lugar geométrico específico no son únicas, ya que el lugar geométrico puede representarse por diferentes pares de ecuaciones.

$x = tv_0 \cos\alpha$, $y = tv_0\,\text{sen}\,\alpha - ½\,gt^2$ — Ecuaciones paramétricas de la parábola
$x = p\,\text{ctg}2\,\alpha$, $y = 2p\,\text{ctg}\,\alpha$

$x = a \sec\theta$, $y = b\,\text{sen}\,\theta$ — Ecuaciones paramétricas cónicas de la parábola

$x = a \sec\theta$, $y = b\,\text{tg}\,\theta$ — Ecuación de la representación paramétrica de la hipérbola

$x = a\,(\theta - \text{sen}\,\theta)$, $y = a\,(1 - \cos\theta)$ — Ecuaciones paramétricas de la cicloide

$$x = (a + b)\cos\theta - b\cos\frac{a+b}{b}\theta,$$

$$y = (a + b)\,\text{sen}\,\theta - b\,\text{sen}\frac{a+b}{b}\theta$$

Ecuaciones paramétricas de la epicicloide

$x = 2a\cos\theta - a\cos 2\theta$,
$y = 2a\,\text{sen}\,\theta - a\,\text{sen}\,2\theta$ — Ecuaciones paramétricas de la cardioide

$$x = (a - b)\cos\theta + b\cos\frac{a-b}{b}\theta,$$ — Ecuaciones paramétricas de la

$$\frac{a-b}{b}$$

$y = (a - b)\ \text{sen}\ \theta - b\ \text{sen}\ \theta$	hipocicloide
$x = a \cos^3 \theta \quad , \qquad y = a\ \text{sen}^3 \theta$	Ecuaciones paramétricas de la astroide
$x = r \cos \theta \quad , \qquad y = r\ \text{sen}\ \theta$	Ecuación polar del lugar geométrico
$p = r \cos (\theta - w)$	Ecuación polar de la recta
$a^2 = r^2 - 2cr\ \cos (\theta - \alpha) + c^2$	Ecuación polar de una circunferencia de centro el punto (c, a) y radio igual a a
$r = a$	Ecuación polar si centro está en el polo
$r = \pm\ 2a \cos \theta$	Ecuación polar si circunferencia pasa por el polo y el centro está sobre el eje polar
$r = \pm\ 2a\ \text{sen}\ \theta$	Ecuación polar si circunferencia pasa por el polo y el centro está sobre el eje a 90°

Se debe tomar el signo positivo o negativo
Según el centro esté arriba o abajo del polo

$r = \dfrac{ep}{1 - e \cos \theta}$	Ecuación de la cónica con directriz a la izquierda del polo
$r = \dfrac{ep}{1 + e \cos \theta}$	Ecuación de la cónica con directriz a la derecha del polo y a p unidades de él

• Características de una superficie cuadrática con ejes paralelos a los coordenados, a partir de su ecuación

Ecuación cuadrática y forma cuadrática

i. Una ecuación cuadrática con dos variables sin términos lineales es una ecuación de la forma

$$ax^2 + bxy + cy^2 = d$$

donde $|a| + |b| + |c| \neq 0$. Es decir, por lo menos uno de los números a, b o c es diferente de cero.

ii. Una forma cuadrática con dos variables es una expresión de la forma

$$F(x, y) = ax^2 + bxy + cy^2$$

donde $|a| + |b| + |c| \neq 0$.

Teorema de los ejes principales en $\Re^2$ Sea $ax^2 + bxy + cy^2 = d^{(*)}$ una ecuación cuadrática en las variables x y y. Entonces existe un solo número θ en $[0, 2\pi]$ tal que la ecuación $^{(*)}$ se puede escribir en la forma

$$a' x'^2 + c' y'^2 = d$$

donde x', y' son los ejes que se obtienen al rotar los ejes x y y un ángulo θ en sentido antihorario. Por otra parte, los números a' y c' son los valores

característicos de la matriz $A = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$. Los ejes x' y x' reciben el nombre de

ejes principales de la gráfica de la ecuación cuadrática ***.**

Si $A = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$, entonces la ecuación cuadrática

$ax^2 + bxy + cy^2 = d$ con $d \neq 0$ es la ecuación de:

i. Una hipérbola si det $A < 0$.
ii. Una elipse, circunferencia o sección cónica degenerada si det $A > 0$.
iii. Un par de líneas rectas o una sección cónica degenerada si det $A = 0$.
iv. Si $d = 0$, entonces la ecuación $ax^2 + bxy + cy^2 = d$ es la de dos líneas rectas si det $A \neq 0$, y es la ecuación de una sola recta si det $A = 0$.

Ecuaciones estándar:

Circunferencia: x2 + y2 = r2

Elipse: $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$

$$\text{Hipérbola:}\quad \begin{cases} \dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1 \\ \text{o} \\ \dfrac{y^2}{b^2} - \dfrac{x^2}{a^2} = 1 \end{cases}$$

A la ecuación (*) se le llama sección **cónica degenerada.**

Hay una gran variedad de superficies tridimensionales de la forma $A\mathrm{v} \bullet \mathrm{v} = d$, siendo $\mathrm{v} \in \Re^2$. A dichas superficies se les llama **superficies cuadráticas.** Las formas cuadráticas pueden definirse con un número cualquiera de variables.

Forma cuadrática Sean $\mathrm{v} = \begin{pmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \end{pmatrix}$ A una matriz simétrica de $n \times n$.

Entonces una **forma cuadrática** en $x_1, x_2, \ldots, x_n$ es una expresión de la forma

$$F(x_1, x_2, \ldots, x_n) = A\mathrm{v} \bullet \mathrm{v}$$

B. Matemáticas aplicadas

I. Probabilidad

1. Fundamentos de la teoría de la probabilidad

- **Probabilidad clásica y probabilidad frecuentista**

Definición clásica de probabilidad.- Consideremos un experimento aleatorio con espacio muestral finito S = {x1, x2, ... , xn} en el que los eventos elementales son igualmente probables. En tal caso, la probabilidad de cualquier evento $A \subset S$ se define como

$$P(A) = \frac{\#A}{\#S} = \frac{\text{No. de casos favorables}}{\text{No. de casos posibles}}$$

Ejemplos:

1. Hay 15 bolas en una bolsa. Todas son del mismo tamaño. Hay 4 rojas, 6 blancas y 5 azules. Encontrar la probabilidad de sacar de la bolsa una bola que sea:
   P(roja) = 4/15

   P(azul) = 5/15 = 1/3

   P(no blanca) = 9/15 = 3/15

2. Martha y cuatro de sus amigos están entre 40 candidatos para visitar Washington, D. C. Se elegirá al azar a un estudiante. ¿Cuál es la probabilidad de que Martha o uno de sus amigos sea elegido?
   P(A) = 1/40 , P(B)= 4/40

   P(A o B) = P(A) + P(B)   Eventos mutuamente exclusivos

   = 1/40 + 4/40
   = 5/40 = 1/8

3. Lisa y su hermano están entre 12 muchachas y 18 muchachos nominados para ocupar puestos en la banda. Se elegirá al azar a una muchacha y a un muchacho para esos puestos. ¿Cuál es la probabilidad de que Lisa o su hermano sean elegidos?
   P(A) = 1/12 , P(B) = 1/18

   $P(A) + P(B) = 1/12 + 1/18 = \frac{3 + 2}{36}$

   = 5/36

Definición frecuencial de probabilidad.- Sea nA el número de ocurrencias del evento A en n repeticiones de un experimento aleatorio. El cociente $\frac{n_A}{n}$ recibe el nombre de frecuencia relativa del evento A y la probabilidad de A se define como el límite:

$$\lim_{n \to \infty} \frac{n_A}{n}$$

P (A) =

Observación.
Como en la práctica es imposible repetir un experimento una infinidad de veces, lo que se hace es repetirlo un número suficientemente grande de veces y tomar las frecuencias relativas como aproximaciones de las probabilidades reales. Así se obtienen las probabilidades en la práctica.

Propiedades de la probabilidad clásica

1. $P(S) = 1$
2. $0 \leq P(A) \leq 1$, para cualquier evento A
3. Si $AB = \varnothing$, entonces $P(A \cup B) = P(A) + P(B)$
4. $P(\varnothing) = 0$
5. $P(AB^C) = P(A) - P(AB)$
6. Si $A \subset B$, entonces $P(A) \leq P(B)$
7. $P(A \cup B) = P(A) + P(B) - P(AB)$
8. $P(A^C) = 1 - P(A)$
9. $P \left\{ \begin{array}{l} \text{ocurra exactamente} \\ \text{uno de los eventos A} \\ \text{y B} \end{array} \right\} = P(A) + P(B) - 2P(AB)$
10. $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(BC)$
11. $P \left\{ \begin{array}{l} \text{ocurra exactamente} \\ \text{uno de los eventos} \\ \text{A, B y C} \end{array} \right\} = P(A) + P(B) + P(C) - 2[P(AB) + P(AC) - P(BC) + 3P(ABC)$
12. $P \left\{ \begin{array}{l} \text{ocurran exactamente} \\ \text{dos de los eventos} \\ \text{A, B y C} \end{array} \right\} = P(AB) + P(AC) + P(BC) - 3P(ABC)$
13. $P \left\{ \begin{array}{l} \text{ocurra exactamente} \\ \text{uno de los eventos} \\ \text{A, B y C} \end{array} \right\} = P(A) + P(B) + P(C) - 2[P(AB) + P(AC) + P(BC) + 3P(ABC)]$

* $P(AB) = P(\varnothing) = 0$, por ser A y B eventos mutuamente excluyentes.

• Problemas que requieren de los axiomas y teoremas fundamentales de la probabilidad

Ejemplo:

Un estudiante toma dos cursos: inglés y matemáticas. Si la probabilidad de que apruebe al menos un curso es 0.8, de que apruebe ambos 0.3, y de que repruebe matemáticas 0.6, determine la probabilidad de que el estudiante

Datos:

A = el estudiante aprueba inglés

B = el estudiante aprueba matemáticas

$P\ (A \cup B) = 0.8\ , \qquad P\ (AB) = 0.3\ , \qquad P\ (B^C) = 0.6$

(a) apruebe matemáticas

Buscamos P (B)

$$\begin{aligned} P\ (B) &= 1 - (B^C) \\ &= 1 - 0.6 = 0.4 \end{aligned}$$

(b) apruebe inglés

Buscamos P (A)

$$P\ (A \cup B) = P\ (A) + P\ (B) - P\ (AB)$$

$$\begin{aligned} P\ (A) &= P\ (A \cup B) - P\ (B) + P\ (AB) \\ &= 0.8 - 0.4 + 0.3 \\ &= 0.7 \end{aligned}$$

(c) repruebe ambos cursos

$$\begin{aligned} P\ (A^C B^C) &= P\ ((A \cup B)^C) \\ &= 1 - P\ (A \cup B) \\ &= 1 - 0.8 \\ &= 0.2 \end{aligned}$$

(d) solo apruebe matemáticas

$$\begin{aligned} P\ (A^C B) &= P\ (B) - P\ (AB) \\ &= 0.4 - 0.3 \\ &= 0.1 \end{aligned}$$

(e) solo apruebe inglés

$$\begin{aligned} P\ (AB^C) &= P\ (A) - P\ (AB) \\ &= 0.7 - 0.3 \\ &= 0.4 \end{aligned}$$

(f) apruebe exactamente uno de los dos cursos

$$\left\{ \begin{array}{l} \text{ocurra exactamente} \\ \text{uno de los eventos} \\ \text{A y B} \end{array} \right\}$$

$$P = P(A) + P(B) - 2P(AB) = 0.7 + 0.4 - 2(0.3) = 0.7 + 0.4 - 0.6 = 0.5$$

Definición axiomática de probabilidad.- La probabilidad es una función P definida para una clase de subconjuntos (eventos) de un conjunto S, con las siguientes propiedades:

1. $0 \le P(A) \le 1$, para cualquier evento $A \subset S$
2. $P(S) = 1$
3. Si a1, A2, . . . es una colección de eventos mutuamente excluyentes, entonces

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sigma_{n=1}^{\infty} (A_n)$$

Definición probabilidad condicional.- Sean A y B dos eventos con $P(B) > 0$. La probabilidad condicional de A dado B, denotada por $P(A|B)$, se define como

$$P(A|B) = \frac{P(AB)}{P(B)}$$

$P(A|B)$ no esta definida cuando $P(B) = 0$.

Teorema 1. (Principio fundamental de conteo). Si cierto experimento E_1 puede ocurrir de n formas y, correspondiendo a cada una de estas formas, un segundo experimento E_2 puede suceder de m modos, entonces el número de maneras diferentes en que ambos pueden ocurrir es igual a mn.

Aún cuando el principio fundamental de conteo sólo se enuncio para dos experimentos, puede extenderse a cualquier número de ellos, si por cada ocurrencia de los experimentos E_1, E_2, . . . , E_{k-1} el experimento E_k se puede representar de n_k Formas, entonces el número de maneras diferentes en que todos ellos pueden ocurrir simultáneamente es $n_1 n_2 \ldots n_k$.

Ejemplos:

(a) Al lanzar tres veces un dado se pueden obtener (6)(6)(6) = 216 posibles resultados.

(b) Si en un Estado las placas de los automóviles constan de dos primeras letras distintas seguidas de cuatro dígitos, el primero de los cuales no puede ser cero, entonces en ese Estado se podrán imprimir (26)(25)(9)(10)(10)(10) = 5,850,000 diferentes placas.

(c) El número de maneras en que un estudiante puede contestar un examen del tipo falso-verdadero de 10 preguntas es $2^{10} = 1{,}024$

Teorema 2. Los números de muestras de tamaño n que pueden Seleccionarse de una población de *M* objetos en cada uno de los diferentes tipos de muestreo están dados por las siguientes fórmulas:

(i) M^n si el muestreo es con orden y con reemplazo.

(ii) $(M)_n$ si el muestreo es con orden y sin reemplazo.

(iii) $C_n^M = \dfrac{M!}{n!(M-n)!}$ si el muestreo es sin orden y sin reemplazo

(iv) $C_n^{M+n-1} = \dfrac{(M+n-1)!}{n!\,(M-1)!}$ si el muestreo es sin orden y con reemplazo.

Ejemplos:

(a) ¿De cuántas maneras puede un grupo de 30 estudiantes elegir un comité de 4?
Aquí el orden en que se lleva a cabo la elección del comité es Irrelevante, además de que los integrantes deben ser 4 personas distintas. Tenemos así un caso de muestreo sin orden y sin reemplazo en el que *M* = 30 y *n* = 4. El número de posibles comités es por lo tanto de

$$C_4^{30} = \frac{30!}{4!26!} = 27{,}405$$

(b) ¿De cuántas maneras puede un grupo de 30 estudiantes elegir un presidente, un vicepresidente, un secretario y un tesorero?
Este es un caso de muestreo con orden y sin reemplazo, pues ahora sí interesa el orden de la elección. El primer miembro elegido deberá cumplir una función específica, digamos la de presidente; el segundo la de vicepresidente; etc. De este modo, el número de formas en que puede hacerse la elección es de $(30)_4 = 30\bullet29\bullet28\bullet27 = 657{,}720$

(c) A un grupo de 30 estudiantes su profesor les plantea 4 preguntas. ¿De cuántas formas se podrían obtener las respuestas de los alumnos?
En esta situación, cada estudiante podría contestar más de una pregunta, además de que se debe distinguir quién de los estudiantes

contesto la primera pregunta, quién la segunda, etc. Con todo esto, el problema puede interpretarse como un muestreo con orden y con reemplazo con M = 30 y n = 4. Así, el número de formas en que se podría dar respuesta a las cuatro preguntas es (30)4 = 810,000

(d) Se van a comparar los efectos de dos medicamentos A y B y una tableta en un estudio farmacéutico en el que participan 100 personas. A 60 se les suministra el medicamento A, a 15 el medicamento B y a las restantes la tableta. ¿De cuántas formas diferentes pueden distribuirse los medicamentos y las tabletas?

El problema puede interpretarse como la realización combinada de tres experimentos. Primero se seleccionan 60 personas de las 100 para suministrarles el medicamento A y luego se eligen 15 de las restantes para darles el medicamento B y a las últimas 25 se les da la tableta. Estos experimentos pueden efectuarse individualmente de C_{60}^{100}, C_{15}^{100} y C_{25}^{25} formas, respectivamente. Finalmente, por el principio fundamental de conteo se tiene que los tres experimentos pueden realizarse de

$$C_{60}^{100}\ C_{15}^{40}\ C_{25}^{25}$$

maneras.

Teorema 3. Consideremos un experimento aleatorio con espacio muestral S. Sea $B \subset S$ un evento con P (B) > 0. Entonces P (•|B) es una función de probabilidades, es decir;

(1) P (S|B) = 1

(2) $0 \leq P(A|B) \leq 1$, para todo $A \subset S$.

(3) Si A1, A2, . . . son eventos mutuamente excluyentes, entonces $P(A_1, \cup A_2 \cup \ldots \mid B) = P(A_1|B) + P(A_2|B) + \ldots$

P (•|B) satisface los axiomas de Kolmogorov. Esto permite pasar a probabilidad condicional las fórmulas de la probabilidad clásica:

(1) $P(\varnothing|B) = 0$

(2) $P(A_1 \cup A_2|B) = P(A_1|B) + P(A_2|B) - P(A_1A_2|B)$

(3) $P(A^C|B) = 1 - P(A|B)$

(4) $P(A_1A_2^C|B) = P(A_1|B) - P(A_1A_2|B)$

(5) Si $A_1 \subset A_2$, entonces $P(A_1|B) \leq P(A_2|B)$

(6) $P\left(\begin{array}{l}\text{exactamente un}\\ \text{evento } A_1 \text{ o } A_2,\\ \text{dado B}\end{array}\right) = P(A_1|B) + P(A_2|B) - 2\,P(A_1A_2|B)$

Ejemplos:

1. Sean A y B dos eventos para los cuales P (A) = 0.5, P (B) = 0.7 y P (A|B) = 0.4. Encuentre:
   (a) P (AB)

   = P (B) P (A|B) = (0.7)(0.4) = 0.28

   (b) P (A $\cup$ B)

   = P (A) + P (B) – P (AB) = 0.5 + 0.7 – 0.28 = 0.92

   (c) P (AB^C) = P (A) – P (AB) = 0.5 – 0.28 = 0.22

2. Sean A y B dos eventos con P (AB^C) = 0.4 y P (B) = 0.2. Encuentre:

   (a) P ($A|B^C$)

   $$= \frac{P(AB^C)}{P(B^C)} = \frac{0.4}{0.8} = 0.5$$

   (b) P ($A^C|B^C$)
   = 1 – P ($A|B^C$) = 1 – 0.5 = 0.5

   (c) P (A^CB^C)
   P (B^C) P ($A^C|B^C$) = (0.8)(0.5) = 0.4

   (d) P (A $\cup$ B)
   = 1 – P $(A \cup B)^C$= 1 – P (A^CB^C) = 1 – 0.4 = 0.6

Regla del producto de probabilidades
Fórmulas para obtener la probabilidad de la intersección de dos eventos.
P (AB) = P (B) P (A|B) (P (B) > 0)

$$\left.\begin{array}{ll} P(AB) = P(A)\,P(A|B) & (P(A) > 0) \\ P(ABC) = P(A)\,P(B|A)\,P(C|AB) & \\ P(ABCD) = P(A)\,P(B|A)\,P(C|AB)\,P(D|ABC) & \end{array}\right\}$$

Casos especiales para 2, 3 y 4 eventos, respectivamente de la regla o fórmula del producto de probabilidades.

Teorema 4. (Regla del producto de probabilidades). Sean $A_1, A_2, \ldots, A_n$ n eventos para los cuales P ($A_1, A_2, \ldots, A_{n-1} > 0$). Entonces
$P(A_1, A_2 \ldots A_n) = P(A_1)\, P(A_2|A_1)\, P(A_3|A_1A_2) \ldots\ P(A_n|A_1A_2 \ldots A_{n-1})$

Ejemplo:
En cierta facultad se ha determinado, de experiencias pasadas, que la probabilidad de que un egresado apruebe su examen profesional en el primer intento es 0.7; que lo apruebe en el segundo intento es 0.8

(naturalmente, dado que no aprobó en su primera oportunidad) y la probabilidad de que apruebe en la tercera oportunidad es 0.9 (por supuesto, dado que falló en sus dos primeras oportunidades).

Datos:

A_1 = el egresado aprueba su examen profesional en la primera oportunidad.

A_2 = el egresado aprueba su examen profesional en la segunda oportunidad.

A_3 = el egresado aprueba su examen profesional en la tercera oportunidad.

Tenemos que $P(A_1) = 0.7$, $P(A_2|A_1^C) = 0.8$ y $P(A_3|A_1^C A_2^C) = 0.9$

(a) ¿Cuál es la probabilidad de que un egresado de tal facultad apruebe su examen profesional hasta la tercera oportunidad?

Buscamos $P(A_1^C A_2^C A_3)$. Aplicando la regla del producto de propiedades tenemos

$$P(A_1^C A_2^C A_3) = P(A_1^C)\, P(A_2^C|A_1^C)\, P(A_3|A_1^C A_2^C)$$

$$= (0.3)(0.2)(0.9)$$
$$= 0.054$$

(b) Si en dicha facultad sólo se conceden tres oportunidades, ¿cuál es la probabilidad de que un egresado se titule?

El evento “el egresado se titulo” puede expresarse en términos de A_1, A_2 y A_3 en la forma $A_1 \cup A_1^C A_2 \cup A_1^C A_2^C A_3$. Así la probabilidad pedida es

$$P(A_1 \cup A_1^C A_2 \cup A_1^C A_2^C A_3) = P(A_1) + P(A_1^C A_2) + P(A_1^C A_2^C A_3)$$

$$= P(A_1) + P(A_1^C)\; P(A_2|A_1^C) + P(A_1^C A_2^C A_3)$$

$$= 0.7 + (0.3)(0.8) + (0.054)$$
$$= 0.954$$

Teorema 5. Probabilidad total. Supongamos que los eventos $B_1, B_2, \ldots, B_n$ forman una partición del espacio muestral S, y que cada uno de ellos tiene probabilidad positiva. Entonces, para cualquier evento $A \subset S$ se tiene que

$$P(A) = P(B_1)\, P(A|B_1) + \ldots P(B_n)\, P(A|B)$$

$$= \sum_{i=1}^{n} P(B_i)\, P(A|B_i)$$

Teorema 6. Regla o fórmula de Bayes. Si los eventos $B_1, B_2, \ldots, B_n$ forman una partición del espacio muestral S y si cada uno de ellos tiene probabilidad positiva, entonces para cualquier evento A con probabilidad positiva se tiene que

$$P(B_i|A) = \frac{P(B_i)\,P(A|B_i)}{P(B_i)\,P(A|B_1) + \ldots P(B_n)\,P(A|B_n)}$$

$$= \frac{P(B_i)\,P(A|B_i)}{\sum_{i=1}^{n} P(B_i)\,P(A|B_i)}$$

Ejemplo:

En un telégrafo inalámbrico los mensajes se transmiten usando señales de punto y raya. De experiencias pasadas se sabe que una señal que se origina como raya, tiene probabilidad de 0.02 de ser recibida como punto; y una señal que se origina como punto tiene probabilidad de 0.01 de ser recibida como raya. Si se sabe, además, que el 60% de las señales transmitidas son raya, encuentre

Datos:

B_1 = la señal es transmitida como raya.
B_2 = la señal es transmitida como punto.
A = la señal es recibida como raya.

$P(A^C|B_1) = 0.02$, $P(A|B_2) = 0.01$, $P(B_1) = 0.6$ y $P(B_2) = 0.4$. Podemos escribir, adicionalmente $P(A|B_1) = 0.98$ y $P(A^C|B_2) = 0.99$.

(a) la probabilidad de que una señal se reciba como raya.

Buscamos $P(A) = P(B_1)\,P(A|B_1) + P(B_2)\,P(A|B_2)$
$= (0.6)(0.98) + (0.4)(0.01)$
$= 0.592$

(b) la probabilidad de que una señal se haya transmitido como raya si se recibió como raya.

Se desea calcular $P(B_1|A)$. Aplicando la fórmula de Bayes se obtiene,

$$P(B_1|A) = \frac{P(B_1)\,P(A|B_1)}{P(A)}$$

$$= \frac{(0.6)(0.98)}{0.592}$$

$$= 0.9932$$

2. <u>Distribuciones</u>

• Problemas aplicando la función de distribución binomial y normal

Definición.- Distribución Binomial o ensayo de Bernoulli. Un ensayo de Bernoulli es cualquier experimento aleatorio que admite "solamente" dos posibles resultados. Uno de ellos se llama éxito (*e*) y el otro fracaso (f). Se denota con p a la probabilidad de éxito y mediante q a la probabilidad de fracaso:

$$p = P(e)$$
$$q = P(f) = 1 - p$$

Características de la distribución binomial

1. El problema consiste en repetir n veces un mismo ensayo de Bernoulli.
2. La probabilidad de éxito p es la misma en cada ensayo.
3. Interesa conocer el número de éxitos o de fracasos en los n ensayos.

Definición.- Si en la descripción anterior *x* representa el número de éxitos logrados en los n ensayos, *x* recibe el nombre de variable aleatoria binomial de parámetros n y p.

Teorema. Para una variable aleatoria binomial de parámetro n y p se tiene que

(a) $p_i = C^n_i \, p^i \, q^{n-i}$, i = 0, 1, 2, . . . , n

(b) $E[X] = np$

(c) $\mathrm{Var}\,[X] = npq$

Ejemplo:

Si un tirador tiene la probabilidad de 0.85 de dar en el blanco, Encuentre la probabilidad de que en 8 disparos

Datos:

El ensayo de Bernoulli que se repite es ejecutar un disparo. Si e = acertar y f = fallar, tendremos los siguientes datos:

n = 8, p = 0.85, q = 0.15 y

X = número de aciertos logrados en los 8 disparos.

La función de probabilidad es entonces

$$P_i = C^8_i \,(0.85)^i \,(0.15)^{8-i}, \qquad i = 0, 1, 2, \ldots, 8.$$

(a) acierte exactamente 3 veces.

Buscamos $p_3 = C_3^8\,(0.85)^3\,(0.15)^5$

(b) acierte al menos una vez
Deseamos calcular $P\,(X \geq 1)$

$P\,(X \geq 1) = 1 - P\,(X = 0) = 1 - (0.15)^8$

(c) acierte a lo más 4 veces
Deseamos calcular $P\,(X \leq 4)$
$P\,(X \leq 4) = p_0 + p_1 + p_2 + p_3 + p_4 = 0.014$

(d) falle en 6 de los 8 disparos

Esto corresponde a $p_2 = 2.8 \times 10^{-6}$

Distribución normal o de Laplace Gauss

Siempre que el valor de una variable sea resultado de la intervención de muchos fenómenos, independientes entre sí, cuyos efectos se suman, se obtiene una curva de distribución para dicha variable de forma acampanada. A dicha curva se le denomina normal, y a la distribución teórica correspondiente, distribución normal o de Laplace-Gauss.

Ley teórica de la distribución normal.- Se trata de una ley de distribución definida para variables continuas, cuya función densidad de probabilidad, p (x), es:

$$y = p\,(x) = \frac{1}{\sigma\sqrt{2\pi}}\; e^{-1/2\left(\frac{x - \mu}{\sigma}\right)^2}$$

En una distribución normal, el área bajo toda la curva vale 1. (El área bajo la curva representa el valor de la probabilidad que, para toda la curva, es la certeza). En consecuencia, el área bajo cada mitad vale 0.5.

Fórmula de la distribución normal estándar.

$$Z = \frac{x - \mu}{\sigma}$$

Ejemplo:
Calificación (0 - 100) N (70, 9)
$\bar{x}$ S^2

$$\downarrow \; \downarrow$$
$$\mu \;\; \sigma^2$$

(a) P ($70 \leq x \leq 76$)

$$Z = \frac{76 - 70}{3} = \frac{6}{3} = 2$$

El valor en tablas de 2 es 0.9772 entonces

0.9772 – 0.5 = 0.4772

(b) P ($73 \leq x \leq 76$)

$$Z_1 = \frac{73 - 70}{3} = \frac{3}{3} = 1$$

$$Z_2 = \frac{76 - 70}{3} = 2$$

0.9772 – 0.8413 = 0.1359

(c) P ($x \geq 76$)

$$Z = \frac{76 - 70}{3} = 2$$

1 – 0.9772 = 0.0228

- **Problemas aplicando la función de distribución geométrica y Poisson**

Definición.- Distribución geométrica.- La distribución geométrica, al igual que la binomial, es producto de repeticiones de ensayos de Bernoulli. Esta distribución se aplica en situaciones como las siguientes: se ejecutan tiros libres a una canasta hasta lograr el primer acierto; se lanza una moneda sucesivamente hasta observar un águila; etc.

Características

(1) Se repite un ensayo de Bernoulli hasta obtener por primera vez un éxito.
(2) La probabilidad de éxito, *p*, es la misma en cada ensayo.
(3) Interesa conocer el número de ensayos necesarios para obtener por primera vez un éxito.

Definición.- Si en la descripción anterior *X* denota el número de ensayos necesarios para lograr el primer éxito, X recibe el nombre de variable aleatoria geométrica de parámetro p.

Para una variable aleatoria geométrica de parámetro p se tiene lo siguiente:
(1) $p_k = pq^{k-1}$, $k = 1, 2, 3, \ldots$

(2) $E[X]$ = ^{1}
(3) Var $[X]$ $\frac{q}{p^2}$

(4) $P(X>K) = q^k$

Ejemplo:

Un matrimonio decide tener hijos hasta que tengan un varón en la familia. Si se reporta que el 48% de los recién nacidos son varones, encuentre

Datos:

El ensayo de Bernoulli que se está repitiendo es el de tener un hijo. Si llamamos éxito al hecho de que un hijo sea varón, entonces $p = 0.48$. Sea

X = el número de hijos que debe tener el matrimonio para lograr el primer varón.

La función de probabilidad de la variable aleatoria geométrica X es

$$P_k = (0.48)(0.52)^{k-1}, \quad k = 1, 2, 3, \ldots$$

(a) la probabilidad de que hasta el tercer hijo logren el varón.

Se desea determinar p_3
$P_3 = (0.48)(0.52)^2 = 0.129$

(b) la probabilidad de que la familia necesite más de 4 hijos.

Buscamos $P(X > 4)$
$P(X > 4) = (0.52)^4 = 0.072$

(c) el número promedio de nacimientos que requerirá el matrimonio.

Esto corresponde a $E[X]$.

$$[X] = \frac{1}{p} = \frac{1}{0.48} = 0.283$$

(d) la probabilidad de que el matrimonio requiera a lo más 6 hijos.

Se está considerando P $(X \leq 6)$.
P $(X \leq 6) = 1 - P(X > 6) = 1 - (0.52)^6 = 0.98$

Definición.- Distribución Poisson. Se dice que un fenómeno aleatorio es un proceso de Poisson si tiene las siguientes características:

(1) El proceso no es hereditario. Esto significa que el número de ocurrencias en un subintervalo dado no tiene efecto alguno sobre el número de ocurrencias del fenómeno en otro subintervalo no transpuesto con el primero.

(2) El proceso es estacionario. Esto significa que el número de ocurrencias del fenómeno en un subintervalo dado sólo depende de su longitud, y no de su posición.

(3) El proceso es ordinario. Esto es, la probabilidad de dos o más ocurrencias del fenómeno en intervalos suficientemente cortos en casi cero.

Definición.- Si en un proceso de Poisson x denota el número de ocurrencias del fenómeno en el "tiempo" de observación, x recibe el nombre de variable aleatoria de Poisson de parámetro λ.

λ = número de ocurrencias del fenómeno en el "tiempo" de observación

Teorema. Para una variable aleatoria de Poisson de parámetro λ se tiene lo siguiente:

(a) $P_k = \dfrac{\lambda^k}{k!} e^{-\lambda}$, $\quad k = 0, 1, 2, \ldots$

(b) $E[X] = \lambda$

(c) $\mathrm{Var}[X] = \lambda$

Ejemplo:

El número de llamadas que llegan al conmutador telefónico de una empresa es un proceso de Poisson con λ = 180 llamadas por hora.

(a) Encontrar la distribución del número de llamadas que llegan en un periodo de un minuto.

Como en promedio se tiene 180 llamadas por hora, entonces

λ = 3 llamadas en promedio por minuto.

Sea

X = número de llamadas recibidas en un minuto.

La distribución de X es

$$p_k = P\,(X = k) = \frac{3^k}{k!}\;e^{-3}, \qquad k = 0, 1, 2, \ldots$$

(b) ¿Cuántas llamadas se espera recibir en 25 minutos?

Sea

X = número de llamadas que llegan al conmutador en 25 minutos.

Buscamos $E[X]$. Puesto que $E[X]$ y

$$\lambda = \frac{(180)(25)}{60} = 75, \text{ entonces}$$

$E[X] = 75$

(c) ¿Cuál es la probabilidad de que en un periodo de un minuto

(i) no se reciban llamadas?
Se desea $p_0 = e^{-3} = 0.049$

(ii) se reciban entre 2 y 5 llamadas?
Se pide $p_2 + p_3 + p_4 + p_5 = 0.7168$

(iii) no se reciban llamadas y en el siguiente minuto se reciba al menos una?
Puesto que los intervalos de tiempo que estamos tomando no se traslapan, el número de llamadas en cada uno de ellos es independiente del otro. Además, como la probabilidad de no recibir llamadas en un minuto es de 0.049, y de recibir al menos una es de 0.951, entonces la probabilidad que buscamos es

$$(0.049)(0.951) = (0.046)$$

II. Estadística

Estadística descriptiva

MUESTRA: Para obtener información sobre una población hay que limitarse a analizar sólo un subconjunto de individuos de la misma. A éste subconjunto se le denomina <u>muestra</u>.

POBLACION: Es todo conjunto de elementos, finito o infinito, definido por una o más características, de las que gozan todos los elementos que lo componen, y sólo ellos.

Métodos de muestreo : para no tener que trabajar con toda la población se utiliza el muestreo . Puede ser :

- Muestreo no probabilístico : no se usa el azar , sino el criterio del investigador , suele presentar grandes sesgos y es poco fiable .
- Muestreo probabilístico : se utilizan las leyes del azar . Puede ser :

1. **Muestreo aleatorio simple** (es el más importante) : cada elemento de la población tiene la misma probabilidad de ser elegido , las observaciones se realizan con reemplazamiento , de manera que la población es identica en todas las extracciones , o sea , que la selección de un individuo no debe afectar a la probabilidad de que sea seleccionado otro cualquiera aunque ello comporte que algún individuoo pueda ser elegido más de una vez . ("se hacen tantas papeletas numeradas como indivuos hay , se coge una y se devuelve , se vuelve a coger otra y se devuelve , etc")
2. **Muestreo sistemático** : es cuando los elementos de la población están ordenados por listas . Se elige un individuo al azar y a continuación a intervalos constantes se eligen todos los demás hasta completar la muestra . Si el oreden de los elementos es tal que los individuos próximos tienden a ser más semejantes que los alejados , el muestreo sistemático tiende a ser más preciso que el aleatorio simple , al cubrir más homogeneamente toda la población .
3. **Muestreo estratificado** : es cuando nos interesa que la muestra tenga la misma composición a la de la población la cual se divide en clases o estratos . Si por ejemplo en la población el 20% son mujeres y el 80% hombres , se mantendrá la misma proporción en la muestra .

Distribuciones de muestreo : al obtener conclusiones de la muestra y las comparamos con las de la población puede que se aproximen o no . No obstante , las medias muestrales se comportan estadísticamente bien y siguen leyes perfectamente previsibles , esto nos permitirá hacer inferencias precisas a partir de ellas , incluso determinar el riesgo que asumimos al hacerlas .
Si una población está formada por N elementos , el nº de muestras diferentes de tamaño n que se pueden obtener , si se pueden repetir los elementos (m.a.s.) sería :
$VR_{N,n}=N^n$

Distribución de medias muestrales : aunque al tomar una muestra no podemos estar seguros de que los parámetros obtenidos sean buenos estimadores de los parámetros poblacionales si se puede afirmar que :

1. La media de las medias muestrales es igual a la media real de la población es decir :
2. La desviación típica de las medias muestrales vale :

Esto significa que la distribución de medias muestrales de tamaño n extraidas de una población (normal o no normal) se distribuye según una N( )

Ejemplo : Supongamos que tenemos los elementos 2,4,6,8 .

En esta población vamos a tomar todas las muestras posibles de tamaño 2 :

Elementos e_1 e_2	Media de la muestra
2 2	2
2 4	3
2 6	4
2 8	5
4 2	3
4 4	4
4 6	5
4 8	6
6 2	4
6 4	5

6	6	6
6	8	7
8	2	5
8	4	6
8	6	7
8	8	8

La media de las medias muestrales será :
La varianza de las medias muestrales será :
Por lo tanto y

Ejemplo : el peso de los recién nacidos en una maternidad se ha distribuido según una ley normal de media =3100gr y de desviación típica =150gr ¿ Cuál será la probabilidad de que la media de una muestra de 100 recién nacidos se superior a 3130gr ?
La distribución muestral sigue una N(3100,15) por lo que p p = 1 - 0'9772 = 0'0228
<u>por lo tanto solo un 2'28% de las muestras tendrá una media por encima de los 3130gr</u>

Intervalos de probabilidad : inferencia estadística ;Como la distribución de medias muestrales es se tendrá por ejemplo que :
Esto significa que por ejemplo el 68'26% de las muestras de tamaño n extraidas de una población de media tendrán una media perteneciente al intervalo
En general el 100·(1-)% de las muestras de tamaño n tendrán una media comprendida entre : siendo el valor de la probabilidad que queda a cada lado del intervalo . O lo que es lo mismo :
(nivel de confianza)
Así por ejemplo si =0'05 entonces el 95% de las muestras tendrán una media comprendida entre =

Sin embargo lo normal será que se desconozca la media y la desviación típica de la población y que mediante técnicas de muestreo se busque estimarlas con la fiabilidad necesaria .
Por lo tanto si nos hacen la pregunta de otra forma (¿ cuál es la probabilidad de que la media <u>poblacional</u> se encuentre entre ...?) podremos transformar la desigualdad obteniendo :
(nivel de confianza)
A este intervalo se le llama **intervalo de confianza** para la media poblacional . A lo que está fuera del intervalo se le llama **regíon crítica** .
Al valor se le llama **nivel de confianza** .
Al valor se le llama **nivel de significación** .

Por lo tanto el nuevo dibujo sería :

Por lo tanto podemos afirmar que en ese intervalo tenemos una probabilidad del 95
% de que está la media poblacional .

Como ya hemos dicho lo normal será que se desconozca la desviación , por lo que debemos sustituir por s_{n-1}= donde s_{n-1}^2 es la cuasivarianza muestral .
La relación entre la varianza muestral y la cuasivarianza muestral es :
Aunque para valores grandes de n (mayores de 30) coinciden aproximadamente la cuasivarianza y la varianza por lo que se puede sustituir por s .

Por lo tanto :

Para n grandes :

Distribución para proporciones : cuando se trata de determinar la proporción de una población que posee un cierto atributo (hombre/mujer , video/no video , éxito/fracaso , etc) su estudio es

equiparable al de una distribución binomial . Así pues si tomamos muestras aleatorias de tamaño n , la media y la desviación típica de las medias muestrales será :
Esta distribución es aproximadamente normal para valores grandes de n (mayor de 30) en consecuencia puede estudiarse como una N
Si hablamos de intervalos de probabilidad entonces :
(nivel de confianza)
Como lo que no se suele saber es la media y la varianza podemos hacer :

Error admitido y tamaño de la muestra :
cuando decimos que estamos admitiendo un error máximo de esto es : la diferncia máxima entre la media poblacional y la media muestral debe ser menor que este valor . Como se puede observar de este valor se puede controlar dos parámetros , n y z .

El tamaño mínimo de una encuesta depende de la confianza que se desee para los resultados y del error máximo que se esté dispuesto a asumir :
E = despejando
Analogamente se puede hacer para la distribución de proporciones .

Ejemplo : se desea realizar una investigación para estimar el peso medio de los hijos de madres fumadoras . Se admite un error máximo de 50 gr , con una confianza del 95% . Si por estudios se sabe que la desviación típica es de 400 gr ¿ Qué tamaño mínimo de muestra se necesita en la investigación ?
El tamaño mínimo de la muestra debe ser n = = 246

Contraste de hipótesis sobre la media poblacional : La media muestral ppuede ser diferente de la media poblacional . Lo normal es que estas diferencias sean pequeñas y estén justificadas por el azar , pero podría suceder que no fuesen debidas al azar sino a que los parámetros poblacionales sean otros , que por los motivos que sea , han cambiado .
El contraste de hipótesis es el instrumento que permite decidir si esas diferencias pueden interpretarse como fluctuaciones del azar (hipótesis nula)o bien , son de tal importancia que requieren una explicación distinta (hipótesis alternativa). Como en los intervalos de confianza las conclusiones se formularán en términos de probabilidad .
Comparando la media poblacional y la media muestral ¿ Podemos asegurar que esa muestra procede de una población de media $_0$? La respuesta será no cuando $_0$ no pertenezca al intervalo de confianza de , para el nivel de significación prefijado , por el contrario la respuesta será sí cuando sí pertenezca a tal intervalo .
Sí pertenece a la población................ <u>se acepta la hipótesis nula</u> . Otra forma de verlo es que :

No pertenece a la población .... <u>se rechaza la hipótesis nula</u> . Otra forma de verlo es que :

Error de tipo I : es el que cometemos cuando rechazamos la hipótesis nula siendo verdadera .
Error de tipo II : es el que cometemos cuando aceptamos la hipótesis nula siendo falsa .

Podemos hacer todavía dos preguntas :
¿ La muestra procede de una población con media mayor que la supuesta ?
Se acepta que la media poblacional es mayor que la supuesta cuando : desarrollando la igualdad obtenemos que :

La media poblacional debe de estar por encima de y por lo tanto por encima de
La rechazamos en caso contrario .

¿ La muestra procede de una población con media menor que la supuesta ?
Se acepta que la media es menor que la supuesta cuando : desarrollando la igualdad obtenemos que :

La media poblacional debe de estar por debajo de y por lo tanto por debajo de
La rechazamos en caso contrario .

Nota : No olvidemos que en todas las ecuaciones anteriores si se desconoce la deviación típica de la población debemos sustituirla por la cuasivarianza de la muestra.

Contraste de hipótesis sobre la proporción p : por analogía con el apartado anterior par responder a la pregunta : ¿ Puede asegurarse que esa muestra de proporción procede de una población con proporción p_0 ?
La respuesta será sí cuando : con una probabilidad de 1 -
Se admite que la media poblacional es mayor que un valor p_0 si :

Se admite que la media poblacional es menor que un valor $\frac{\bar{p} - p_0}{\sqrt{pq/n}} \leq -z_\alpha$

PRUEBAS DE HIPÓTESIS

Si queremos decidir entre dos hipótesis que afectan a un cierto parámetro de la población, a partir de la información de la muestra usaremos el contraste de hipótesis, cuando optemos por una de estas dos hipótesis, hemos de conocer una medida del error cometido, es decir, cuantas veces de cada cien nos equivocamos.

En primer lugar, veremos cómo se escribirían las hipótesis que queremos contrastar:

H0 se llama hipótesis nula y es lo contrario de lo que sospechamos que va a ocurrir (suele llevar los signos igual, mayor o igual y menor o igual)

H1 se llama hipótesis alternativa y es lo que sospechamos que va a ser cierto (suele llevar los signos distinto, mayor y menor)

Los contrastes de hipótesis pueden ser de dos tipos:

Bilateral: En la hipótesis alternativa aparece el signo distinto.

Unilateral: En la hipótesis alternativa aparece o el signo > o el signo <.

Podemos aceptar una hipótesis cuando en realidad no es cierta, entonces cometeremos unos errores, que podrán ser de dos tipos:

Error de tipo I: Consiste en aceptar la hipótesis alternativa cuando la cierta es la nula.

Error de tipo II: Consiste en aceptar la hipótesis nula cuando la cierta es la alternativa.

Estos errores los aceptaremos si no son muy grandes o si no nos importa que sean muy grandes.

alfa: Es la probabilidad de cometer un error de tipo I.

beta: Es la probabilidad de cometer un error de tipo II.

De los dos, el más importante es alfa que llamaremos nivel de significación y nos informa de la probabilidad que tenemos de estar equivocados si aceptamos la hipótesis alternativa.

Debido a que los dos errores anteriores a la vez son imposibles de controlar, vamos a fijarnos solamente en el nivel de significación, este es el que nos interesa ya que la hipótesis alternativa que estamos interesados en probar y no queremos aceptarla si en realidad no es cierta, es decir, si aceptamos la hipótesis alternativa queremos equivocarnos con un margen de error muy pequeño.

El nivel de significación lo marcamos nosotros. Si es grande es más fácil aceptar la hipótesis alternativa cuando en realidad es falsa. El valor del nivel de significación suele ser un 5%, lo que significa que 5 de cada 100 veces aceptamos la hipótesis alternativa cuando la cierta es la nula.

Solamente vamos a estudiar el contraste bilateral para la media.

Pruebas de Chi-Cuadrada

Pruebas de bondad de ajuste (goodness of fit)

Esta es una prueba que determina si una muestra pudo haberse obtenido de una distribución poblacional hipotética. Se analizan las frecuencias (o proporciones) de una variable categórica.

Ejemplo: En la ultima encuesta para determinar la intención de voto, se seleccionó una muestra de 678 hombres y 522 mujeres1[1][1]. Si la proporción de la población de hombres y mujeres es de 53.87 % y 46.13%2[2][2] respectivamente, podemos preguntarnos si las proporciones de hombres y mujeres en la muestra corresponden a las de la población. Es decir, la hipótesis nula sería

H_0: La muestra proviene de la población especificada

H_1: La muestra no proviene de la población especificada

En la muestra de 1200 personas, tendríamos dos frecuencias, la observada y la esperada.

	Hombres i = 1	Mujeres i = 2	Totales
Frecuencia observada $f_{o,i}$	678	522	1200
Frecuencia esperada $f_{e,i}$	646.44	553.56	1200

Se puede calcular un estadístico chi-cuadrado para probar la hipótesis nula, el cual se calcula de la siguiente forma

$$\chi^2 = \sum_{i=1}^{k} \frac{(f_{o,i} - f_{e,i})^2}{f_{e,i}}$$

el cual procede de una distribución chi-cuadrado de grados de libertad *gl* dados por

$$gl = k - c - 1$$

donde *k* es el número de categorías y *c* es el número de parámetros poblacionales desconocidos estimados por estadísticos muestrales.

Este estadístico puede ser comparado con un valor crítico χ^2_α para un nivel de significancia dado α, o bien se calcula el valor p para determinar si se rechaza o no la hipótesis nula. El criterio es si $\chi^2 > \chi^2_\alpha$, entonces la hipótesis nula se rechaza. El valor p correspondiente es la probabilidad de que si se toma otra muestra de 1200 personas, el valor de χ^2 exceda el valor dado para esta muestra en particular.

En este ejemplo, para una significancia del 5%

$$\chi^2 = \frac{(678-646.44)^2}{646.44} + \frac{(522-553.56)^2}{553.56} \cong 3.34$$

$$gl = 2 - 0 - 1 = 1$$

$$\chi^2_\alpha = CHIINV(\alpha, gl) = CHIINV(0.05,1) \cong 3.84$$

$$\text{valor p} = CHIDIST(\chi^2, gl) \cong 0.0676 = 6.76\%$$

por lo que no se rechaza la hipótesis nula con una significancia del 5%. El valor p indica que se puede rechazar la hipótesis nula con una significancia del 6.76%. Este valor p también se puede calcular en Excel, utilizando el comando CHITEST (Excel 2000, versión en inglés), teniendo como entradas las frecuencias observadas y las esperadas (sin los totales).

Ejemplo:

La revista *Shopping Journal* publicó un estudio en el que se indicaba el desglose de clientes de centros comerciales (en EEUU) según la edad, en una ciudad grande, era

Categoría de Edad	< 21 $i = 1$	21-35 $i = 2$	36-50 $i = 3$	51-65 $i = 4$	> 65 $i = 5$
Proporción	10%	32%	31%	16%	11%

La cadena de tiendas Marshall obtuvo una muestra aleatoria de 250 observaciones de edades de clientes, con las cuales se observaron las siguientes frecuencias

Categoría de Edad	< 21 $i = 1$	21-35 $i = 2$	36-50 $i = 3$	51-65 $i = 4$	> 65 $i = 5$

Frecuencia observada $f_{o,i}$	18	51	42	89	50

Lo que se quiere determinar es si esta muestra pertenece o no a la población de clientes de centros comerciales en ciudades grandes. Para esto, podemos establecer las hipótesis nula y alternativa en la siguiente forma:

H_0: $p_1 = 0.10$, $p_2 = 0.32$, $p_3 = 0.31$, $p_4 = 0.16$, $p_5 = 0.11$

H_1: Alguna de las igualdades no se cumple

Se procede a calcular las proporciones esperadas, asumiendo que la hipótesis nula es cierta, lo cual da como resultado la tabla de frecuencias esperadas

Categoría de Edad	< 21 $i = 1$	21-35 $i = 2$	36-50 $i = 3$	51-65 $i = 4$	> 65 $i = 5$
Frecuencia esperada $f_{e,i}$	25	80	77.5	40	27.5

Se calcula el estadístico

$$\chi^2 = \sum_{i=1}^{k} \frac{\left(f_{o,i} - f_{e,i}\right)^2}{f_{e,i}} \cong 107.2$$

el cual tiene grados de libertad $gl = 5 - 0 - 1 = 4$. El valor crítico, para una significancia del 1% se calcula como

$$\chi^2_\alpha = CHIINV(\alpha, gl) = CHIINV(0.01,4) \cong 13.277$$

por lo que se rechaza la hipótesis nula. El valor p de esta prueba es

valor p = *CHIDIST*(107.2, 4) ≅ 0

lo cual quiere decir que la hipótesis nula se rechaza casi sin la posibilidad de cometer un error de tipo I.

TEORIA DE PEQUEÑAS MUESTRAS

DISTRIBUCION JI-CUADRADA (X^2)

En realidad la distribución ji-cuadrada es la distribución muestral de s^2. O sea que si se extraen todas las muestras posibles de una población normal y a cada muestra se le calcula su varianza, se obtendrá la distribución muestral de varianzas.

Para estimar la varianza poblacional o la desviación estándar, se necesita conocer el estadístico X^2. Si se elige una muestra de tamaño *n* de una población normal con varianza σ^2, el estadístico:

$$\frac{(n-1)s^2}{\sigma^2}$$

tiene una distribución muestral que es una **distribución ji-cuadrada** con gl=n-1 **grados de libertad** y se denota X^2 (X es la minúscula de la letra griega ji). El estadístico ji-cuadrada esta dado por:

$$X^2 = \frac{(n-1)s^2}{\sigma^2}$$

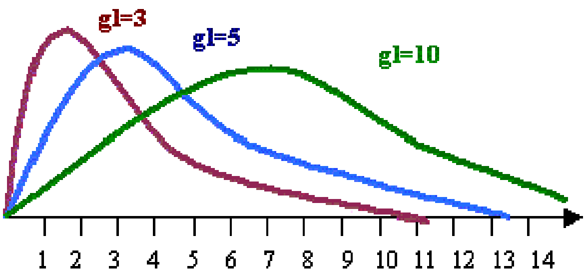
donde *n* es el tamaño de la muestra, s^2 la varianza muestral y σ^2 la varianza de la población de donde se extrajo la muestra. El estadístico ji-cuadrada también se puede dar con la siguiente expresión:

$$X^2 = \frac{\sum (x-\bar{x})^2}{\sigma^2}$$

Propiedades de las distribuciones ji-cuadrada

1. Los valores de X^2 son mayores o iguales que 0.
2. La forma de una distribución X^2 depende del gl=n-1. En consecuencia, hay un número infinito de distribuciones X^2.
3. El área bajo una curva ji-cuadrada y sobre el eje horizontal es 1.
4. Las distribuciones X^2 no son simétricas. Tienen colas estrechas que se extienden a la derecha; esto es, están sesgadas a la derecha.
5. Cuando n>2, la media de una distribución X^2 es n-1 y la varianza es 2(n-1).
6. El valor modal de una distribución X^2 se da en el valor (n-3).

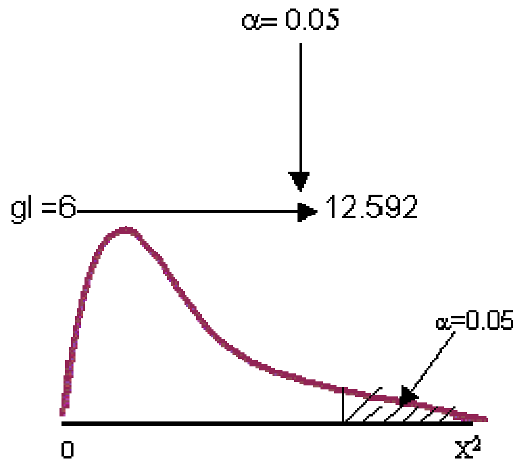
La siguiente figura ilustra tres distribuciones X^2. Note que el valor modal aparece en el valor (n-3) = (gl-2).



La función de densidad de la distribución X^2 esta dada por:

$$f(x) = \frac{1}{2^{v/2}\Gamma\left(\frac{v}{2}\right)} x^{v/2-1} e^{-x/2}$$ para x>0

La tabla que se utilizará para estos apuntes es la del libro de probabilidad y estadística de Walpole, la cual da valores críticos X^2_α(gl) para veinte valores especiales de α. Para denotar el valor crítico de una distribución X^2 con gl grados de libertad se usa el símbolo X^2_α(gl); este valor crítico determina a su derecha un área de α bajo la curva X^2 y sobre el eje horizontal. Por ejemplo para encontrar $X^2_{0.05}(6)$ en la tabla se localiza 6 gl en el lado izquierdo y $\alpha = 0.05$ a o largo del lado superior de la misma tabla.



Cálculo de Probabilidad

El cálculo de probabilidad en una distribución muestral de varianzas nos sirve para saber como se va a comportar la varianza o desviación estándar en una muestra que proviene de una distribución normal.

Ejemplos:

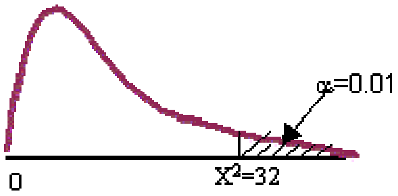
1. Suponga que los tiempos requeridos por un cierto autobús para alcanzar un de sus destinos en una ciudad grande forman una distribución normal con una desviación estándar σ =1 minuto. Si se elige al azar una muestra de 17 tiempos, encuentre la probabilidad de que la varianza muestral sea mayor que 2.

   Solución:

   Primero se encontrará el valor de ji-cuadrada correspondiente a s^2=2 como sigue:

$$X^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(17-1)(2)}{(1)^2} = 32$$

   El valor de 32 se busca adentro de la tabla en el renglón de 16 grados de libertad y se encuentra que a este valor le corresponde un área a la derecha de 0.01. En consecuencia, el valor de la probabilidad es P(s^2>2)



2. Encuentre la probabilidad de que una muestra aleatoria de 25 observaciones, de una población normal con varianza $\sigma^2 = 6$

, tenga una varianza muestral:

a. Mayor que 9.1

b. Entre 3.462 y 10.745

Solución.

a. Primero se procederá a calcular el valor de la ji-cuadrada:

$$X^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(25-1)(9.1)}{6} = 36.4$$

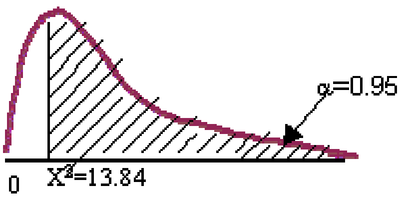
Al buscar este número en el renglón de 24 grados de libertad nos da un área a la derecha de 0.05. Por lo que la P(s^2 >9.1) = 0.05

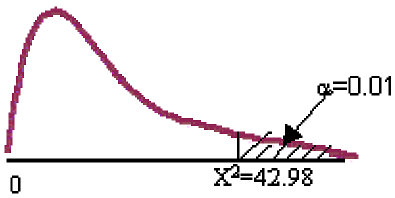
1. Se calcularán dos valores de ji-cuadrada:

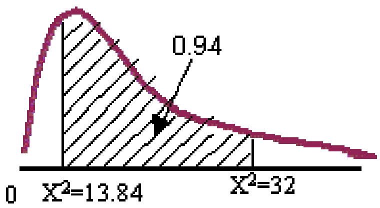
$$X^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(25-1)(3.462)}{6} = 13.847 \quad \text{y} \quad X^2 = \frac{(25-1)(10.745)}{6} = 42.98$$

Aquí se tienen que buscar los dos valores en el renglón de 24 grados de libertad. Al buscar el valor de 13.846 se encuentra un área a la derecha de 0.95. El valor de 42.98 da un área a la derecha de 0.01. Como se está pidiendo la probabilidad entre dos valores se resta el área de 0.95 menos 0.01 quedando 0.94.

Por lo tanto la P(3.462 $\leq s^2 \leq$ 10.745) = 0.94







Estimación de la Varianza

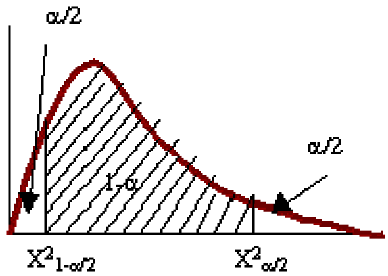
Para poder estimar la varianza de una población normal se utilizará la distribución ji-cuadrada.

$$X^2 = \frac{(n-1)s^2}{\sigma^2}$$

Al despejar esta fórmula la varianza poblacional nos queda:

$$\sigma^2 = \frac{(n-1)s^2}{X^2}$$

Los valores de X^2 dependerán de nivel de confianza que se quiera al cual le llamamos $1-\alpha$. Si nos ubicamos en la gráfica se tiene:



Ejemplos:

1. Los siguientes son los pesos, en decagramos, de 10 paquetes de semillas de pasto distribuidas por cierta compañía: 46.4, 46.1, 45.8, 47.0, 46.1, 45.9, 45.8, 46.9, 45.2 y 46. Encuentre un intervalo de confianza de 95% para la varianza de todos los paquetes de semillas de pasto que distribuye esta compañía, suponga una población normal.

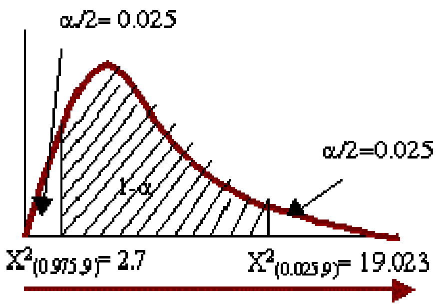
   Solución:

   Primero se calcula la desviación estándar de la muestra:

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{(46.4-46.12)^2 + (46.1-46.12)^2 + \ldots + (46-46.12)^2}{10-1}} = 0.5347$$

   al elevar este resultado al cuadrado se obtiene la varianza de la muestra $s^2 = 0.286$.

   Para obtener un intervalo de confianza de 95% se elige un $\alpha = 0.05$. Después con el uso de la tabla con 9 grados de libertad se obtienen los valores de X^2.



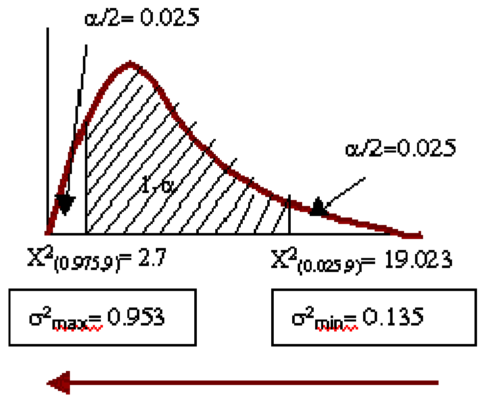
Se puede observar en la gráfica anterior que el valor de X^2 corre en forma normal, esto es de izquierda a derecha.

Por lo tanto, el intervalo de confianza de 95% para la varianza es:

$$\sigma^2 \max = \frac{(10-1)(0.286)}{2.7} = 0.953$$

$$\sigma^2 \min = \frac{(10-1)(0.286)}{19.023} = 0.135$$

Graficamente:



Se observa que la varianza corre en sentido contrario, pero esto es sólo en la gráfica. La interpretación quedaría similar a nuestros temas anteriores referentes a estimación. Con un nivel de confianza del 95% se sabe que la varianza de la población de los pesos de los paquetes de semillas de pasto esta entre 0.135 y 0.935 decagramos al cuadrado.

2. En trabajo de laboratorio se desea llevar a cabo comprobaciones cuidadosas de la variabilidad de los resultados que producen muestras estándar. En un estudio de la cantidad de calcio en el agua potable, el cual se efectúa como parte del control de calidad, se analizó seis veces la misma muestra en el laboratorio en intervalos aleatorios. Los seis resultados en partes por millón fueron 9.54, 9.61, 9.32, 9.48, 9.70 y 9.26. Estimar la varianza de los resultados de la población para este estándar, usando un nivel de confianza del 90%.

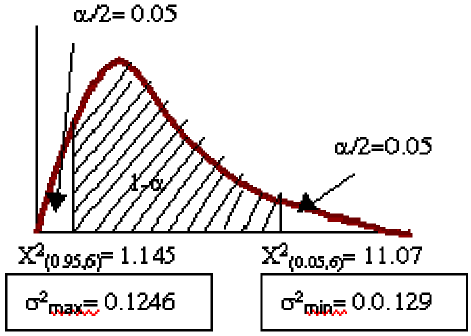
Solución:

Al calcular la varianza de la muestra se obtiene un valor de s^2= 0.0285.

Se busca en la tabla los valores correspondientes con 5 grados de libertad, obteniéndose dos resultados. Para $X^2_{(0.95,5)}$= 1.145 y para $X^2_{(0.0,5)}$= 11.07.

Entonces el intervalo de confianza esta dado por:

$$\sigma^2 \max = \frac{(6-1)(0.0285)}{1.145} = 0.1246 \quad \text{y} \quad \sigma^2 \min = \frac{(6-1)(0.0285)}{11.07} = 0.0129$$



Ensayo de Hipótesis para la Varianza de una Población Normal

En la mayoría de los casos se tiene el problema de desconocer la varianza o desviación estándar de la población, en donde las distribuciones son normales. Si se desea probar una hipótesis acerca de la varianza se puede hacer utilizando las medidas estadísticas con las que se construyó el intervalo de confianza σ^2, esto es con la distribución Ji- cuadrada.

Ejemplos:

1. Una compañía que produce una parte maquinada para un motor, afirma que tiene una varianza de diámetro no mayor a 0.0002 pulgadas. Una muestra aleatoria de 10 de dichas partes dio una varianza de muestra s^2 = 0.0003. Si se supone que las medidas del diámetro se distribuyen en forma normal, ¿hay evidencia para refutar lo que afirma el proveedor? Use α = 0.05.

Solución:

Como en todos los ensayos de hipótesis que se han realizado anteriormente el procedimiento es el mismo. Después de que se identifican los datos, se plantea la hipótesis para determinar el tipo de ensayo.

Datos:

σ^2 = 0.0002

n = 10

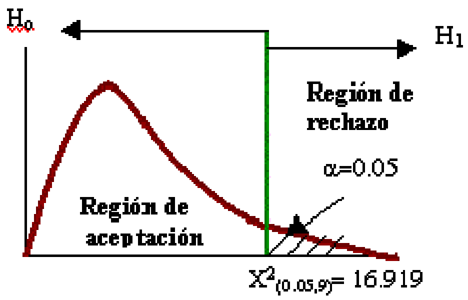
s^2 = 0.0003

α = 0.05

Ensayo de hipótesis:

H_o; σ^2 = 0.0002

H_1; σ^2 > 0.0002



Regla de decisión:

Si $X^2_R \leq 16.919$ no se rechaza H_o.

Si $X^2_R > 16.919$ se rechaza H_o.

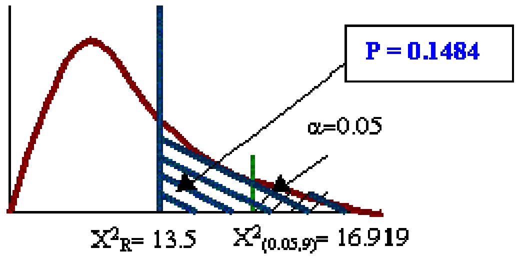
Cálculos:

$$X^2{}_R = \frac{(n-1)s^2}{\sigma^2} = \frac{(10-1)(0.0003)}{0.0002} = 13.5$$

Justificación y decisión:

Como 13.5 no es mayor que 16.919 por lo tanto no se rechaza H_o y se concluye con un nivel de significancia de 0.05 que no se puede refutar la afirmación del proveedor.

Este ejercicio se puede aprovechar para calcular el valor de P. En la tabla se busca el valor de 13.5 en el renglón de 9 grados de libertad. Interpolando entre 0.10 y 0.20 se obtiene un valor de P de 0.1484.



2. El contenido de azúcar del almíbar de los duraznos enlatados tiene una distribución normal, donde se cree que la varianza es σ^2 = 18 mg^2. Se toma una muestra de 10 latas dieron una desviación estándar de 4.8 mg. ¿Muestran estos datos suficiente evidencia para decir que la varianza ha cambiado?. Use un α = 0.05 y calcule el valor de P.

Solución:

Datos:

σ^2 = 18

n = 10

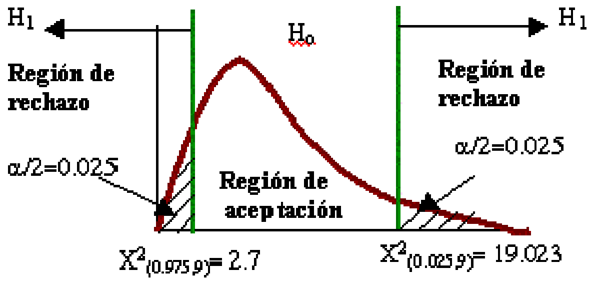
s = 4.8

α = 0.05

Ensayo de hipótesis:

H_o; σ^2 = 18

H_1; $\sigma^2 \neq$ 18



Regla de decisión:

Si $2.7 \leq X^2_R \leq 19.023$ no se rechaza H_o.

Si $X^2_R < 2.7$ ó si $X^2_R > 19.023$ se rechaza H_o.

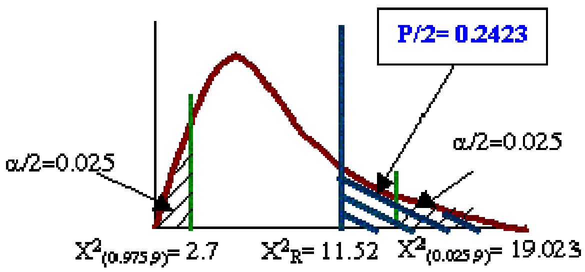
Cálculos:

$$X^2_R = \frac{(n-1)s^2}{\sigma^2} = \frac{(10-1)(4.8)^2}{18} = 11.52$$

Justificación y decisión:

Como 11.52 está entre 2.7 y 19.023, no se rechaza H_o, y se concluye con un nivel de significancia de 0.05 que la varianza del contenido de azúcar del almíbar no ha cambiado, esto es es de 18 mg^2.

Si recordamos al principio de este tema se dijo que la media de la distribución ji-cuadrada es (n-1), por lo tanto la media de este ejercicio es de 9. Como el valor real de X^2_R = 11.52 este número se encuentra a la derecha de la media, lo cual quiere decir que el valor de P/2 será el área a la derecha del valor de X^2_R. Al buscar el valor de 11.52 en la tabla se obtiene un área de 0.2423, por lo tanto P/2 = 0.2423 y **P= (2)(0.2423) = 0.4846**



3. Experiencia anterior indica que el tiempo que se requiere para que los estudiantes de último año de preparatoria completen una prueba estandarizada es una variable aletoria normal con una desviación estándar de seis minutos. Se toma una muestra aleatoria de 20 estudiantes de último año de preparatoria y se obtiene una desviación estándar de 4.51. ¿Muestran estos datos suficiente evidencia para decir que la desviación estándar disminuyó?. Utilice el valor de P para su decisión.

Solución:

Datos:

$\sigma = 6$

n = 20

s = 4.51

Ensayo de hipótesis:

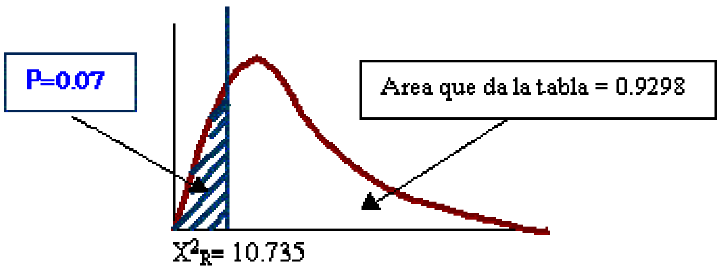
H_o; $\sigma = 6$

H_1; $\sigma < 6$

Cálculos:

$$X^2_R = \frac{(n-1)s^2}{\sigma^2} = \frac{(20-1)(4.51)^2}{(6)^2} = 10.735$$

Para obtener el valor de P, se busca en la tabla el 10.735 con 19 grados de libertad, y el área que se encuentra es la que está a la derecha de este valor. Como la media de esta distribución ji-cuadrada es de 19, por lo tanto el valor de 10.735 queda a la izquierda de la media. El valor de P es de 0.07, y con esto se puede concluir que si hubiéramos utilizado un nivel de significancia de 0.10, se rechaza H_o y se concluye que la desviación estándar disminuyo, pero si se utiliza un valor de $\alpha = 0.05$, entonces no se rechaza H_o y se concluiría que la desviación estándar no disminuyó. La decisión depende del error tipo I que esté dispuesto a tolerar el investigador.



Error tipo II ó β

El error tipo II se calcula de la misma forma en la que se calculó con la distribución z. Se realizarán algunos ejercicios en los cuales se determinará la probabilidad de cometer el error tipo II, utilizando la tabla de la distribución Ji-cuadrada.

1. Se tiene un ensayo de hipótesis unilateral derecho, con n=20 y $\alpha = 0.05$

   H_o; $\sigma = 0.10$

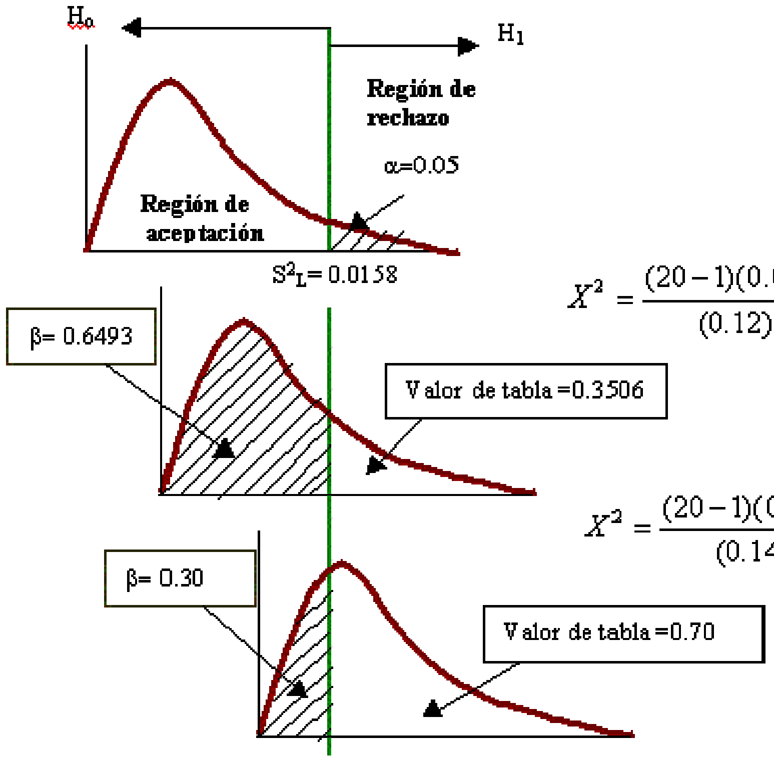
   H_1; $\sigma > 0.10$

   Se quiere calcular el error tipo II ó β si las desviaciones estándar verdaderas fueran de 0.12 y 0.14.

   Solución:

Para poder calcular el error tipo II, primero se debe encontrar el valor de la varianza muestral límite, esto es s^2_L, para poder calcular los valores de X^2 y posteriormente calcular el área. Al buscar en la tabla $X^2_{(0.05,19)}$=30.144, este valor se sustituirá en la formula. Al despejar de la fórmula original de X^2 se obtiene:

$$s_L^2 = \frac{X_L^2 \sigma^2}{(n-1)} = \frac{(30.144)(0.10)^2}{(20-1)} = 0.0158$$



$$X^2 = \frac{(20-1)(0.0158)}{(0.12)^2} = 20.84$$

$$X^2 = \frac{(20-1)(0.0158)}{(0.14)^2} = 15.31$$

2. Encontrar el error tipo II para el ejercicio 2 de esta sección, en donde el ensayo es bilateral pues se quiere ver si la varianza del contenido de azúcar en el almíbar de los duraznos ha cambiado. Suponga una varianza real de 20 y 26.

Solución:

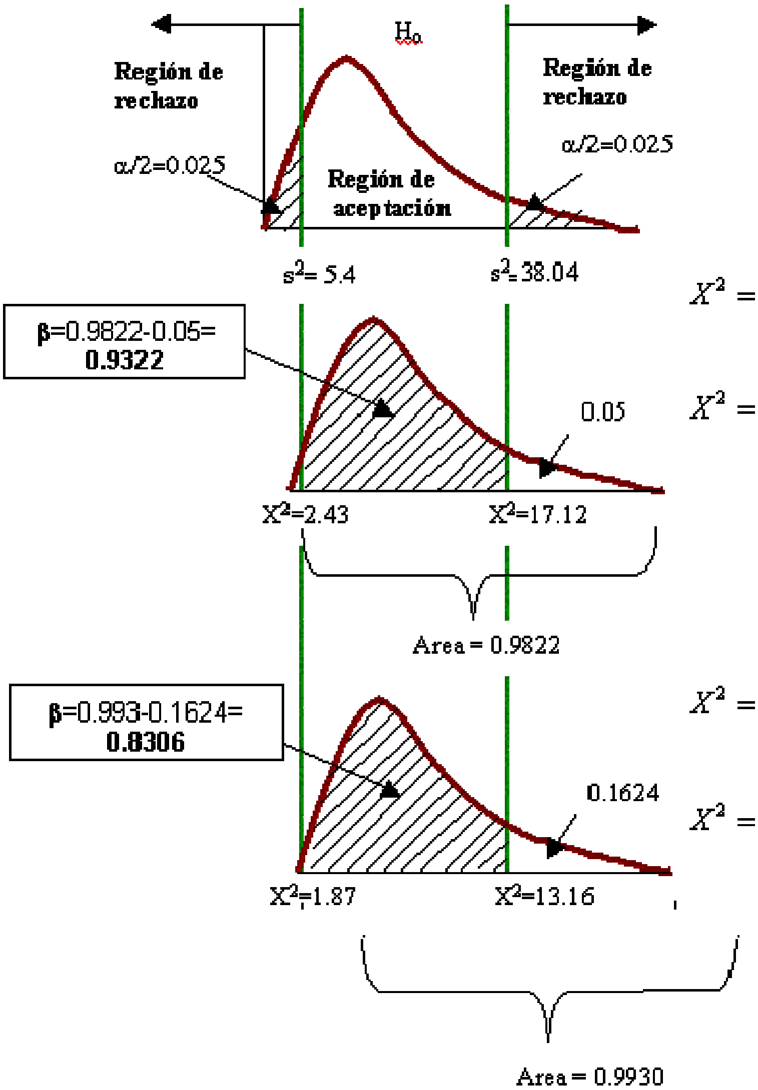
Como este es un ensayo bilateral se tendrán dos valores de s^2_L. Los cuales se calcularán utilizando las ji-cuadradas límites que eran de de 2.7 y 19.023.

$$s_L^2 = \frac{X_L^2 \sigma^2}{(n-1)} = \frac{(2.7)(18)}{(10-1)} = 5.4$$

y

$$s_L^2 = \frac{X_L^2 \sigma^2}{(n-1)} = \frac{(19.023)(18)}{(10-1)} = 38.04$$

Estos dos valores se utilizarán para calcular las nuevas ji-cuadradas para calcular el valor de β



$$X^2 = \frac{(10-1)(5.4)}{20} = 2.43$$

$$X^2 = \frac{(10-1)(38.04)}{20} = 17.12$$

$$X^2 = \frac{(10-1)(5.4)}{26} = 1.87$$

$$X^2 = \frac{(10-1)(38.04)}{26} = 13.16$$

3. Regresión y correlación

- **Técnicas de regresión lineal simple**

Se denomina regresión a la estimación de una variable Y (variable dependiente) a partir de otra variable X (variable independiente) o bien de varias variables (x_1, x_2, . . . , x_n) relaciones entre sí.

Cuando se estima una variable Y dependiente, a partir de otra variable X se habla de regresión lineal.

Se denomina correlación al grado de relación de interdependencia, que existe entre dos variables.

III IInvestigación de operaciones

Programación lineal

Consiste en una función objetivo lineal a maximizar o minimizar en precencia de un conjunto de restricciones lineales y su máximo exponente es 1.

Ejemplo:

Min Z= 3x1 + 2x2

s.a. -x1 + x2 >/ 5

3x1 - 18x2 >/ -3

x2 >/ 2

x1, x2 >/ 0

En general,

Max Z= C1X1 + C2x2 + ... +C4x4

s.a. a11x1 + a12x2 + ... + a14x4

En forma matricial:

Max Z = c . x

s.a. Ax

x

Donde: c t Mixn, x t mnxi, A t Mmxn, b t Mmxi, "Al conjunto de valores númericos X1°, X2°, ..., Xn° que satisfacen al conjunto de restricciones, se le denomina punto factible [X1°, X2°, ..., Xn°]

Nota: Al conjunto de puntos factibles, se les llama región factible.

PROBLEMA DE DIETA

Lupita gómez está preocupada por su sobrepeso y el costo de la comida diaria, ella sabe, que para bajar de peso, debe consumir a lo más, 1350 kilokalorías, pero requiere de un mínimo de 500 miligramos de vitamina A, 350 mm. de Calcio, 200 mm. de proteinas y 150 mm de minerales.

De a cuerdo a esto, ella a elegido 6 alimentos, que a su criterio son ricos en nutrientes y de bajo costo:

ALIMENTO	PORCION	VITAM. A	CALCIO	PROTEINAS	MINERALES	COSTO	KALORIAS
LECHE	1 TAZA	105	75	50	35	$ 5	60
HUEVO	2 PIEZAS	75	80	50	15	$ 7	50
ESPINACAS	1 RACION	100		125	78	$ 2	
CHULETAS	2 CHULETS	25	10	55		$45	175
PESCADO	1 MOJARRA	150	50	100	50	$60	150
PASTEL	2 REB.	30	05	08		$50	200

Lupita se ha dado cuenta de que es muy posible que si ella se comiera cinco mojarras diarias, tendría satisfechas sus necesidades de nutrientes y de kilokal.; pero no está dispuesta a tal sacrificio, por tanto, ella ha decidido que lo más que puede comerce en porciones de leche son tres, de huevo dos, de espinacas uno, una de chuletas de cerdo, dos de pescado y de pastel una y media porciones.

Proporcionar a Lupita el modelo de P.L. que determine la dieta más económica.

xi:= (Es un valor de decición binaria.) Cantidad de porción del alimeto a incluir en la dieta diaria, donde i= 1, ..., 6.

Min Z=	005x1	+	007x2	+	002x3	+	045x4	+	060x5	+	050x6			
s.a.	060x1	+	050x2	+	000x3	+	175x4	+	150x5	+	200x6			
	060x1	+	050x2	+	000x3	+	175x4	+	150x5	+	200x6	>/	1350	Kalorias
	105x1	+	075x2	+	100x3	+	025x4	+	150x5	+	030x6	>/	0500	Vitam. A
	075x1	+	080x2	+	------	+	100x4	+	050x5	+	005x6	>/	0350	Calcio
	050x1	+	050x2	+	125x3	+	055x4	+	100x5	+	008x6	>/	0200	Proteinas
	035x1	+	015x2	+	078x3	+	000x4	+	050x5	+	000x6	>/	0150	Minerales
0	x1												3	
0			x2										2	
0					x3								1	
0							x4						1	
0									x5				2	
0											x6		3	

PROBLEMA DOS

CONTRATACION DE PERSONAL

La cadena de reataurantes California, trabaja las 24 hrs. del día, han abierto un nuevo restaurante en Div. del Norte, y por ello requiere contratar meseras. El administrador ha dividido las 24 horas en horarios de tres y determina el número mínimo requerido de meseras para dichos horarios.

Número	Horario	# mínimo meseras
1	0-3	4
2	3-6	3
3	6-9	8
4	9-12	6
5	12-15	7
6	15-18	14
7	18-21	10
8	21-24	5

Si cada mesera trabaja 3 horarios consecutivos, le regalan una hora de comida, determinar el P.L. que determine el menor número de meseras por contratar.

xi:= Será número de meseras a contratar para el horario i, donde i = 1, ..., 8.

Min Z =	x1	+	x2	+	x3	+	x4	+	x5	+	x6	+	x7	+	x8			
s.a.	x1	+	x2	+	x3											>/	08	
			x2	+	x3	+	x4									>/	06	
					x3	+	x4	+	x5							>/	07	
							x4	+	x5	+	x6					>/	14	

									x5	+	x6	+	x7			>/	10	
											x6	+	x7	+	x8	>/	05	
	x1	+											x7	+	x8	>/	04	
	x1	+	x2												x8	>/	03	
Donde: Xi >/ 0																		

EJEMPLO 3: PROBLEMA BINARIO (Se invierte todo o nada)

Samnonite es una empresa que se dedica a la fabricación de mochilas y tiene disponible un millón de pesos para invertir. El lic. en Administración Omar, tiene a su cargo, la díficil tarea de decidir en cuales de los cinco proyectos siguientes desea invertir:

PROYECTO	COSTO	UTILIDAD
1	500000	325000
2	200000	122000
3	195000	095000
4	303000	111000
5	350000	150000

Si el elegir un proyecto implica, pagar el costo total del mismo, hacer el modelo P.L. que defina la mejor inversión para el licenciado.

Xi = 1; Si se invierte en el proyecto i.

Xi = 0; Si No se invierte en el proyecto i.

Donde i = 1, ..., 5 proyectos (-:Este si es fácil amigos, es sólo poner lo siguiente:-)

Max Z =	325x1	+	122x2	+	095x3	+	011x4	+	150x5		
s. a.	500x1	+	200x2	+	195x3	+	303x4	+	350x5		1000

EJEMPLO 4: PROBLEMA DE PRODUCCION

La compañía W, trabaja 3 tipos distintos de salas: Económicas, comerciales y de lujo. El tiempo que requieren de carpintería para cada sala son: 12 hrs., 14 hrs. y 18 hrs. respectivamente. El tiempo de tapicería que se requiere es de 13 Hrs, 14 hrs. y 25 hrs. respectivamente.

Si se dispone de 1385 hrs mensuales de carpintería, 1500 de Tapicería y 50 de Supervisión, y cada sala al ser vendida, nos proporciona una utilidad neta de 500, 850 y 1500 respectivamente; pero según un estudio de mercado, no se venderán más de 10 salas de lujo para el próximo mes y ya se tiene un pedido de 27 salas económicas, proporcionar el modelo de P.L. que reditúe las mejores ganancias.

Xi : = Tipo de salas del tipo i, a producir en el mes, donde i = 1,2,3 tipos de salas.

Max Z =	500x1	+	850x2	+	1500x3			
s.a.	012x1	+	014x2	+	0018x3		1385	Hrs. Carpintería
	013x1	+	014x2	+	0025x3		1500	Hrs. Tapicería
	001x1	+	001x2	+	03/2x3		0050	Hrs. Supervisión
	x1					>/	0027	
					x3	>/	0010	

Donde : x1, x2, x3 >/ 0 y enteras.

OPTIMO	*Lo mejor posible dadas las restricciones del sistema.*
EFICAZ	*Quién logra cumplir el objetivo.*
EFICIENTE	*Quién logra cumplir el objetivo al menor costo posible, en tiempo, en dinero, etc.*

EL METODO SIMPLEX PARA SOLUCIÓN DE PROBLEMAS DE PROGRAMACIÓN LINEAL

Es un procedimiento iterativo que permite ir mejorando la solución a cada paso. El proceso concluye cuando no es posible seguir mejorando más dicha solución.

Partiendo del valor de la función objetivo en un vértice cualquiera, el método consiste en buscar sucesivamente otro vértice que mejore al anterior. La búsqueda se hace siempre a través de los lados del polígono (o de las aristas del poliedro, si el número de variables es mayor). Cómo el número de vértices (y de aristas) es finito, siempre se podrá encontrar la solución.

El método del simplex se basa en la siguiente propiedad: si la función objetivo, *f*, no toma su valor máximo en el vértice A, entonces hay una arista que parte de A, a lo largo de la cual *f* aumenta.

El método del simplex fue creado en 1947 por el matemático George Dantzig .

El método del simplex se utiliza, sobre todo, para resolver problemas de programación lineal en los que intervienen tres o más variables.

El álgebra matricial y el proceso de eliminación de Gauss-Jordan para resolver un sistema de ecuaciones lineales constituyen la base del método simplex.

Con miras a conocer la metodología que se aplica en el Método SIMPLEX, vamos a resolver el siguiente problema:

Maximizar $Z = f(x,y) = 3x + 2y$

sujeto a:
$$2x + y \leq 18$$
$$2x + 3y \leq 42$$
$$3x + y \leq 24$$
$$x \geq 0 \text{ , } y \geq 0$$

Se consideran las siguientes fases:

1. Convertir las desigualdades en igualdades

Se introduce una *variable de holgura* por cada una de las restricciones, para convertirlas en igualdades, resultando el sistema de ecuaciones lineales:

$$2x + y + h = 18$$
$$2x + 3y + s = 42$$
$$3x + y + d = 24$$

2. Igualar la función objetivo a cero

$$-3x - 2y + Z = 0$$

3. Escribir la tabla inicial simplex

En las columnas aparecerán todas las variables del problema y, en las filas, los coeficientes de las igualdades obtenidas, una fila para cada restricción y la última fila con los coeficientes de la función objetivo:

Tabla I . Iteración nº 1

Base	Variable de decisión		Variable de holgura			Valores solución
	x	*y*	*h*	*s*	*d*	
h	2	1	1	0	0	18
s	2	3	0	1	0	42
d	**3**	1	0	0	1	24
Z	-3	-2	0	0	0	0

4. Encontrar la variable de decisión que entra en la base y la variable de holgura que sale de la base

A. Para escoger la variable de decisión que entra en la base, nos fijamos en la última fila, la de los coeficientes de la función objetivo y escogemos la variable con el coeficiente negativo mayor (en valor absoluto).
En nuestro caso, la variable *x* de coeficiente - 3.

Si existiesen dos o más coeficientes iguales que cumplan la condición anterior, entonces se elige uno cualquiera de ellos.

Si en la última fila no existiese ningún coeficiente negativo, significa que se ha alcanzado la solución óptima. Por tanto, lo que va a determinar el final del proceso de aplicación del método del simplex, es que en la última fila no haya elementos negativos.

La columna de la variable que entra en la base se llama *columna pivote* (En color azulado).

B. Para encontrar la variable de holgura que tiene que salir de la base, se divide cada término de la última columna (valores solución) por el término correspondiente de la columna pivote, siempre que estos últimos sean mayores que cero. En nuestro caso:
18/2 [=9] , 42/2 [=21] y 24/3 [=8]

Si hubiese algún elemento menor o igual que cero no se hace dicho cociente. En el caso de que todos los elementos fuesen menores o iguales a cero, entonces tendríamos una solución no acotada y no se puede seguir.

El término de la columna pivote que en la división anterior dé lugar al menor cociente positivo, el 3, ya 8 es el menor, indica la fila de la variable de holgura que sale de la base, *d*. Esta fila se llama *fila pivote* (En color **azulado**).

Si al calcular los cocientes, dos o más son iguales, indica que cualquiera de las variables correspondientes pueden salir de la base.

C. En la intersección de la fila pivote y columna pivote tenemos el elemento pivote operacional, **3**.

5. Encontrar los coeficientes de la nueva tabla.

Los nuevos coeficientes de *x* se obtienen dividiendo todos los coeficientes de la fila *d* por el pivote operacional, 3, que es el que hay que convertir en 1.

A continuación mediante la reducción gaussiana hacemos ceros los restantes términos de su columna, con lo que obtenemos los nuevos coeficientes de las otras filas incluyendo los de la función objetivo *Z*.

También se puede hacer utilizando el siguiente esquema:

Fila del pivote:

Nueva fila del pivote= (Vieja fila del pivote) : (Pivote)

Resto de las filas:

Nueva fila= (Vieja fila) - (Coeficiente de la vieja fila en la columna de la variable entrante) X (Nueva fila del pivote)

Veámoslo con un ejemplo una vez calculada la fila del pivote (fila de x en la Tabla II):

Vieja fila de s	2	3	0	1	0	42
	-	-	-	-	-	-
Coeficiente	2	2	2	2	2	2
	x	x	x	x	x	x
Nueva fila pivote	1	1/3	0	0	1/3	8
	=	=	=	=	=	=
Nueva fila de s	0	7/3	0	1	-2/3	26

Tabla II . Iteración nº 2

Base	Variable de decisión	Variable de holgura	Valores solución

	x	y	h	s	d	
h	0	**1/3**	1	0	-2/3	2
s	0	7/3	0	1	-2/3	26
x	1	1/3	0	0	1/3	8
Z	0	-1	0	0	1	24

Como en los elementos de la última fila hay uno negativo, -1, significa que no hemos llegado todavía a la solución óptima. Hay que repetir el proceso:

A. La variable que entra en la base es *y*, por ser la variable que corresponde al coeficiente -1

B. Para calcular la variable que sale, dividimos los términos de la última columna entre los términos correspondientes de la nueva columna pivote:
2:1/3 [=6] , 26:7/3 [=78/7] y 8:1/3 [=8]
y como el menor cociente positivo es 6, tenemos que la variable de holgura que sale es *h*.

C. El elemento pivote, que ahora hay que hacer 1, es **1/3**.

Operando de forma análoga a la anterior obtenemos la tabla:

Tabla III . Iteración nº 3

Base	Variable de decisión		Variable de holgura			Valores solución
	x	y	h	s	d	
y	0	1	3	0	-2	6
s	0	0	-7	0	**4**	12
x	1	0	-1	0	1	6
Z	0	0	3	0	-1	30

Como en los elementos de la última fila hay uno negativo, -1, significa que no hemos llegado todavía a la solución óptima. Hay que repetir el proceso:

A. La variable que entra en la base es *d*, por ser la variable que corresponde al coeficiente -1

B. Para calcular la variable que sale, dividimos los términos de la última columna entre los términos correspondientes de la nueva columna pivote:
6/(-2) [=-3] , 12/4 [=3], y 6:1 [=6]
y como el menor cociente positivo es 3, tenemos que la variable de holgura que sale es *s*.

C. El elemento pivote, que ahora hay que hacer 1, es **4**.

Obtenemos la tabla:

Tabla IV . Final del proceso

Base	Variable de decisión		Variable de holgura			Valores solución
	x	y	h	s	d	
y	0	1	-1/2	0	0	**12**
d	0	0	-7/4	0	1	3
x	1	0	-3/4	0	0	**3**
Z	0	0	5/4	0	0	**33**

Como todos los coeficientes de la fila de la función objetivo son positivos, hemos llegado a la solución óptima.

Los solución óptima viene dada por el valor de Z en la columna de los valores solución, en nuestro caso: **33.** En la misma columna se puede observar el vértice donde se alcanza, observando las filas correspondientes a las variables de decisión que han entrado en la base: **D(3,12)**

C. Matemáticas discretas

I. Lógica

1. <u>**Lógica proposicional**</u>

- **Reglas de inferencia**

$[p \wedge (p \rightarrow p)] \Rightarrow q$	Modus Ponens
$[(p \rightarrow q) \wedge \overline{q}\,] \Rightarrow \overline{p}$	Modus Tollens o demostración por contradicción
$[(p \rightarrow q) \wedge (q \rightarrow r)] \Rightarrow (p \rightarrow r)$	Ley del silogismo

Ejemplo:

2. <u>**Lógica de predicados**</u>

- **Reglas para fórmula bien formadas**

Una fórmula esta bien formada cuando esta bien estructurada de acuerdo con el conjunto de reglas establecidas y con el dominio de discurso.

- **Formas clausales: resolución, unificación**

La **resolución** es una técnica para demostrar teoremas en el lenguaje de la lógica.

La resolución considera 2 ó más cláusulas e intenta deducir hechos o relaciones no explícitas en la cláusulas originales del problema.

Las demostraciones por resolución involucran los siguientes pasos:

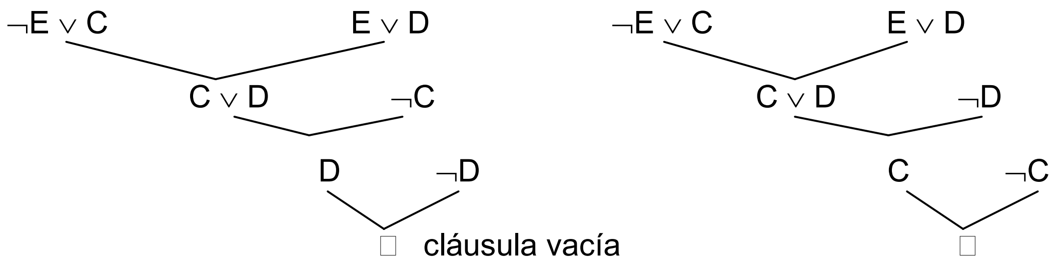
1. Poner las premisas o axiomas en forma clausular
2. Agregar al conjunto de axiomas la negación de lo que se demostrará en forma clausular
3. Resolver las cláusulas entre sí, produciendo nuevas cláusulas que se sigan lógicamente de ellas
4. Producir una contradicción, generando la cláusula vacía

Ejemplo:
Si estudio obtengo buenas calificaciones, si no estudio me divierto

Conclusión: Por lo tanto u obtengo buenas calificaciones o me divierto

E – Estudio
C – Calificaciones buenas
D – Divierto

$$\begin{array}{c} \neg E \vee C \\ E \vee D \\ \hline C \vee D \end{array}$$



II. Combinatoria

1. Análisis combinatorio

- ### Teoría de conteo

Definición .- Un conjunto de objetos en donde el orden de los objetos no importa se llama combinación.

Siempre que el enunciado tenga la palabra “elegir, escoger, tomar, sacar una muestra, de cuántas formas” será conteo o combinaciones.

Principio de conteo

Para encontrar el número total de opciones para un evento, se multiplica el número de opciones para cada parte.

Ejemplos:

Una cantante tiene 4 blusas y 5 pantalones para usar en conciertos. ¿De cuántas maneras diferentes se puede vestir para un concierto?
R = $4 \times 5 = 20$

Un identificado tipo etiqueta de un programa de computadora, se forma con una letra seguida de tres dígitos. Si se permiten las repeticiones, ¿cuántos identificadores distintos son posibles?
Solución: Hay 26 posibilidades para la letra inicial y 10 para cada uno de los tres dígitos. Entonces, por el principio de multiplicación generalizado, hay

$26 \times 10 \times 10 \times 10 = 26\ 000$ identificadores posibles

El nombre tradicional combinatorio para un subconjunto de r elementos de un conjunto S de n elementos es una combinación de n elementos tomados de r en r.

Fórmula:

$$C^n_r = \frac{n!}{r!\,(n - r)!}$$

- **Relaciones de recurrencia**

Definición.- Una relación de recurrencia para una sucesión $a_0, a_1, \ldots,$ es una ecuación que relaciona a_n, con alguno de sus antecesores $a_0, a_1, \ldots, a_{n-1}$.

Una relación de recurrencia define indirectamente el término n-ésimo de una sucesión.

Ejemplo:

Una serie geométrica es una sucesión infinita de números, como 5, 15, 45, 135, . . . , donde la división de cualquier término, distinto del primero, entre su predecesor inmediato es una constante llamada razón común. Para el ejemplo anterior esta razón común es 3, ya que 15 = 3(15), 45 = 3(15), etc. Si $a_0, a_1, a_2, \ldots,$ es una serie geométrica, entonces $a_1/a_0 = a_2/a_1 = \ldots = a_{n+1}/a_n = \ldots = r$, la razón común. En esta serie geométrica particular, resulta $a_{n+1} = 3a_n$, $n \geq 0$.

La relación de recurrencia $a_{n+1} = 3a_n$, $n \geq 0$ no define una serie geométrica única. Para especificar la sucesión particular descrita por $a_{n+1} = 3a_n$, es necesario conocer uno de los términos de dicha sucesión. De ahí que $a_{n+1} = 3a_n$, $n \geq 0$, $a_0 = 5$.

III. Relaciones y grafos

1. <u>Relaciones</u>

- **Ordenes parciales**

Definición.- Una relación puede considerarse como un cuadro que muestra las correspondencias de unos elementos con respecto a otros.

Definición.- Una relación R en un conjunto A recibe el nombre de relación de orden parcial si R es reflexiva, antisimétrica y transitiva.

- **Relaciones de equivalencia**

Definición.- Una relación que es reflexiva, simétrica y transitiva en un conjunto X se conoce con el nombre de relación de equivalencia sobre X.

2. Gráficas y árboles

- **Recorridos, números cromáticos, coloración de aristas y vértices**

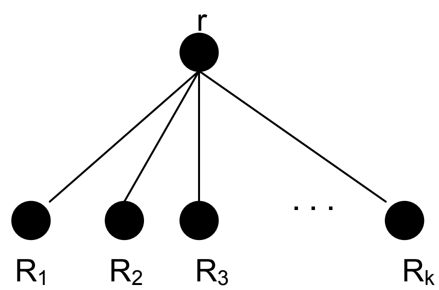
Definición.- Sea G = (V, A) un grafo no dirigido. G se denomina árbol si es conexo y no contiene ciclos.

Como un lazo es un ciclo de longitud uno, un árbol no tiene lazos.

Cuando un grafo es no conexo no puede ser un árbol, pero cada componente del grafo es un árbol y se denomina bosque.

Cuando un grafo es un árbol se escribe R en lugar de G para destacar su estructura.

Definición.- Sea R = (V, A) un árbol con raíz r. Si R no tiene otros vértices, entonces la raíz misma constituye el recorrido en orden previo, simétrico y posterior de R. Si |V| > 1, sean R_1, R_2, R_3, . . . , R_k los subárboles de R según se va de izquierda a derecha.

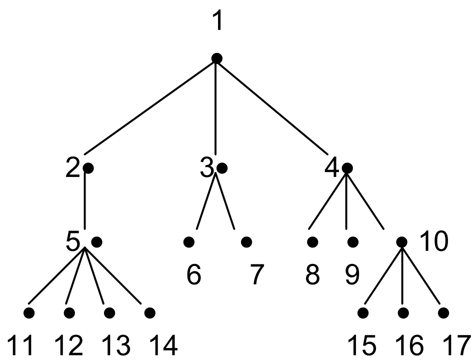


a) El recorrido en orden previo de R comienza en r y después pasa por los vértices de R_1 en orden previo, a continuación por los vértices de R_2 en orden previo, y así sucesivamente hasta que se pasa por los vértices de R_k en orden previo.

b) El recorrido en orden simétrico de R primero, se pasa por los vértices de R_1 en orden simétrico, después por la raíz r y a continuación por los vértices de los subárboles R_2, R_3, . . . , R_k en orden simétrico.

c) El recorrido en orden posterior de R pasa por los vértices de los subárboles $R_1, R_2, \ldots, R_k$ en orden posterior y a continuación por la raíz.

Ejemplo:



a) Recorrido en orden previo: 1, 2, 5, 11, 12, 13, 14, 3, 6, 7, 4, 8, 9, 10, 15, 16, 17.

b) Recorrido en orden simétrico: Comenzando en el vértice 1, se recorren los vértices del subárbol R1, con raíz del vértice 2 en orden simétrico. Esto lleva al vértice 5 y después al 11, una hoja. De ahí que al pasar en orden simétrico por los vértices del subárbol con raíz del vértice 5, se comience listando los vértices visitados como 11, 5, 12, 13, 14. A continuación, se pasa por el vértice 2, la raíz del subárbol R1 y después por la raíz del vértice 1. Para completar este recorrido, ahora se debe pasar por los vértices de los subárboles R2 y R3 en orden simétrico. Para R2 esto lleva del vértice 6 al vértice 3 (la raíz de R2) y después al vértice 7. Por último, para el subárbol R3 con raíz en el vértice 4 primero, se pasa por el vértice 8, después por la raíz en el vértice 4 y s continuación, por el 9. Después de que se recorre el subárbol con raíz del vértice 10; resulta el listado en orden simétrico 15, 10, 16 y 17. En consecuencia, el recorrido en orden simétrico del árbol determina la sucesión (en orden simétrico) 11, 5, 12, 13, 14, 2, 1, 6, 3, 7,8, 4, 9, 15, 10, 16, 17 para los vértices.

c) Recorrido en orden posterior: Para el recorrido en orden posterior de un árbol se comienza en la raíz y se construye el camino más largo, yendo al hijo situado más a la izquierda de cada vértice interno al que se llegue. Al llegar a una hoja h se pasa por este vértice y se retrocede hasta su padre p. Sin embargo, no se pasa por p hasta después de pasar por todos sus descendientes. El siguiente vértice por el que se pasa se halla aplicando el mismo procedimiento a p que a r, para obtener h. Nunca se pasa por un vértice más de una vez o antes que por sus descendientes. El recorrido en orden posterior pasa por los vértices en el orden 11, 12, 13, 14, 5, 2, 6, 7, 3, 8, 9, 15, 16, 17, 10, 4, 1.

Definición.- Si G = (V, A) es un grafo no dirigido, aparece una coloración apropiada de G cuando se colorean los vértices de G de modo que si {a, b} es una arista de G, entonces a y b se pintan con colores distintos. (De ahí que vértices adyacentes tengan colores distintos).

El número mínimo de colores necesarios para colorear de forma apropiada G se denomina **número cromático** de G y se escribe γ (G).

- **Vértices de corte**

Definición.- La división, o separación, de un vértice en un grafo da lugar a un nuevo grafo con más componentes, ese vértice se denomina **vértice de corte** o punto de articulación.

Se llaman vértices los círculos o puntos de un grafo o árbol y se les llama aristas a las líneas que unen un vértice con otro vértice.

D. Teoría matemática de la computación

1. Máquinas finitas

Autómatas Finitos Deterministas

Se llama Autómata Finito Determinista (AFD) a la quíntupla:

(Σ, Q, f, q_0, F)

- Σ es un alfabeto, llamado "alfabeto de entrada".
- Q es un conjunto finito, no vacío llamado "conjunto de estados".
- f es una función $f: Q \times \Sigma \to Q$ que se llama "función de transición".
- $q_0 \in Q$ es el "estado inicial".
- $F \subseteq Q$ es el conjunto de "estados finales", o "estados de aceptación", no vacío.

Un AFD puede considerarse como una máquina secuencial de Moore, cuyo alfabeto de entrada sea Σ, su alfabeto de salida $\Sigma_S=\{0,1\}$, y donde F será el conjunto de los estados tales que g(q)=1.

Tabla de Transición

Será una tabla cuyas filas están encabezadas por los estados (elemen-tos de Q). Los encabezamientos de las columnas son los símbolos del alfabeto de entrada (los elementos de). Cumpliéndose que el elemento i, j de la tabla de transición corresponde al valor de $f(q_i, e_j)$, donde q_i es el elemento i-ésimo de Q, y e_j es el elemento j-ésimo de . Tanto el estado inicial como el final estarán marcados por y por * respectivamente. *Nota*: el estado final puede ser indicado también rodeando el estado por un círculo.

	( ): $a_1 ... a_n$
** q_0** **.....** **q_i** **......** **$*q_f$**	

Ejemplo:

AF=(0,1 , q_0,q_1,q_2 , f, q_0, q_1)

f	**0**	**1**
** q_0**	q_1	q_0
$*q_1$	q_0	q_2
q_2	q_1	-

Diagrama de Transición

Es un grafo dirigido que se forma de la siguiente manera:

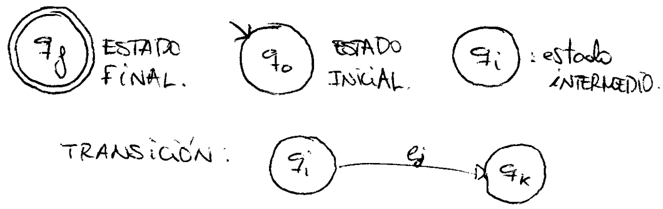
1. El grafo tendrá tantos nodos como Q , cada nodo estará etiquetado por un elemento de Q.

1. Si $f(q_i, e_j) = q_k$, dibujaremos una rama dirigida desde el nodo de etiqueta q_i hasta el nodo de etiqueta q_k. La etqueta de la rma será e_j.

2. El estado inicial estará señalado mediante el símbolo .

3. Los estados finales estarán señalados mediante el simbolo *, o doble círculo alrededor de la etiqueta del estado final.

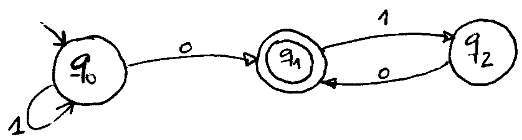
Representamos los estados como:



Lenguaje asociado a un autómata finito determinista L(AFD)

Sea un AFD (, Q, f, q_0, F). Decimos qie una palabra x   * es "aceptada" o "reconocida" por el autómata si f(q_0,x) F. Se llama lenguaje asociado al autómata finito, o conducta del autómata finito al conjunto de todas las palabras aceptadas por éste. Es decir: L = x x * & f(q_0,x) F

Ejemplo:



0 L(AF), ya que f(q_0, 0) = q_1 F

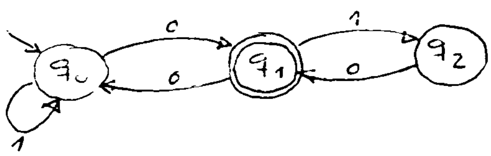
10 L(AF)

1*0 L(AF)

1*010 L(AF)

Luego L(AF) = 1*0(10)*

Ejemplo 2 :



Calculo posibles caminos desde q_0 a q_0 (por ser el estado inicial): (1*+(10)*0)*

Caminos desde el estado inicial hasta el estado final (el mas corto): 0

Caminos desde el estado final al estado final sin pasar por el inicial: (10)*

Luego L(AF)= (1*+(10)*0)* 0 (10)*

Estado Accesible

Un estado de un AF es accesible si existe una palabra x que permite que el autómata se pare en ese estado partiendo del estado inicial. Es decir, un estado *q* es accesible desde otro estado *p* del autómata si existe una cadena x que permite que el autómata transitando desde p con x se pare en el estado q:

$q \in Q$ es accesible si $\exists\, x \mid x \in \Sigma^* \mid q = f(q_0, x)$

Por definición de función de transición, todo estado es accesible desde sí mismo, y es accesible mediante la palabra vacía.

LEMA del estado accesible

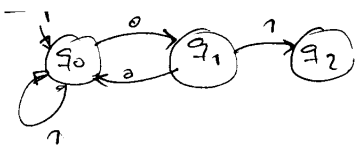
Dado un AF que tiene n estados, |Q|= n, un estado $q \in Q$ es accesible desde otro estado p si y sólo si $\exists\, x \in \Sigma^* \mid |x|<n$ (siendo n el nº de estados de autómata) | f(p,x)=q.

Supongamos |x|>=n | f(p,x=q $\Rightarrow$ siempre es posible escontrar otra cadena $x' \in \Sigma^*$ | |x|<n | f(p,x')=q.

Es decir, que si yo tengo un autómata con n estados, si transito por todos los estados, pasando una vez por cada estado, habré realizado n-1 transiciones, que es menos que n estados.

Si |x|>=n para la transición se repetirá al menos un estado.

Ejemplo :



q_2 es accesible desde q_0, por ejemplo con la palabra: x=110001, |x|=6, como |Q|=3, decimos que siempre podremos encontrar una palabra x' tal que |x'|<|x| (puesto que |x|>|Q|)

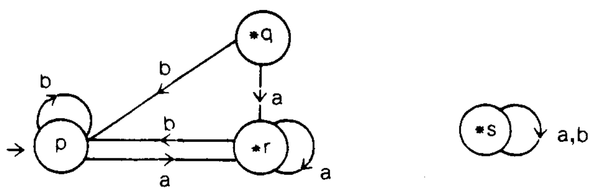
y ademá cumple que |x'|<|Q|. Esta x' sería: x'=01

Autómatas conexos

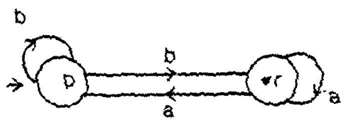
Sea (Σ , Q, f, q_o, F) un autómata finito determinista. Diremos que el autómata es "conexo" si todos los estados de Q son accesibles desde el estado inicial q_o.

Dado un autómata no conexo, podemos obtener a partir de él otro autómata que sea conexo eliminando todos los estados que no sean accesibles desde qo. Es evidente que los dos autómatas aceptarán el mismo lenguaje.

Ejemplo: sea el autómata definido por el diagrama de transición siguiente:



Es evidente que los estados q **y** s son inaccesibles desde el estado inicial p. Por lo tanto, el autómata siguiente aceptará el mismo lenguaje:



Minimización de AF

Equivalencia de estados en AF

Sea el autómata finito determinista (□ , Q, f, q_0, F). Decimos que dos estados p,q□ Q son equivalentes (se representa por pEq) si para toda palabra x□ □ *, se verifica que f(p,x)□ F□ f(q,x)□ F.

Nota: Si pEq entonces pE_nq, para todo n.

Propiedades de la equivalencia de estados

Dado AF=(□ , Q, f, q_0, F), con p,q□ Q

1. , pEq si x□ □ * □ f(p,x)□ F□ f(q,x)□ F
1. pE_kq (p y q son k-equivalentes) si: x□ □ * | |x|<=k □ f(p,x)□ F□ f(q,x)□ F
2. Ambas relaciones de equivalencia E y E_k inducen una partición de equivalencia sobre el autómata finito. Es decir: E induce P_E, y E_k induce P_k.
3. La primera partición que se establece es P_0, se hace con: p E_0q □ (p□ F□ q□ F). Nota: Entonces P_0=□ (Q-F),F□
4. E_k,si pE_kq □ pE_vq v<k; si pEq □ k pE_kq.

LEMA1(propiedad de estados)

Si $pE_{k+1}q$ □ pE_kq y f(p,a)E_kf(q,a) a□ □

LEMA2

$P_k=P_{k+1}$ □ $P_k= P_E$

LEMA3 (Se va a establecer cuando se estabiliza el valor de la partición)

Si |Q| = n >1 □ pE_nq □ $pE_{n-2}q$

Es decir, n-2 es el grado de equivalencia menor al que habría que llegar para conseguir confirmar la equivalencia de dos estados en un autómata con n estados.

TEOREMA

Si |Q|=n>1 □ j<=n-2 | $P_j= P_{j+1}$

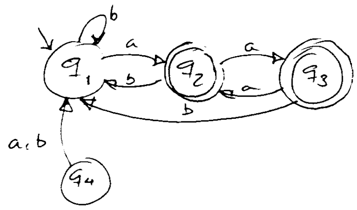
Algoritmo para construir P_E

1. P_0=□ □ q_i□ F□ ,□ q_i□ F□ □

1. P_{n+1}=□ $pE_{n+1}q$ □ pE_nq y $f(p,a)E_nf(q,a)$ a□ □ □

2. $P_n = P_{n+1} = P_E$

Ejemplo :



Para hallar las partes de equivalencia del autómata primero debemos partir de un autómata finito conexo. Para ello vamos a eliminar q_4, ya que no es posible acceder a él desde el estado inicial (q_1). Ahora vamos a establecer las partes de equivalencia:

□ P_0=□ Q-F, F□ =□ □ q_1□ , □ q_2, q_3□ □

Ahora hacemos P_1, para ello sabemos que □ q_1□ va a seguir igual, es decir, va a ser equivalente con sigo mismo, puesto que accede a los mismo estados(finales o no finales).

Miramos □ q_2, q_3□ y comprobamos si acceden al mismo tipo de estado para las mismas entradas, es decir, sabiendo que $q_2 E_0 q_3$ vamos a ver si $q_2 E_1 q_3$:

$f(q_2,a)=q_3$□ F $f(q_2,b)=q_1$□ F

$f(q_3,a)=q_2$□ F $f(q_3,b)=q_1$□ F

luego $q_2 E_1 q_3$.

P_1=□ □ q_1□ , □ q_2, q_3□ □ = P_0= P_E

Algoritmo para hallar el autómata mínimo de uno dado

1. Calcular el autómata conexo

1. Construir Q/E del autómata conexo obtenido en el punto anterior.

2. Calcular A'=(□ ,Q/E,f',C_0,F'), donde:

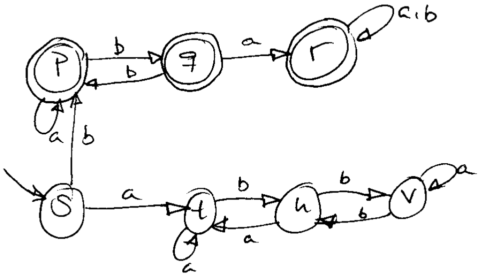
a. *C_0* es la clase donde se encuentra q_0 ;

b. *F'*=□ c | c contiene al menos un estado de F(es decir, existe un q□ F, tal que q□ c)□ ;

c. *f'*: Q/E x □ □ Q/E, donde

f'(c_i, a)= c_j □. q_i □ c_i y. q_j □ c_j | $f(q_i,a)=q_j$

Ejemplo:



Es un autómata conexo, pues todos sus estados son accesibles, vamos a hallar las partes de equivalencia:

P_0=□ □ s,t,u,v□ ,□ p,q,r□ □ (=□ Q-F,F□)

miramos la equivalencia de las clases de equivalencia establecidas en P_0.

f(p,a) = p□ F; f(p,b)= q□ F;

f(q,a)= r□ F; f(q,b)=p□ F;

f(r,a)= r□ F; f(r,b)=r□ F;

aqui vemos que □ p,q,r□ se conserva pues tienen iguales transiciones a estados finales a partir de las entradas.

f(s,a)=t□ F; f(s,b)=p□ F

f(t,a)=t□ F; f(t,b)=u□ F

f(u,a)=t□ F; f(u,b)=v□ F

f(v,a)=v□ F; f(v,b)=u□ F

ahora vemos que mientras t,u,v tienen iguales transiciones a estados no finales para las mismas entradas, siguen siendo equivalentes, mientras que s para b transita a un estado final, ya será equivalente a los tres estados anteriores:

P_1=□ □ p,q,r□ ,□ s□ ,□ t,u,v□ □

Ahora hallamos P_2:

f(p,a) = p□ F; f(p,b)= q□ F;

f(q,a)= r□ F; f(q,b)=p□ F;

f(r,a)= r□ F; f(r,b)=r□ F;

f(t,a)=t□ F; f(t,b)=u□ F

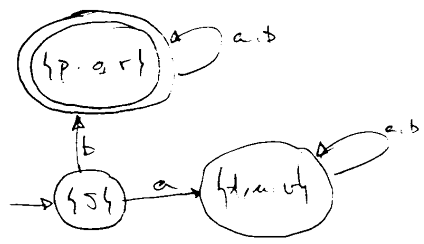
f(u,a)=t□ F; f(u,b)=v□ F

f(v,a)=v□ F; f(v,b)=u□ F

P_2= □ □ p,q,r□ ,□ s□ ,□ t,u,v□ □ = P_1

luego $P_E = \{\{p,q,r\},\{s\},\{t,u,v\}\}$

El autómata mínimo nos va a quedar:



Si nos pidieran el lenguaje del autómata: $L(A)=L(A')=\{b(a+b)^*\}$, (donde A es el autómata inicial, y A' es el autómata mínimo).

Autómatas Equivalentes

Suma directa de autómatas finitos deterministas

Sean los AFD $A_1=(\Sigma, Q_1, f_1, q_{01}, F_1)$ y $A_2=(\Sigma, Q_2, f_2, q_{02}, F_2)$, tales que Q_1 y Q_2 no tienen elementos comunes. Llamamos suma directa de A_1 y A_2 al autómata $A=A_1+A_2=(\Sigma, Q_1 U Q_2, f, q_0, F_1 U F_2)$, dinde q_0 es uno cualquiera de los dos estados q_{01} o q_{02}, y f se define así:

$f(q,a)=f_1(q,a)$ si $q \in Q_1$

$f(q,a)=f_2(q,a)$ si $q \in Q_2$

Equivalencia de autómatas

Sean los AFD $A_1=(\Sigma, Q_1, f_1, q_{01}, F_1)$ y $A_2=(\Sigma, Q_2, f_2, q_{02}, F_2)$. Decimos que dos estados $p \in Q_1$ y $q \in Q_2$ son equivalentes si $f(p,x) \in F_1 \Leftrightarrow f(q,x) \in F_2$, para todo $x \in \Sigma^*$.

Decimos que los dos autómatas son equivalentes si reconocen el mismo lenguaje. Es decir: si $f(q_{01},x) \in F_1 \Leftrightarrow f(q_{02},x) \in F_2$, para todo $x \in \Sigma^*$. Dicho de otro modo, decimos que dos autómatas son equivalentes si sus estados iniciales los son: $q_{01} E q_{02}$.

Teorema.- Sean dos autómatas finitos deterministas $A_1=(\Sigma, Q_1, f_1, q_{01}, F_1)$ y $A_2=(\Sigma, Q_2, f_2, q_{02}, F_2)$, tales que Q_1 y Q_2 no tienen estados comunes y $|Q_1|=n_1$ y $|Q_2|=n_2$. Entonces, A_1 y A_2 son equivalentes si q_{01} yq_{02} son equivalentes en el autómata $A=A_1+A_2$, es decir, si ambos autómatas aceptan las mismas palabras de longitud menor que n_1+n_2-1. Además, en general, n_1+n_2-2 es el valor más pequeño que cumple siempre este teorema.

ALGORITMO PARA VER SI DOS AUTÓMATAS SON EQUIVALENTES

1. Calcular la suma directa de autómatas
1. Construir P_E del autómata suma obtenido
2. Comprobar que los estados iniciales son equivalentes, es decir, están en la misma clase de equivalencia. Cuando se da la condición: $p_{01}, p_{02} \in C_0 \Rightarrow A_1$ y A_2 son equivalentes

Isomorfismo entre Autómatas

Se dice que A_1 es isomorfo a A_2, es decir, $A_1 \cong A_2$ si $\exists$ i $:Q_1 \rightarrow Q_2$, (i : imagen). Por lo tanto :

1. $i(p_{01}) = p_{02}$ (la imagen del estado inicial de A_1 es el estado inicial de A_2.

1. Dados p□ F_1, q□ F_2 : i(p)□ F_2, y i(q)□ F_1.(es decir, la imagen de los estados finales de uno de los autómatas, es un estado final del otro autómata).

2. $i(f_1(p_1,e)) = f_2(i(p_1),e) = f_2(i(p_2),e)$.La imagen de la transición es la transición de la imagen.

Por lo tanto A_1 y A_2 son iguales renombrando estados.

Entonces si dos autómatas son isomorfos van a ser equivalentes, es decir, los lenguajes que van a generar ambos autómatas van a ser el mismo : $L(A_1) = L(A_2)$. Con esto comprobamos que la isomorfía implica la equivalencia.

Teorema : Si dos autómatas son equivalentes, entonces sus autómatas mínimos son isomorfos, es decir : A_1EA_2 □ $\hat{A}_1$□ $\hat{A}_2$ (siendo $\hat{A}_i$ □ autómata mínimo de A_i).

Autómatas Finitos No Deterministas

Llamaremos "autómata finito no determinista" (AFND) a la séxtupla:

A = (□ , Q, f, q_0, F, T), donde:

- □ , Q,q_0, F significan lo mismo que en un autómata finito determinista.
- f es una aplicación de Q x □ en el conjunto de las partes de Q
- T es una relación definida sobre pares de elementos de Q, donde p, q □ Q están en relación T(se representará por pTq, o (p,q)□ T) si existe una transición del estado p al estado q por medio de la palabra vacía □ (□ -transición).

Ejemplo: Sea el autómata finito no determinista:

(□ a, b□ , □ p, q, r, s□ , f, p, □ p, s□ , □ (q,s), (r,r), (r,s), (s,r)□), donde f se define asi:

f(p,a)=□ q□ f(p,b)=□

f(q,a)=□ p,r,s□ f(q,b)=□ p,r□

f(r,a)=□ f(r,b)=□ p,s□

f(s,a)=□ f(s,b)=□

Representación

<u>En Diagrama</u>: Con tantos nodos como estados.

Pueden sobrar o faltar arcos de un estado a otro

Si $f(p,a)=Q_1$, trazaremos una rama dirigida desde p hasta cada uno de los estados del conjunto de estados Q_1, con la etiqueta a.

Pueden existir arcos de un estado a otro por medio de □ , en el caso de pTq, que trazaremos una rama desde p hasta q con la etiqueta □ .

Relación T

Sabemos que: (p,q) □ T*, donde T* = T^0 □ T^1 □ T^2 □ ....

Utilizaremos las *matrices booleanas de representación T*, para representar si un estado q es accesible desde un estado p mediante una □ -transición.

Primero se representará la matriz booleana de representación de T, o, lo que es lo mismo, la matriz booleana que representa la posibilidad de transitar a otro estado a partir de una □ -transición:

T	P	q	r	s
p	0	0	0	0
q	0	0	0	1
r	0	0	1	1
s	0	0	1	0

T*	P	q	r	s
p	0	0	0	0
q	0	0	1	1
r	0	0	1	1
s	0	0	1	1

Ahora se hace la clausura transitiva de la relación T, es decir, la relación T*, que no es más que calcular la accesibilidad de unos estados a otros mediante un número de □ -transiciones mayor.

Lo que se ha hecho en esta clausura no es más que calcular aquellos estados q, que a partir de un estado p, es posible acceder mediante □ -transiciones. Esto se puede ver tanto en el diagrama de estados, como en la matriz booleana de la relación T. Se puede observar que el estado r no accesible desde q mediante una □ -transición, sin embargo si será accesible mediante dos □ -transiciones. Este también será el caso del estado s y su acceso a sí mismo mediante □ -transiciones.

Extensión de f a palabras

Ahora vamos a definir f’ como una nueva función de transición que, en lugar de actuar sobre letras del alfabeto de entrada, actúe sobre palabras (que incluyan la palabra vacía □). Es decir: f’: Q x □ * □ P(Q) (P(Q), partes de Q)

a) f’(q,□) = □ p □ q T* p□ , para todo q□ Q

b) Sea x = $a_1a_2...a_n$, n > 0 . Entonces, f’(q,x) = □ p □ p es accesible desde q mediante la cadena de entrada □ ^{i0}a_1□^{i1}a_2... a_n□^{in}□ para todo q□ Q (la sucesión anterior es idéntica a x).

Ejemplo (en el autómata anterior):

f’(p,□) = □ p□ f’(q,□) = □ q,r,s□

f’(r,□) = □ r,s□ f’(s,□) = □ r,s□

f’(p,a) = □ q,r,s□ f’(p,b) = □

f’(q,a) = □ p,r,s□ f’(q,b) = □ p,r,s□

f’(r,a) = □ f’(r,b)=□ p,r,s□

f’(s,a) = □ f’(s,b) = □ p,r,s□

f’(p,aa) = □ p,r,s□ f’(p,ab) = □ p,r,s□

f’(q,aa) = □ q,r,s□ f’(r,ba) = □ q,r,s□

Lenguaje aceptado por un AFND

El L(AFND) estará formado por el conjunto de cadenas que acceden a un estado final: L(AFND) = □ x □ x □ □ * □ f (q_0, x) = C_i □. q_f □ C_i □ q_f □ F □

La cadena vacía □ , es aceptada por el AFND , cuando el estado final es un estado final: □ □ L(AFND) si q_0 □ F □ q_0 T* q_f □ q_f □ F (esto quiere decir que si □ pertenece al lenguaje asociado al AFND entonces: O el estado inicial es también un estado final, o que el estado final es accesible mediante □ -transiciones desde el estado inicial)

Autómata finito determinista equivalente a uno no determinista

Dado un autómata finito no determinista, siempre es posible construir otro finito determinista que acepta el mismo lenguaje que el primero. Puesto que el conjunto de lenguajes aceptados por los autómatas finitos no deterministas es idéntico al conjunto de los lenguajes finitos deterministas: L(AFD) = L (AFND)

Sea el autómata finito no determinista A = (□ , Q, f, q_0, F, T). Definiremos a partir de él el siguiente autómata finito determinista: B = (□ , Q’, f’, q'_0, F’) donde:

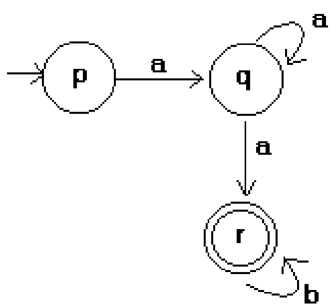
- Q’ = P(Q), el conjunto de las partes de Q, que incluye a Q y al conjunto vacío, □ .
- q'_0 = f’(q_0, □) (Estados a los que podemos ir desde q_0 mediante □ -transiciones)
- F’ = □ c □ c□ Q’ y existe un q □ c tal que q □ F□ .
- f’(c,a) = □ c’ □ c’ = $U_{q□ c}$ f’(c,a) □ . Es decir, como c es un subconjunto de Q, f’(c,a) es el subconjunto de Q formado por todos los estados a los que puede ir el autómata A, partiendo de un estado c, mediante la entrada a.

Ejemplo :

Sea el autómata finito no determinista definido por la tabla siguiente :

f	**a**	**b**
□ **p**	p	
q	q, r	
***r**		r

tendrá el diagrama siguiente :



□ El autómata finito determinista equivalente será (por pasos) :

1º) p será el estado inicial

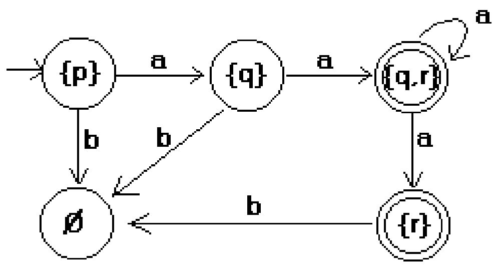
2º) (hallamos las partes de Q)

P(Q) = □ □ , □ p□ , □ q□ , □ r□ , □ p, q□ , □ q, r□ , □ r, p□ , □ p, q, r□ □

3º) Hago una tabla de transiciones, y pongo en la columna de estados aquellos que sean accesibles desde el estado inicial, no se ponen todas las partes de Q (P(Q)). Podría hacerse primero de esa forma, poniendo en los estados P(Q), pero después tendríamos que hallar el autómata conexo asociado.

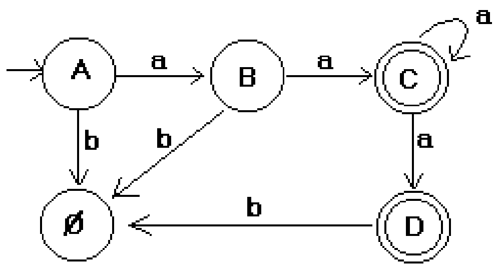
	a	**b**
□ □ **p**□	□ p□	□
□ **q**□	□ q, r□	□
*□ **r**□	□	□ r□
*□ **q,r**□	□ q,r□	□ r□

Y este es el autómata finito determinista equivalente al no determinista.



Si nos pidieran minimizar el AFD obtenido:

Renombramos:



Hallamos el autómata equivalente hallando la partición de equivalencia correspondiente:

P_0=□ □ A,B,E□ , □ C,D□ □

P_1=□ □ A,E□ , □ B□ , □ C□ , □ D□ □

P_2=□ □ A□ , □ B□ , □ C□ , □ D□ , □ E□ □ = P_E luego el autómata ya es mínimo

Autómata finito no determinista equivalente a uno determinista

Un AFD es igual a un AFND sin □ -transiciones.

Autómata Finito asociado a una Gramática Regular(Gramática tipo 3)

Notas

- El alfabeto no terminal de la Gramática y el conjunto de estados del AF van a "memorizar" el proceso.
- El conjunto de producciones de P "marchará en paralelo" con la función de transición f.
- En las gramáticas hemos hablado de depuración, lo cual podría considerarse como un símil a la minimización en un AF:

Gramáticas	Autómatas
Quitamos variables inaccesibles	Creamos autómata finito conexo
Variables no generativas	Estados trampa

Producciones reductoras (A□ □)	□ -transición a estado final (A □ □ q_f)
Producción A□ B	□ -transición(A□ □ B)
Producción a□ A	□ -transición(A□ □ A)

- Cuando tenemos un AFD vamos a tener una gramática depurada (excepción de GRLI)
- De un AFND □ gramática ambigua
- De un AFND □ AFD □ gramática no ambigua
- Dada una gramática ambigua, a partie de ella podemos obtener un AFND de donde puedo obtener el AFD y a partir de aquí una gramática no ambigua.

2.Lenguajes formales

- Gramáticas formales: definiciones, operaciones, tipos de lenguajes, ambigüedad, equivalencia, la jerarquización de Chomsky (R)

J. Lenguajes formales

La idea básica es considerar a un lenguaje como un conjunto compuesto por cadenas de longitud finita formadas por símbolos tomados de un alfabeto. Es decir, dado un alfabeto $\Sigma = \{\sigma_1, \sigma_2, ...\sigma_n\}$, decimos que $x = \sigma_{i_1}\sigma_{i_2}...\sigma_{i_m}$, (para un m dado, que llamamos longitud de x) es una cadena construida con símbolos de Σ. Si $m = 0$tenemos la cadena llamada *vacia*, y denotada por Λ. Al conjunto formado por de todas las cadenas que se pueden construir con las letras de dicho alfabeto se le llama lenguaje universal y se le denota por Σ^* . Es evidente que el número de elementos de Σ^* es infinito numerable. Entre las cadenas, como elementos de Σ^* , se define una operación, llamada *concatenación*, mediante la cual a dos cadenas dadas $x, y \in \Sigma^*$ se le asocia otra cadena $z$obtenida por yuxtaposición de x, y; es decir, si $x = x_1x_2...x_p \text{ con } x_i \in \Sigma^*$, e $y = y_1y_2...y_q \text{ con } y_j \in \Sigma^*$, entonces la cadena $z = xy$, sería : $z = x_1x_2...x_py_1y_2...y_q$.

Llamamos *lenguaje* construido con el alfabeto Σ, a cualquier subconjunto L de Σ^*. Diremos que un lenguaje es finito, si es finito el número de cadenas de que consta; en otro caso diremos que es infinito. El *lenguaje vacio* no tiene ninguna cadena (que no debe confundirse con el lenguaje que solo tiene la cadena vacia).

Una vez dada esta definición tan general de lenguaje, el problema inicial que se nos presenta es como distinguir con precisión las cadenas que pertenecen a un lenguaje L de aquellas que no pertenecen. Hay dos formas de hacerlo: una consiste en dar las reglas mediante las cuales construir las cadenas que forman el lenguaje, y otra es dar un procedimiento para determinar si una cadena dada pertenece o no al lenguaje considerado. En el primer caso se trata de definir las *gramáticas formales*, o generadores, y en el otro los *autómatas* o reconocedores.

Para fijar un poco las ideas de lo que se entiende por lenguaje definido por una gramática formal, demos ligeramente algunos detalles técnicos. Una gramática es un procedimiento finito para formar o generar cualquiera de todas las posibles cadenas de un lenguaje. Formalmente diremos que una gramática es un sistema :

$$G = (V_N, V_T, P, S)$$

Donde V_N es un alfabeto llamado *vocabulario no terminal*; V_T es un alfabeto, llamado *vocabulario terminal*. Las reglas de P son de la forma $\alpha_i \to \alpha_j$, donde

$$\alpha_i \in (V_N \cup V_T)^+, \qquad \alpha_j \in (V_N \cup V_T)^*$$

y S, un símbolo de V_N, llamado símbolo inicial. Para generar una palabra mediante la gramática G, nos apoyamos en secuencias de cadenas de símbolos, $v_1, v_2, ...v_m$. Estas secuencias (cada una de las cuales es llamada *derivación*), se construyen del siguiente modo:

1:

El primer elemento de la derivación es siempre el símbolo inicial S (es decir, siempre $v_1 = S$).

2:

A partir de un elemento v_i se pasa al siguiente v_{i+1}, (donde v_i y v_{i+1} pertenecen a $(V_N \cup V_T)^+$) , mediante una regla de P, es decir, si $v_i = \delta\alpha_i\delta'$, y la regla $\alpha_i \to \alpha_j \in P$, entonces será $v_{i+1} = \delta\alpha_j\delta'$, (escribiremos $v_i \Rightarrow v_{i+1}$)

3:

Si v_m, obtenida aplicando m veces el procedimiento anterior, pertenece a V_T^*, entonces decimos que v_m es una palabra generada por G.

A la derivación $S = v_1, v_2, ...v_m$, se la puede expresar:

$$S \Rightarrow v_1 \Rightarrow v_2 \Rightarrow .... \Rightarrow v_m,$$

o simplemente

$$S =* \Rightarrow v_m$$

que podría leerse: v_m se obtiene por derivación de S aplicando reglas de G, o simplemente: v_m *en una derivada de* S.

Podemos sintetizar lo anterior diciendo que el lenguaje generado por la gramática G, es:

$$L(G) = \{x | x \in V_T^*, S = * \Rightarrow x\}$$

Decimos que dos gramáticas son equivalentes, si y solo si generan el mismo lenguaje, es decir

$$G_1 \equiv G_2 \Leftrightarrow L(G_1) = L(G_2)$$

II. Sistemas formales

1. Maquinas de Turing.

Otra forma de caracterizar la clase de las funciones computables se hace mediante las llamadas *máquinas de Turing* [33]. Estas son máquinas formales, es decir sin cables ni componentes físicos. Su finalidad es definir los cálculos a partir de las operaciones más sencillas posibles, y utilizando las cuales calcular efectivamente funciones dadas. Las funciones que así pueden realizarse se llaman *funciones calculables* o *computables*.

Este tipo de máquina consta hipoteticamente de una unidad de control capáz de interpretar las instrucciones que reciba, y de una cabeza lectora que permite leer el contenido de una de las casillas en que esta dividida una memoria lineal, ilimitada en ambas direcciones de sus extremos.

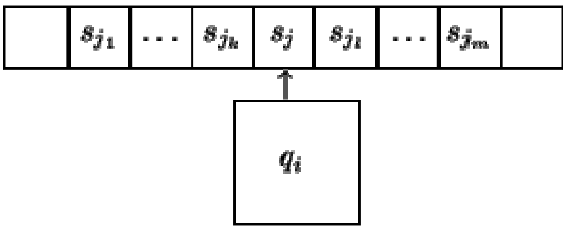


Figura 4

La máquina en su funcionamiento pasa por diferentes estados q_i en instantes t sucesivos. El argumento de la función que queremos calcular estará almacenado previamente en la memoria, y en el instante inicial, antes de que la máquina comience a funcionar, la cabeza lectora apuntará al símbolo más a la izquierda del argumento. A partir de ese momento la máquina realizará operaciones, que dependerán del estado en que ella se encuentre, y del símbolo que en ese momento lea la cabeza lectora. Con estas operaciones se podrán realizar las siguientes áreas: sustituir el símbolo leído por otro, pasar a leer el símbolo que esta en memoria a la derecha del símbolo leído, pasar a leer el símbolo que esta en memoria a la izquierda, o saltar directamente a ejecutar otra instrucción si se cumple una condición especificada; en todos los casos, y una vez ejecutada la tarea indicada, se pasaría al estado que también se indica en la propia instrucción.

Estas instrucciones, o cuaternas, las podemos representar e interpretar así:

$q_i\ s_j\ s_k\ q_l$:	si estando en el estado q_i se lee el simbolo s_j, entonces s_j se cambia por s_k y se pasa al estado q_l
$q_i\ s_j$ R q_l :	si desde q_i se lee s_j, entonces se pasa a leer la casilla situada a su derecha y se cambia al estado q_l
$q_i\ s_j$ L q_l :	si desde q_i se lee s_j, entonces se pasa a leer la casilla a su izquierda y se cambia al estado q_l
$q_i\ s_j\ q_k\ q_l$:	si desde q_i se lee s_j, entonces se cambia al estado q_k o al q_l según una cierta condición

donde: q_i indica un estado de la máquina tomado del conjunto de estados Q; s_j indica uno de los símbolos que pueden aparecer en la memoria tomado del alfabeto S; Res un símbolo que indica pasar a leer la casilla de memoria situada a la derecha; Les un símbolo que indica pasar a leer la casilla de memoria situada a la izquierda.

Se define una *máquina de Turing* (M) como un conjunto finito, no vacío, de cuaternas, de forma que no contenga dos cuaternas con los dos primeros símbolos correspondientes iguales.

Indicaremos por $Q = \{q_i\}$ al conjunto de estados de la máquina, por $S = \{s_j\}$ al alfabeto de símbolos de la memoria.

En Q siempre existe un estado particular, que se suele indicar por q_0 y llamarse *estado inicial*, en el que se supone está la máquina al comenzar a operar. En el alfabeto S siempre figurarán, entre otros posibles, dos símbolos: el blanco (B), y el uno (expresado por un palote $|$) . A las cadenas formadas por símbolos de este alfabeto se las llama *expresiones de memoria*. Si en una expresión de memoria incluimos un único símbolo q_i (que indique un estado de Q) siempre que no le situemos a la derecha de cualquier otro, obtendremos una expresión que llamaremos *descripción instantánea* de la maquina M .

Los números naturales se representan en estas máquinas, de la forma más simple, es decir usando un solo símbolo ($|$). Se emplean dos clases de representación numérica; una llamada *de entrada* mediante la cual con un solo palote representaremos el cero, y cualquier otro número $n$se representará por $n+1$ palotes; y la otra, llamada *de salida*, consiste en contar el número de palotes que hay en la expresión de memoria (ningún palote se interpretaría, en este caso, como el 0).

La descripción instantánea de la maquina Men el instante inicial podría representarse por

$$q_0\, s_{i1}\, s_{i2}\, s_{i3} \ldots s_{in}$$

En cualquier otro instante una descripción instantánea sería de la forma

$$\alpha = s_{j1} \ldots s_{jk}\, q_i\, s_j s_{jl} \ldots s_{jm}$$

Para ver como funciona una máquina de Turing, veamos como se realizan las transiciones de unas descripciones instantaneas a otras. Dada una descripción instantanea α, en la que aparezcan el par de símbolos $q_i\, s_j$, buscaríamos la cuaterna de la máquina que comience con estos dos símbolos y operaríamos en consecuencia, es decir, para cada uno de los tipos de cuaternas, construiríamos la descripción instantánea siguiente, β, como indicamos a continuación:

para $q_i s_j s_k q_l$ sería $\beta = s_{j_1} \ldots s_{j_k} q_i s_k s_{j_l} \ldots s_{j_m}$

para $q_i s_j R q_l$ sería $\beta = s_{j_1} \ldots s_{j_k} s_j q_i s_{j_l} \ldots s_{j_m}$

para $q_i s_j L q_l$ sería $\beta = s_{j_1} \ldots q_i s_{j_k} s_j s_{j_l} \ldots s_{j_m}$

para $q_i s_j q_k q_l$ sería

$$\beta = s_{j_1} \ldots s_{j_k} q_k s_j s_{j_l} \ldots s_{j_m}$$

si el número representado por cinta en ese instante *pertenece* a un conjunto dado A, o bien

$$\beta = s_{j1} \ldots s_{jk} q_l s_j s_{jl} \ldots s_{jm}$$

si el número representado por la cinta en ese instante *no pertenece* a un conjunto dado A

En el último caso aparece el conjunto A, que es un conjunto de números naturales, que no forma parte de la máquina, y al que hay que consultar antes de tomar la decisión sobre una de las alternativas propuestas (es el ``oráculo" a que nos referiamos en una nota de la pagina anterior).

Si no existiese en M ninguna cuaterna que comenzase por $q_i s_j$ diríamos que la descripción instantánea α es *terminal*.

En cualquiera de los casos anteriores decimos que de la descripción instantanea α se pasa a la β mediante la maquina de Turing M, lo que se denota $\alpha \rightarrow \beta(M)$.

Llamamos computación o cálculo de una máquina de Turing M, a una sucesión finita de descripciones instantáneas $\alpha_1, \alpha_2, \ldots \alpha_n$, de forma que para $1 \leq i < n$, sea $\alpha_i \rightarrow \alpha_{i+1}(M)$, y sea α_n, una descripción instantánea terminal.

Para calcular una función $f : N^n \rightarrow N$, mediante una maquina de Turing, se comenzará representando la n-upla de los argumentos mediante una expresión de memoria de la forma $||..|B||..|B||..|B...B||..|$, a partir de la cual se formará la descripción instantánea inicial

$$\alpha_1 = q_0||..|B||..|B||..|B...B||..|,$$

a la que se aplicaran sucesivamente las cuaternas que correspondan de la maquina de Turing hasta llegar a una descripción instantánea terminal. En esta contamos el numero de palotes que contiene, y este será el valor de la función. [35]

Dada una función $f : N^n \rightarrow N$, no siempre existe una Máquina de Turing, mediante la cual, a partir de cualquier argumento, podamos llegar desde el estado inicial a una descripción instantánea terminal; cuando existe una tal máquina decimos que la función es *computable*.