ANEXO Tema 2: Conceptos básicos de probabilidad aplicada al Reconocimiento de Patrones

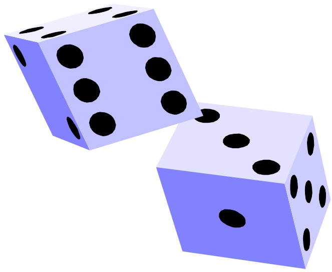
Grado en Ingeniería Informática

Curso 2013 / 2014

UCA Universidad de Cádiz

Contenidos

1. Conceptos
2. Probabilidad clásica
3. Distribuciones más frecuentes (discretas)
4. Distribuciones más frecuentes (contínuas)



Conceptos

Concepto de probabilidad

- **PROBABILIDAD:** es la disciplina científica que estudia las leyes del azar.

- ¿Cómo osamos hablar de leyes del azar? ¿No es, acaso, el azar la antítesis de cualquier ley?
Bertrand Russell

- Es un hecho destacable que una ciencia que empezó analizando juegos de azar acabe convirtiéndose en el más importante objeto del conocimiento humano.

  Pierre Simon Laplace

Concepto de probabilidad

- Problemas simples
  - Sacar cartas de un mazo.
  - Tirar una moneda.
  - Arrojar un dado.

- Problemas complejos
  - Genética
  - Bolsa
  - Elecciones generales
  - Física nuclear
  - Calificaciones de exámenes de algunas asignaturas
  - CASI CUALQUIER PROBLEMA DE LA NATURALEZA

Experimento aleatorio

- Cualquier situación que, realizada en las mismas condiciones, proporcione un resultado imposible de predecir *a priori*.
- Por ejemplo:
  - Lanzar un dado.
  - Extraer una carta de una baraja.
  - Se lanza una moneda. Si sale cara se extrae de una urna U1, con una determinada composición de bolas de colores, una bola y si sale cruz de extrae de una urna U2, con otra determinada composición de bolas de colores, una bola.

Experimento aleatorio

- Diremos que un **experimento** es **aleatorio** si se verifican las siguientes condiciones:
  - Se puede repetir indefinidamente, siempre en las mismas condiciones;
  - Antes de realizarlo, no se puede predecir el resultado que se va a obtener;
  - El resultado que se obtenga, ***e***, pertenece a un conjunto conocido previamente de resultados posibles. A este conjunto, de resultados posibles, lo denominaremos **espacio muestral**.

Espacio muestral

- Es la colección de posibles resultados del experimento

- EJEMPLOS:
  - Lanzar una moneda al aire y observar los resultados:

    E = { CARA(C) , CRUZ(X) }

  - Lanzar un dado:

    E = {1 , 2 , 3 , 4 , 5 , 6}

  - Lanzar dos monedas

    E = { CC , CX , XC , XX }

Espacio muestral

- Experimento: Lanzamiento de DOS dados

E = (1,1) , (1,2) , (1,3) , (1,4) , (1,5) , (1,6)
(2,1) , (2,2) , (2,3) , (2,4) , (2,5) , (2,6)
(3,1) , (3,2) , (3,3) , (3,4) , (3,5) , (3,6)
(4,1) , (4,2) , (4,3) , (4,4) , (4,5) , (4,6)
(5,1) , (5,2) , (5,3) , (5,4) , (5,5) , (5,6)
(6,1) , (6,2) , (6,3) , (6,4) , (6,5) , (6,6)

Concepto de evento (ó suceso)

- Evento: cualquier subconjunto de E.
- Ejemplos:
  - Cuando se lanza 3 veces una moneda, el espacio muestral es:

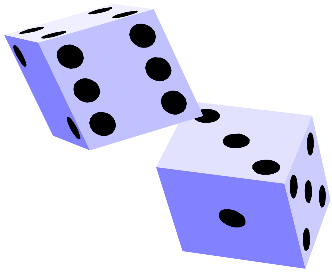
E ={ CCC, CCX, CXC, CXX, XCC, XCX, XXC, XXX }

  - El suceso o evento sale “al menos dos caras” es:

S ={ CCC, CCX, CXC, XCC }

  - El suceso aparece “al menos una cruz” es:

S = { CCX, CXC, CXX, XCC, XCX, XXC, XXX }



Probabilidad clásica

Probabilidad clásica

- Laplace define la probabilidad de un suceso A como:

$$P(A) = \frac{\text{Número de casos favorables}}{\text{Número de casos posibles}}$$

- Si lanzamos un dado ¿Cuál es la probabilidad $P(A)$ de A = *mayor o igual a 5*? ¿Y la probabilidad de B = *impar*?
- Solución: Los seis casos posibles son igualmente probables, cada uno tiene probabilidad 1/6.
- $P(A) = 2/6 = 1/3$ pues $A = \{5,6\}$ tiene dos casos favorables.
- $P(B) = 3/6 = 1/2$ pues $B = \{1, 3, 5\}$ tiene tres casos favorables

Axioma de probabilidad

- Se llama probabilidad a cualquier función P que asigna a cada suceso A del espacio muestral E un valor numérico $P(A)$, verificando los siguientes axiomas:

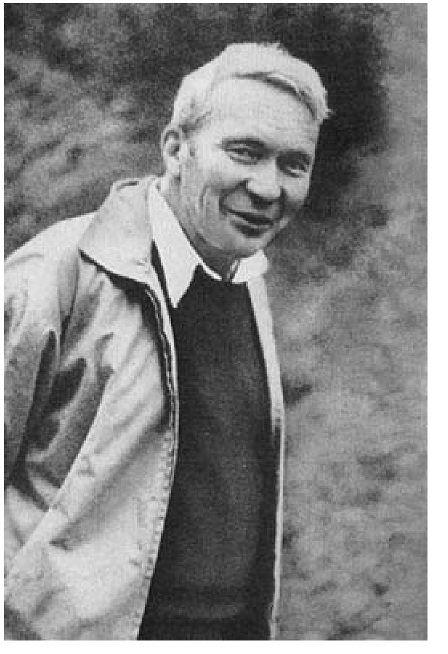
(1) No negatividad: $0 \leq P(A)$

(2) Normalización: $P(E) = 1$

(3) Aditividad: $P(A \cup B) = P(A) + P(B)$

si $A \cap B = \emptyset$

(donde $\emptyset$ es el conjunto vacío).



Kolmogorov, 1933

Función de probabilidad

- La función de Probabilidad asigna a cada valor de la variable su correspondiente probabilidad.

- En el experimento “Lanzar un dado”, la función de probabilidad f(k) es:

X	1	2	3	4	5	6
f(k)=P(X=k)	1/6	1/6	1/6	1/6	1/6	1/6

Función de probabilidad

- En el experimento "Lanzar un dado **trucado**", la función de probabilidad f(k) podría ser:

X	1	2	3	4	5	6
f(k)=P(X=k)	**1.5**/6	1/6	1/6	1/6	1/6	**0.5**/6

Caso continuo

- Problema: ¿qué hacer si los sucesos no son una variable discreta, sino continua?
- Ejemplos:
  - Sacar un pez del agua, y medir su longitud
  - La cotización en la bolsa de un valor
  - El tiempo transcurrido en una carrera de 1.500 m.
- En este caso, ¿no habría infinitos casos posibles?
- Para resolver el problema introducimos un nuevo concepto

Función de distribución de probabilidad

- Dada una variable aleatoria *X* la **función de distribución de probabilidad** *F*(*x*) asigna a un evento definido sobre *x* una probabilidad.

$$F(a) = P(X \leq a)$$

- F ha de cumplir que:
  - F ha de ser continua y monótona creciente

$$\lim_{x \to -\infty} F(x) = 0$$

$$\lim_{x \to \infty} F(x) = 1$$

Función de distribución de probabilidad

- Caso discreto

$$F(x) = \sum_{t=-\infty}^{x} f(t)$$

- Caso continuo

$$F(x) = \int_{t=-\infty}^{x} f(t)$$

Función de densidad de probabilidad

- Matemáticamente la función densidad de probabilidad es la derivada de F.
- La función densidad de probabilidad debe cumplir que f(x)>=0, y que la integral de f en [- ∞,∞] sea 1

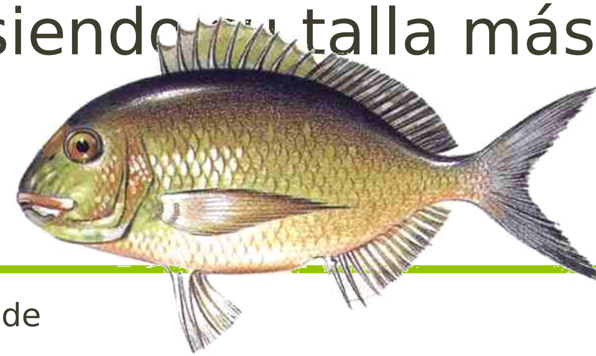
$$F(a) = P(X \le a) = \int_{-\infty}^{a} f(x) \cdot dx$$

$$f(x) \ge 0$$

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

Ejemplo: la dorada

- Tiene un cuerpo alto y compacto.
- Se suele conocer por presentar una banda amarilla en la parte frontal de la cabeza y entre los ojos.
- Su dorso es de color gris plateado, los flancos amarillo grisáceo con algún reflejo dorado, también presenta a la altura de la abertura branquial una mancha oscura.
- Puede alcanzar unos 70 cms, siendo su talla más común de 20 a 50 cms.



Distribución de doradas (Variable discreta)

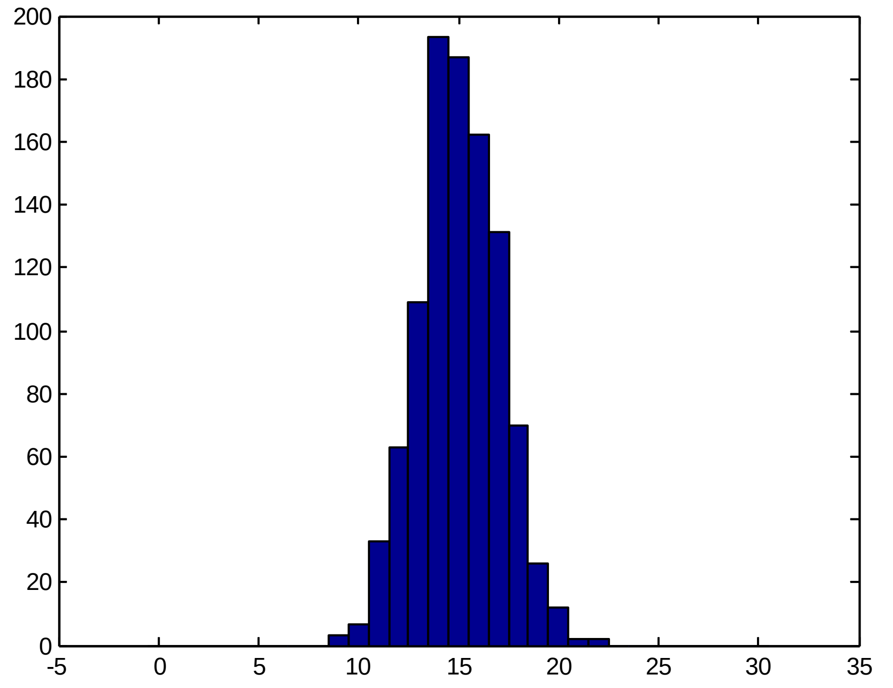
```
x=floor(15+2*randn(1,1000));

x= [ 11   16   13   15   11   18   18   18
     15   14   17   14   19   15   17   12
     16   11   15   13   15   13   15   14
     12   18   16   15   14 ...
```

Concepto de histograma (Variable discreta)

hist(x,0:30)

s=hist(x,0:30)

bin = 12; length(find(x==bin))



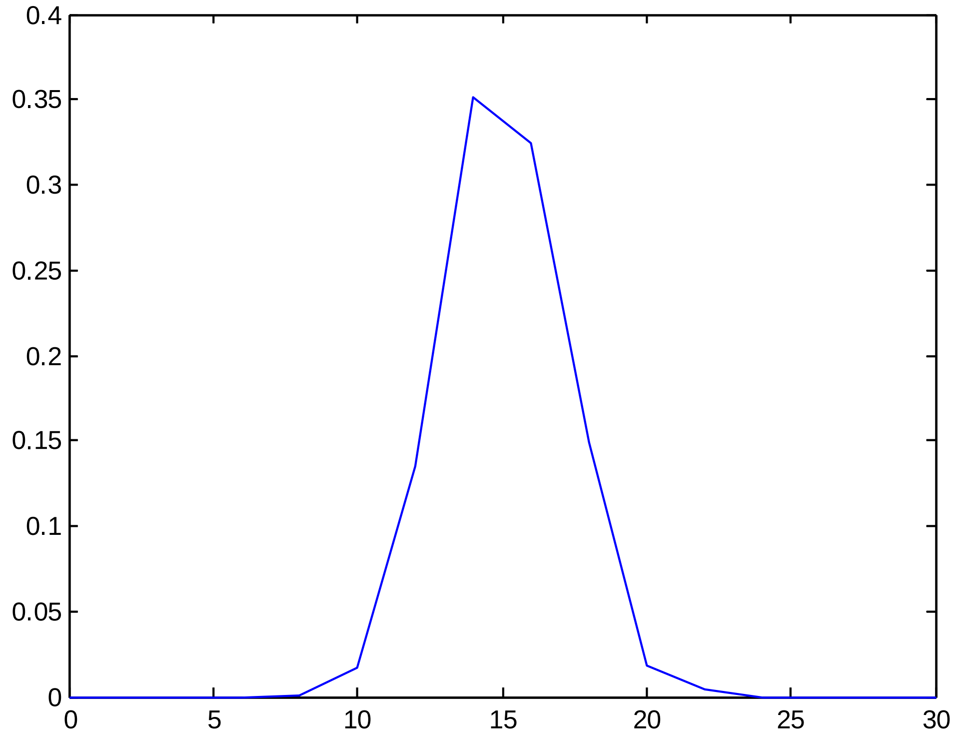
pregunta

- ¿Cómo podemos convertir el histograma en una función de probabilidad?

Concepto de probabilidad (Variable discreta)

```
s = hist(x,0:2:30)
f = s / length(x)
plot(0:2:30,f)
```

f(X) ≈ P(x==X)




```
F = cumsum(f)
plot(0:2:30,F)
```

F(X) ≈ P(x<=X)

Distribución de doradas (Variable continua)

x=~~floor~~(15+2*randn(1,1000));

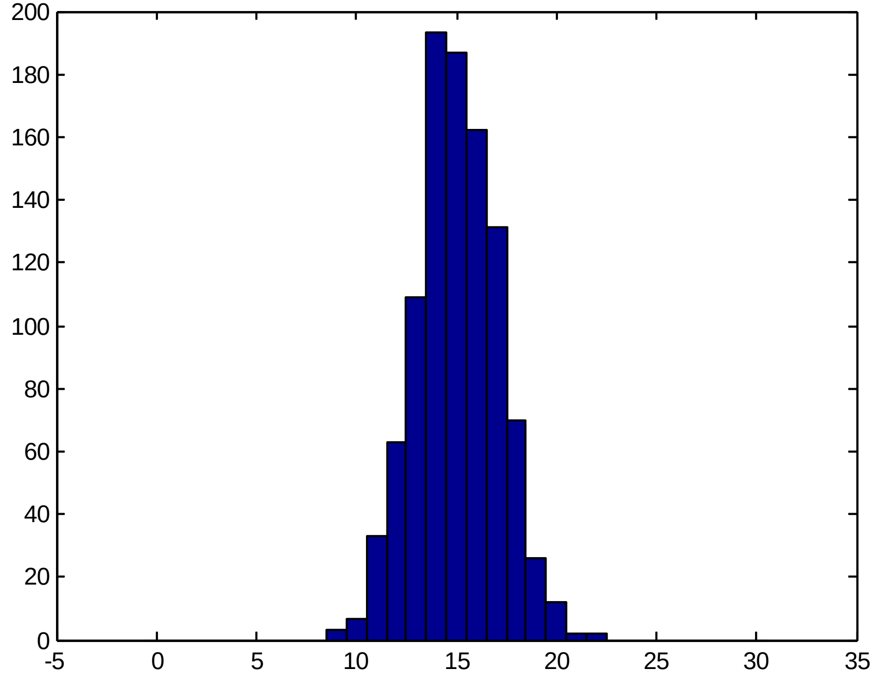
x= [11.21 16.14 13.18 15.16 11.22
16.19 11.21 15.14 13.07 15.09
12.14 18.92 16.81 15.27 ...

Concepto de histograma (Variable continua)

hist(x,0:30)

s=hist(x,0:30)

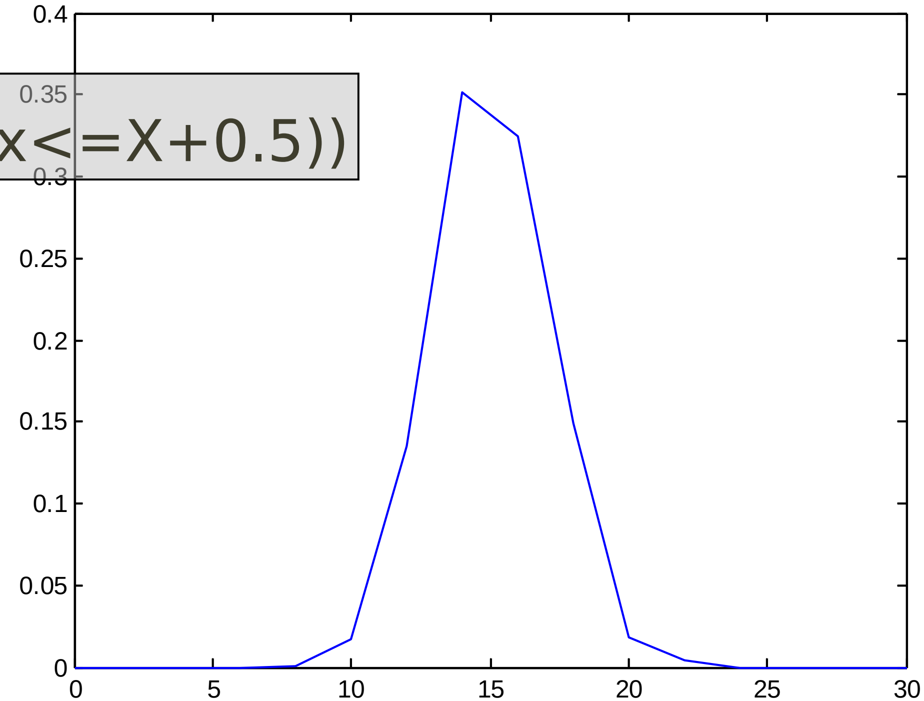
bin = 10; length(find((x>bin-0.5) & (x<bin+0.5)))



Concepto de probabilidad (Variable continua)

```
s = hist(x,0:2:30)
f = s / length(x)
plot(0:2:30,f)
```

f(X) ≈ P((x>=X-0.5) & (x<=X+0.5))

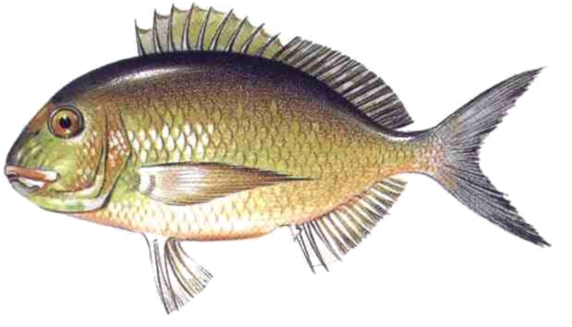



```
F = cumsum(f)
plot(0:2:30,F)
```

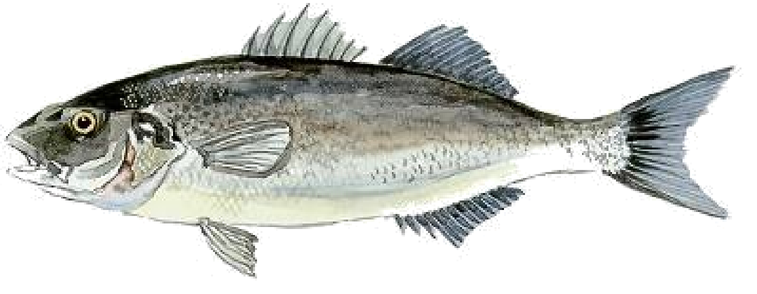
F(X) ≈ P(x<=X)

Problema : clasificar peces por su longitud

```
x=15+2*randn(1,1000);
plot(x,0,'.b','MarkerSize',5)
hold on
```

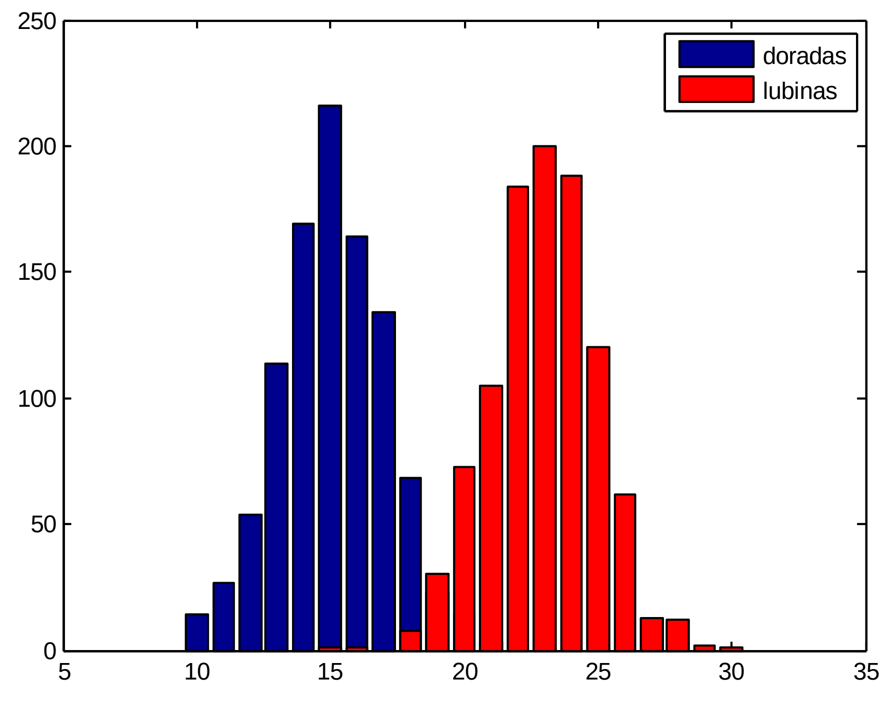


```
y=23+2*randn(1,1000);
plot(y,1,'.r','MarkerSize',5)
axis([0 40 -10 10])
```



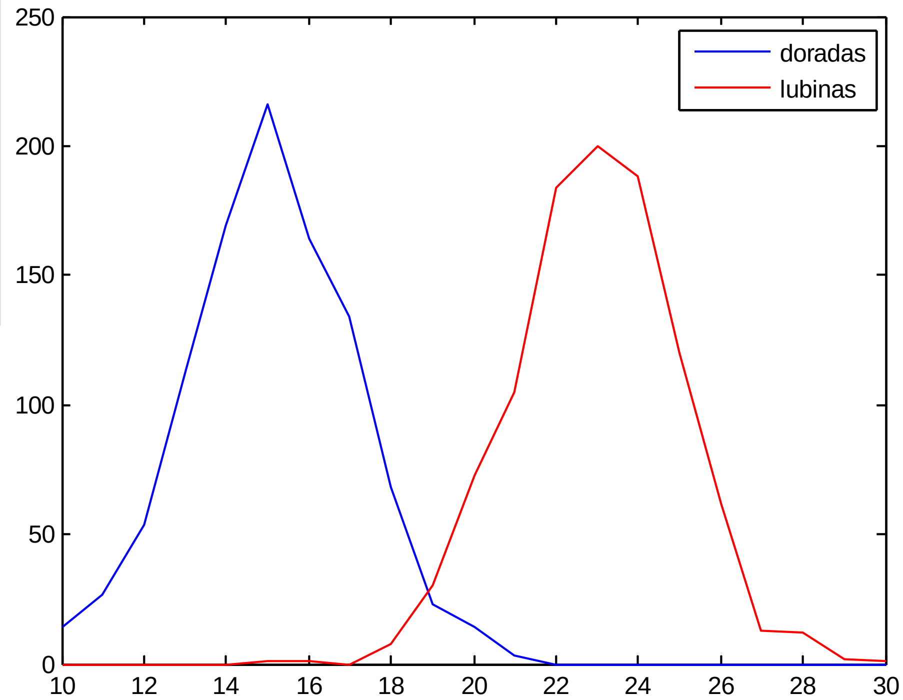
Histogramas de ambas especies

- interv=10:30;
- vx=hist(x,interv);
- vy=hist(y,interv);
- bar(interv,vx);hold on;
- bar(interv,vy,'r');hold off
- legend('doradas','lubinas')



Histogramas de ambas especies

- interv=10:30;
- vx=hist(x,interv);
- vy=hist(y,interv);
- plot(interv,vx);hold on;
- plot(interv,vy,'r');hold off
- legend('doradas','lubinas')

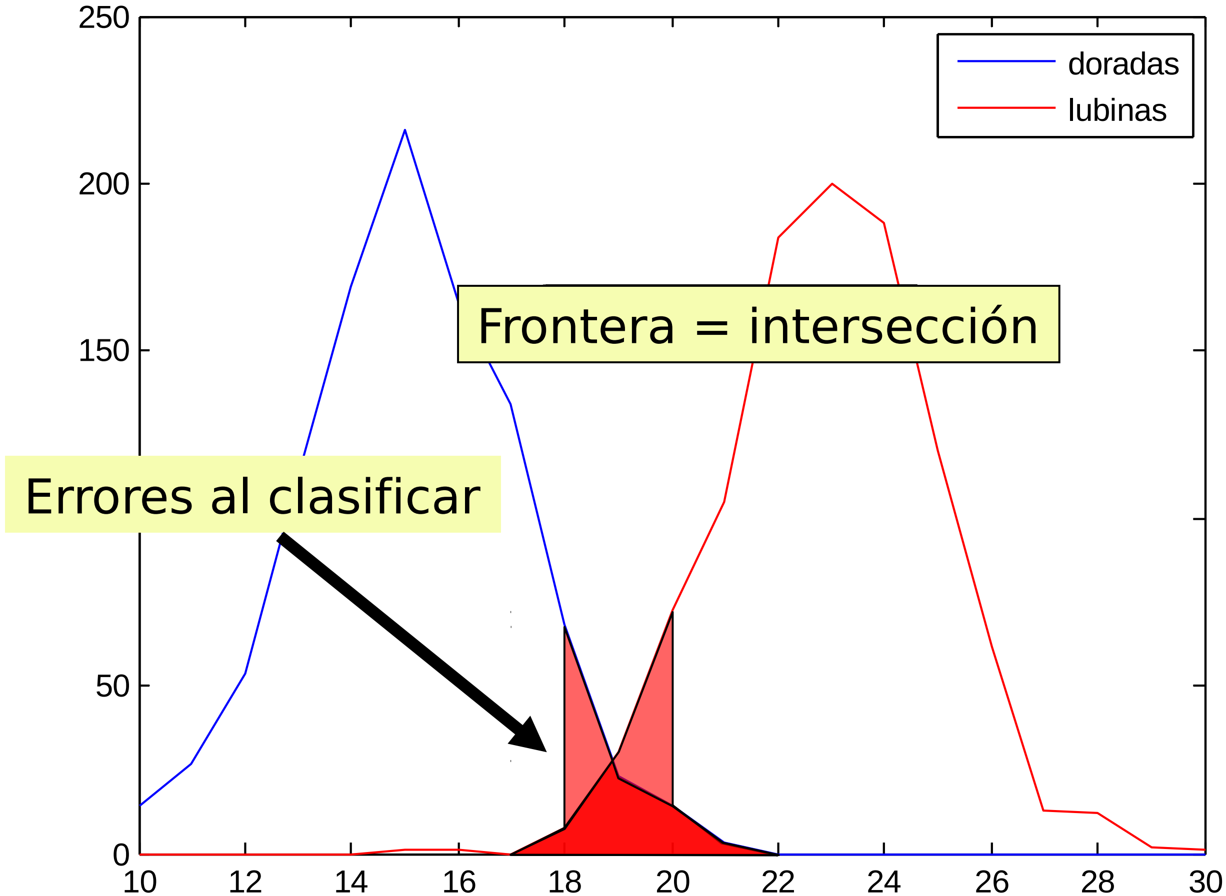


PREGUNTA

- Si elegimos un pez al azar y lo medimos, y su longitud es 15.23 cm. ¿sabemos de qué especie se trata?
- ¿Y si mide 19.54 cm. ?
- ¿A partir de qué valor podemos decir que es más probable que sea una lubina que una dorada?

!!ESTO YA ES UN SISTEMA DE CLASIFICACIÓN!!

Valor óptimo = Mínimo Error



Resultados numéricos

```
clc
x=15+2*randn(1,1000);
y=23+2*randn(1,1000);
for Frontera = 17:21
    Err1 = length(find(x>Frontera));
    Err2 = length(find(y<Frontera));
    ErrTotal = Err1 + Err2;
    disp([' Frontera = ' num2str(Frontera)])
    disp(['   Err1 Err2 ErrTot'])
    disp([Err1 Err2 ErrTotal])
end
```

VALOR EXACTO

- En este caso, debido a que las distribuciones son iguales, simétricas, y el número de datos de cada especie es igual, se puede demostrar que el valor exacto es:

Frontera = 19 cm.

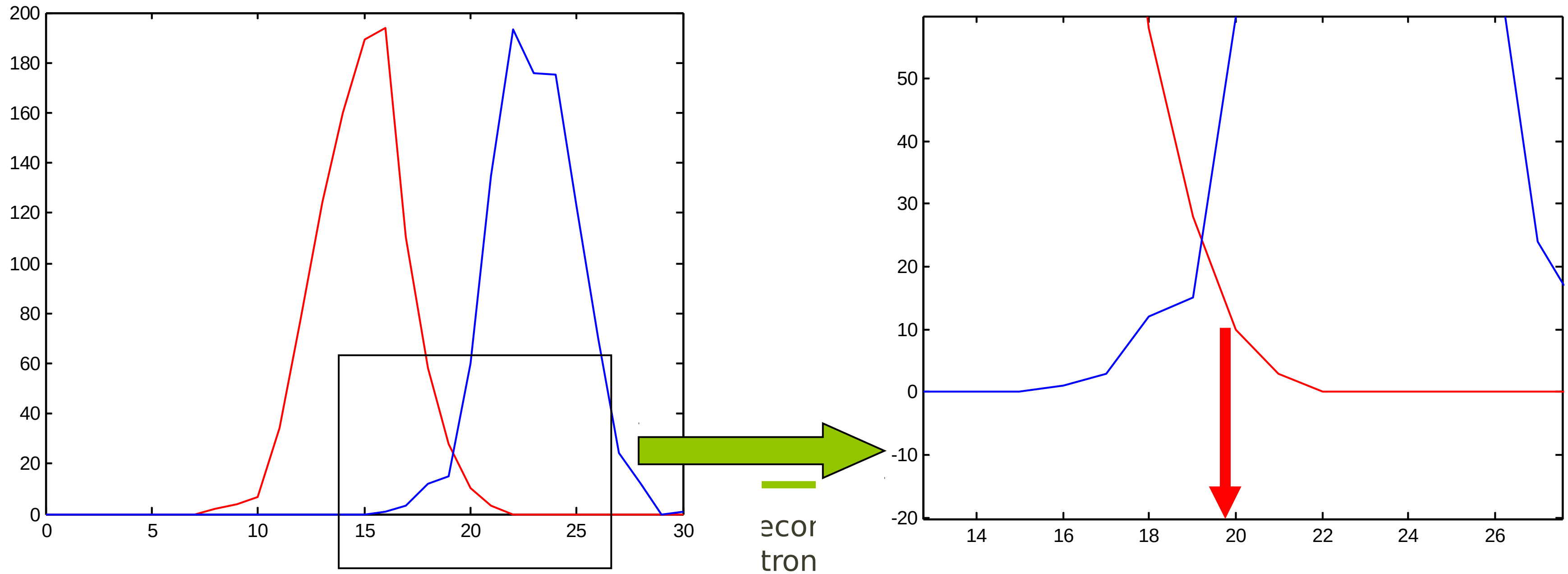
- Normalmente no es éste el caso

- ¿Qué hacer si no conozco la solución exacta?

SOLUCIÓN 1

- Dibujar los histogramas de ambas especies, y buscar el punto de corte

```
sx=hist(x,0:30);plot(0:30,sx,'r');hold on;
sy=hist(y,0:30);plot(0:30,sy, 'b');hold off;
```



Problemas de la solución 1

- Inexactitud debido a la discretización realizada
  - Cuanto más estrecha es cada barra del histograma, más precisión
  - Pero tengo menos valores por cada barra
  - COMPROMISO

- **!! Necesito sacar todos los peces del mar para conocer el valor exacto de la frontera óptima !!**

SOLUCIÓN 2

- Suponemos que las distribuciones son simétricas e iguales en forma
  - mx = mean(x)
  - my = mean(y)
  - Frontera = (mx+my) /2

- PREGUNTAS
  - ¿Qué método da mejores resultados?
  - ¿Cuál es más sencillo?
  - ¿Cuál es más rápido de calcular?
  - ¿Qué método es más riguroso?

Problemas de la solución 2

- ¿Qué ocurre si las distribuciones no son simétricas?

```
x=15+sum(randn(5,1000).^2);
y=23+sum(randn(5,1000).^2);
sx=hist(x,0:40);plot(0:40,sx,'r');hold on;
sy=hist(y,0:40);plot(0:40,sy, 'b');hold off;
title(num2str(0.5*(mean(x)+mean(y))))
```

Problemas de la solución 2

- ¿Qué ocurre si las distribuciones no son iguales en forma?

```
x=15+2*randn(1,1000);
y=23+4*randn(1,1000);
sx=hist(x,0:40);plot(0:40,sx,'r');hold on;
sy=hist(y,0:40);plot(0:40,sy, 'b');hold off;
title(num2str(0.5*(mean(x)+mean(y))))
```

Problemas de la solución 2

- ¿Qué ocurre si hay un nº de individuos diferente para cada especie?

```
x=15+2*randn(1,3000);
y=23+2*randn(1,97000);
sx=hist(x,0:40);plot(0:40,sx,'r');hold on;
sy=hist(y,0:40);plot(0:40,sy, 'b');hold off;
title(num2str(0.5*(mean(x)+mean(y))))
```

PROBLEMA GENERAL

- Para que todo funcione lo mejor posible, tendríamos que sacar y medir **TODAS** las doradas y lubinas del mar
- Por ser imposible, trabajaremos con una **MUESTRA** representativa de la población
- Cuanto más elementos haya en mi muestra, y más representativa sea, mejores serán las estimaciones

SOLUCIÓN 3

- Sacar una muestra representativa de cada tipo de peces
- Estimar lo mejor posible la f. de densidad de probabilidad (forma y parámetros) para cada tipo de pez
- Ajustar las f. de densidad para tener en cuenta el nº de individuos de cada clase
- Encontrar la intersección entre ambas curvas de **forma analítica** y dar ese valor como la frontera óptima

Cuidado con los casos especiales

- ¿Qué ocurre si las distribuciones están centradas en el mismo punto, pero tienen diferente anchura?

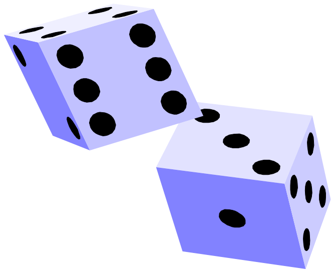
```
x=15+4*randn(1,10000);
y=15+2*randn(1,10000);
sx=hist(x,0:40);plot(0:40,sx,'r');hold on;
sy=hist(y,0:40);plot(0:40,sy, 'b');hold off;
title(num2str(0.5*(mean(x)+mean(y))))
legend('doradas', ‘herreras')
```

Problemas de la solución 3

- Estimar lo mejor posible la f. de densidad de probabilidad de una población a partir de una muestra

- Estimar lo mejor posible el nº de individuos en la población de cada clase a partir de una muestra

!!! MÁS ESTADÍSTICA !!!



Distribuciones más frecuentes (discretas)

Distribución uniforme

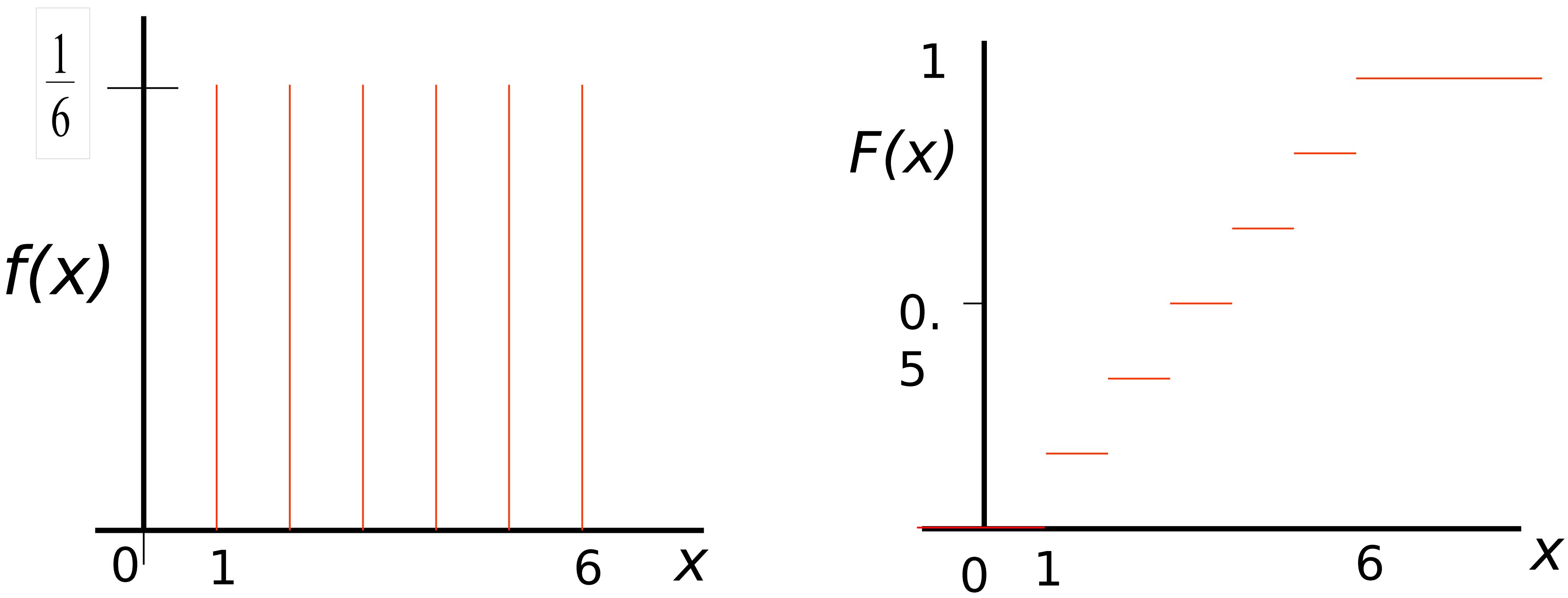
- Si todos los sucesos elementales tienen la misma probabilidad decimos que X **se distribuye uniformemente**.
- Si el espacio muestral consta de n sucesos simples ($0 < n < \infty$), entonces la función de probabilidad discreta se define como

$$p(x) = 1/n$$

  - Ejemplos:
    - Lanzamiento de un dado (n=6)
    - Lanzamiento de una moneda (n=2)

Distribución uniforme
Ejemplo: lanzamiento de un dado

- *X* tiene como posibles valores *x* = 1, 2, 3, 4, 5, 6 cada uno con probabilidad 1/6



Función de probabilidad f(x)

Función de distribución F(x)

Distribución de Bernouilli

- Si el resultado de un experimento solo admite dos resultados: éxito(1) o fracaso(0)
- Un típico experimento de Bernoulli es el lanzamiento de una moneda con probabilidad p para cara y (1-p) para cruz.
- Si la probabilidad de éxito es p:

$$f(x) = p^x (1-p)^{1-x} \qquad x = 0,1$$

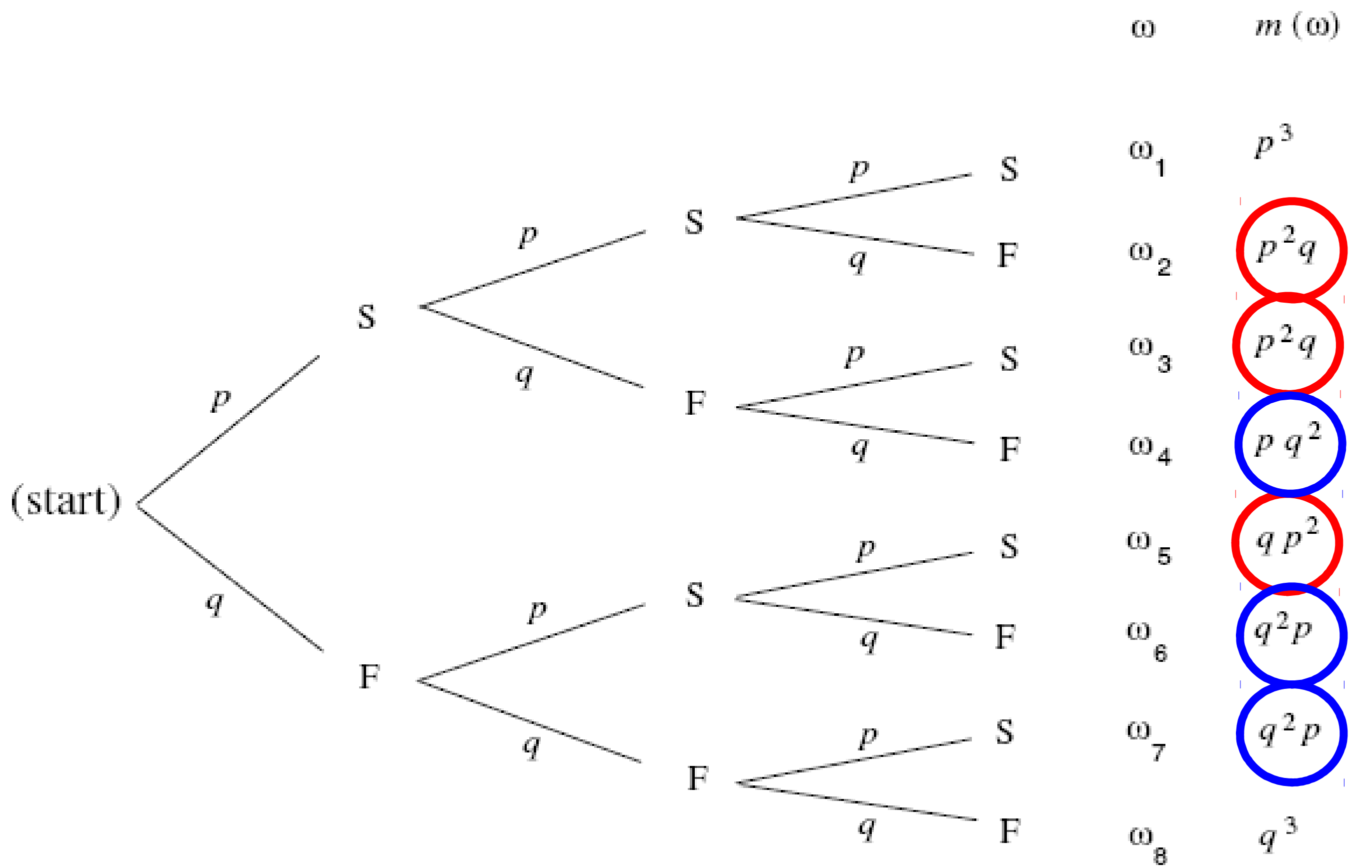
$$F(x) = \begin{cases} 1-p, & \text{para } x = 0 \\ 1, & \text{para } x = 1 \end{cases}$$

Distribución binomial

- La distribución binomial aparece cuando estamos interesados en el **número de veces que un suceso *A* ocurre (éxitos) en *n* intentos independientes de un experimento de Bernouilli.**

- *P. ej.: nº de caras en 3 lanzamientos de una moneda.*

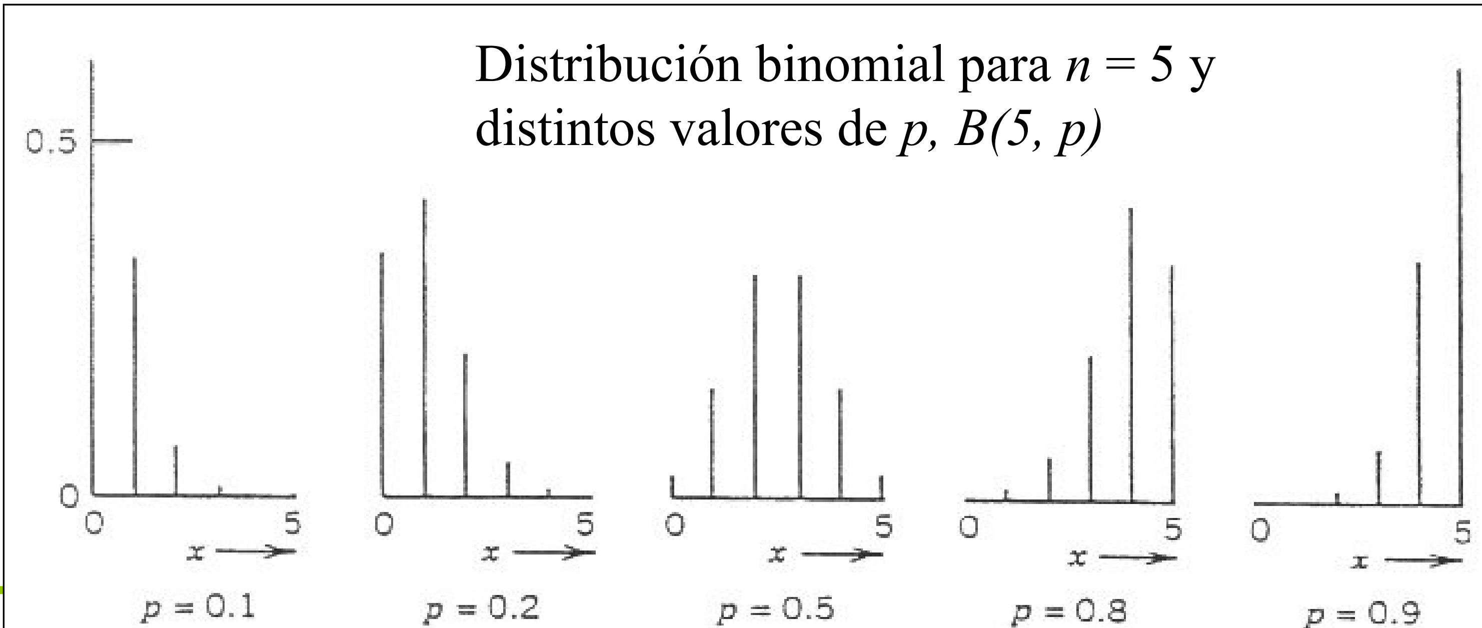
Distribución binomial
Nº de caras en 3 lanzamientos



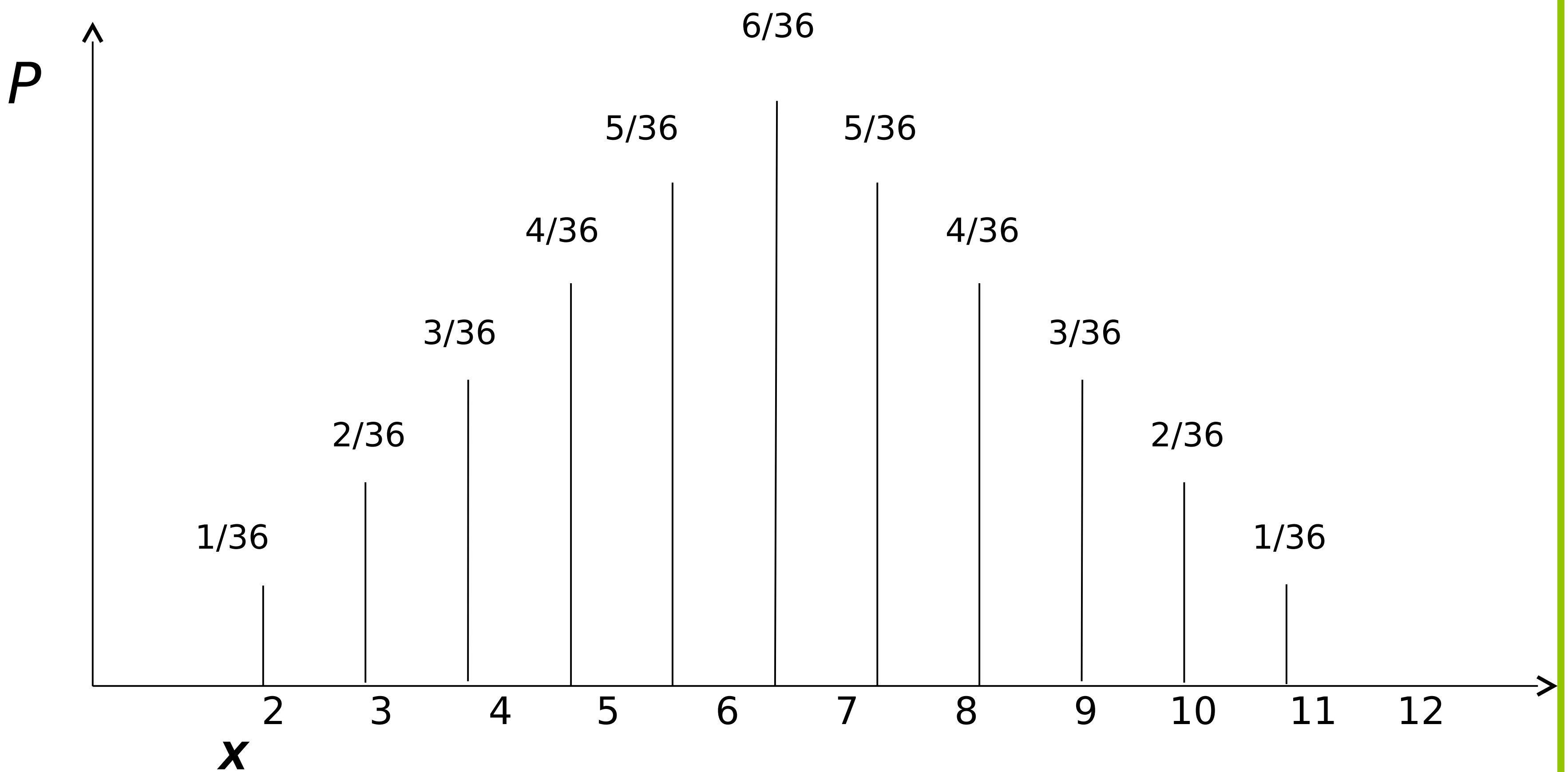
Distribución binomial

- En este caso, la función de probabilidad es:

$$B(n,p) = p(x) = \binom{n}{x} p^x (1-p)^{n-x} = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$



Distribución binomial
Ej: lanzamiento de dos dados



Problemas propuestos

1. El 30% de los estudiantes de la UCA son miopes. Si se coge a 20 estudiantes aleatoriamente, ¿cuál es la prob. de que como máximo haya dos miopes?
2. Una moneda se tira 10 veces. ¿Cuál es la prob. de que salgan menos de 3 caras?
3. Un granjero planta 12 tomateras. En promedio, el 15% mueren el primer invierno. Calcula la prob. de que muera este invierno más de una
4. Unas bombillas se empaquetan en cajas de 20. Una bombilla de cada 10 es defectuosa de media. ¿Cuál es la prob. de que una caja tenga dos bombillas defectuosas?

Soluciones : (1) 0.0355 (2) 0.0547 (3) 0.5565 (4) 0.285

Distribución de Poisson

- Obtiene la probabilidad de que un suceso ocurra x veces en un cierto periodo de tiempo, sabiendo que el nº medio de que ocurra esto es λ
- Dado que x es el nº de ocurrencias, x = 0,1,2,3, …
- *La función de probabilidad es:*

$$p(x) = \frac{e^{-\lambda}\lambda^{x}}{x!} \qquad x = 0,1,2,\ldots$$

- *Ejemplos: Llegada de fotones a un detector.*

Poisson = límite de la Binomial

- Cuando en una distribución binomial el número de intentos (*n*) es grande y la probabilidad de éxito es pequeña, la distribución binomial converge a la distribución de Poisson (λ = n·p)

$$p(x) = \frac{n!}{x!\cdot(n-x)!}\cdot p^x \cdot (1-p)^{n-x} \qquad \text{tomando } \lambda = n \cdot p$$

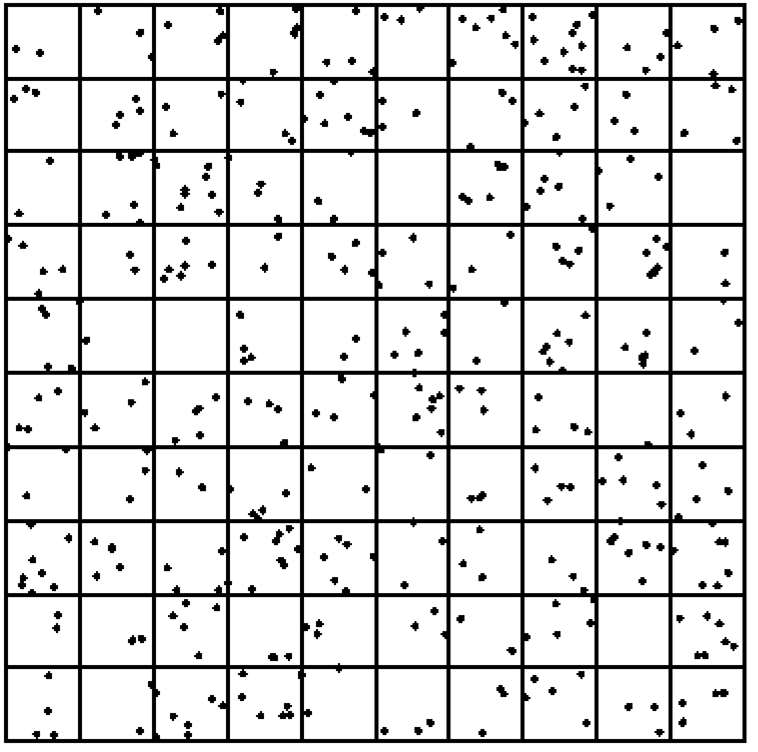
$$\lim_{n\to\infty} p(x) = \lim_{n\to\infty} \frac{n(n-1)(n-x+1)}{x!}\cdot\frac{\lambda^x}{n^x}\left(1-\frac{\lambda}{n}\right)^n\cdot\left(1-\frac{\lambda}{n}\right)^{-x}$$

$$\lim_{n\to\infty} p(x) = \lim_{n\to\infty} \frac{n(n-1)(n-x+1)}{n^x}\cdot\frac{\lambda^x}{x!}\left(1-\frac{\lambda}{n}\right)^n\cdot\left(1-\frac{\lambda}{n}\right)^{-x}$$

$$\lim_{n\to\infty} p(x) = 1\cdot\frac{\lambda^x}{x!}e^{-\lambda}\cdot 1 = \frac{\lambda^x}{x!}e^{-\lambda}$$

Ejemplo: **Bombas sobre Londres en la II Guerra Mundial (Feller)**

400 bombas



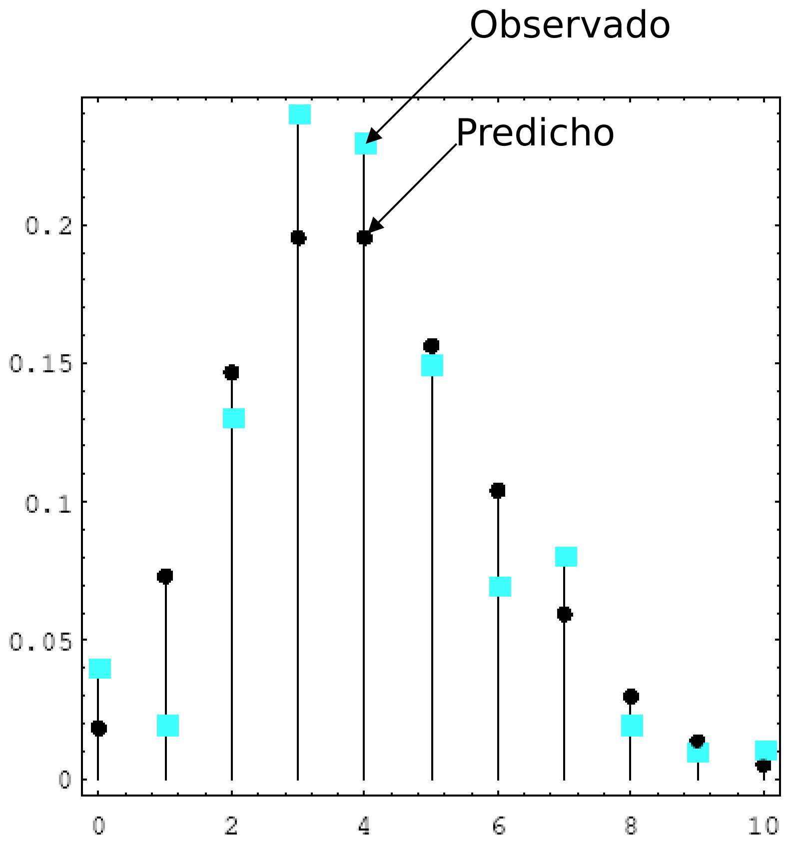
10 x 10

Ejemplo: **Bombas sobre Londres en la II Guerra Mundial (Feller)**

- Supón que vivías en uno de los 100 bloques que aparecen en la gráfica inferior.
- La probabilidad de que una bomba cayera en tu bloque era p=1/100.
- Como cayeron 400 bombas, podemos entender el número de impactos en tu bloque como el número de éxitos en un experimento de Bernoulli con n = 400 y p = 1/100.
- Podemos usar una Poisson con $\lambda=n*p = 400 * 1/100 = 4$.

Ejemplo: **Bombas sobre Londres en la II Guerra Mundial (Feller)**

$$p(x) = \frac{e^{-4}4^x}{x!}$$



Problemas propuestos

1. Una secretaria comete una media de dos errores por página. ¿Qué probabilidad hay de que escriba una página sin ningún error?
2. Un ordenador tiene de media una "caída" cada dos días. ¿Cuál es la probabilidad de que haya 2 caídas en una semana?
3. Están vendiendo juguetes con un nº medio de fallos de 8. ¿Cuál es la prob. de comprar un juguete con un solo fallo?

Soluciones : (1) 0.135 (2)0.185 (3) 0.0027

Otras distribuciones:multinomial

- Cuando hay más de dos acontecimientos posibles (A1, A2, A3 ...) con probabilidades **p1** , **p2** , **p3 ...** constantes y tales que:

$$\sum_i p_i = 1$$

$$p(x_1, x_2, x_3 \ldots) = \frac{n!}{x_1!x_2!x_3!\ldots} p_1^{x_1} . p_2^{x_2} . p_3^{x_3} \ldots.$$

Otras distribuciones: geométrica

- Consiste en repetir un experimento de Bernouilli hasta conseguir el primer éxito.
- Definimos la variable aleatoria X, como el **número de fracasos hasta que se obtiene el primer éxito**.

$$f(x) = (1-p)^x p \qquad x = 0, 1, 2, \ldots$$

$$F(n) = \sum_{x=0}^{n} (1-p)^x p = 1 - (1-p)^{n+1}$$

- Ejemplo : sea cruz = éxito, cuántas veces tengo que tirar una moneda hasta sacar una cruz

Otras distribuciones: binomial negativa

- Repetimos un experimento de Bernouilli **hasta conseguir el r-ésimo éxito**, el número de fracasos hasta que se obtiene el r-ésimo éxito sigue la distr. binomial negativa.

$$BN(r,p) = P(X = x) = \binom{x+r-1}{x} p^r (1-p)^x,$$
$$x = 0, 1, 2, \ldots$$

- Ejemplo: lanzar una moneda hasta sacar 3 caras

Otras distribuciones: binomial negativa

- La distribución binomial negativa también se puede definir como el número de pruebas *x* hasta la aparición de *r* éxitos.

- Como el número de pruebas *x*, en este caso, contabiliza tanto los éxitos como los fracasos se tendría según ésta definición que:

$$BN(r,p) = P(X = x) = \binom{x-1}{r-1} p^r (1-p)^{x-r},$$

$$x = r,\ r+1,\ r+2,\ \ldots$$

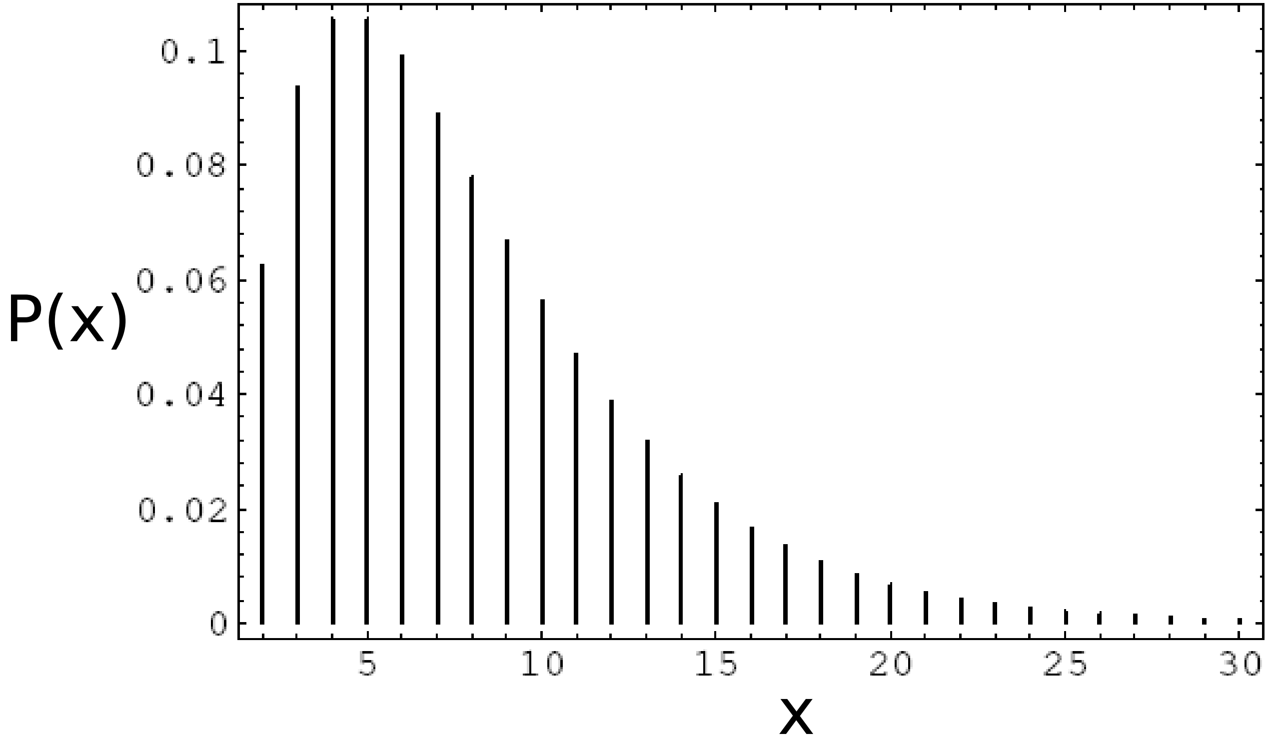
Ejemplo

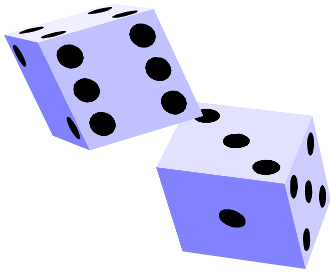
- Disponemos de una moneda trucada con probabilidad de cara igual a p=0.25. La lanzamos hasta que obtenemos 2 caras. La distribución del número de lanzamientos *x* será:

$$BN(r=2, p=0.25) = P(X=x) = \binom{x-1}{2-1} 0.25^2 (1-0.25)^{x-2},$$

$$x = 2,3,4,\ldots$$

ejemplo





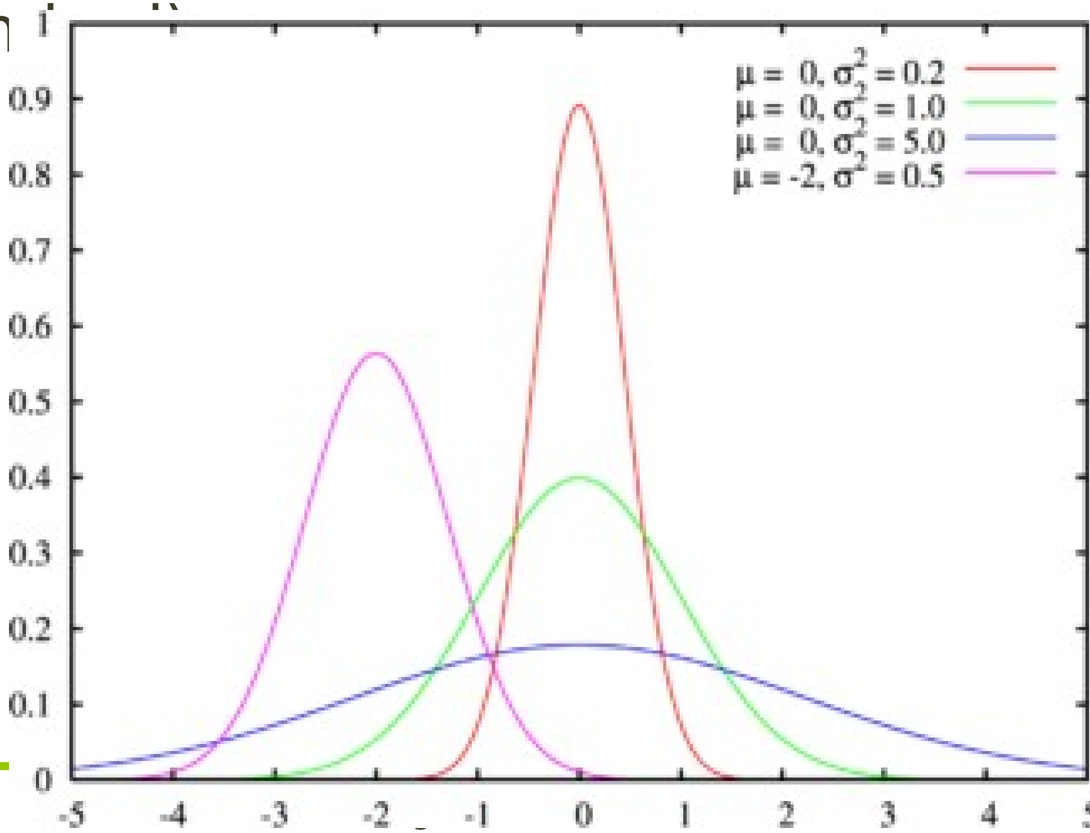
Distribuciones más frecuentes (continuas)

Distribución normal

- Normal, también llamada de Gauss o gaussiana
- Es la distribución de probabilidad que con más frecuencia aparece debido a fundamentalmente a que hay muchas variables asociadas a fenómenos naturales que siguen el modelo de la normal
  - Caracteres morfológicos de individuos
  - Caracteres fisiológicos como el efecto de un fármaco
  - Caracteres sociológicos como el consumo de cierto producto por un mismo grupo de individuos
  - Caracteres psicológicos como el cociente intelectual
  - Nivel de ruido en Telecomunicaciones
  - Errores cometidos al medir magnitudes en experimentos
  - Valores estadísticos muestrales como la media
- Es, además, límite de otras distribuciones

Distr. normal : Forma de campana

- Dos factores, el punto donde está centrada (media) y la anchura de la campana (desv. están



Función de densidad de prob. normal

- La función de densidad de probabilidad se define matemáticamente como:

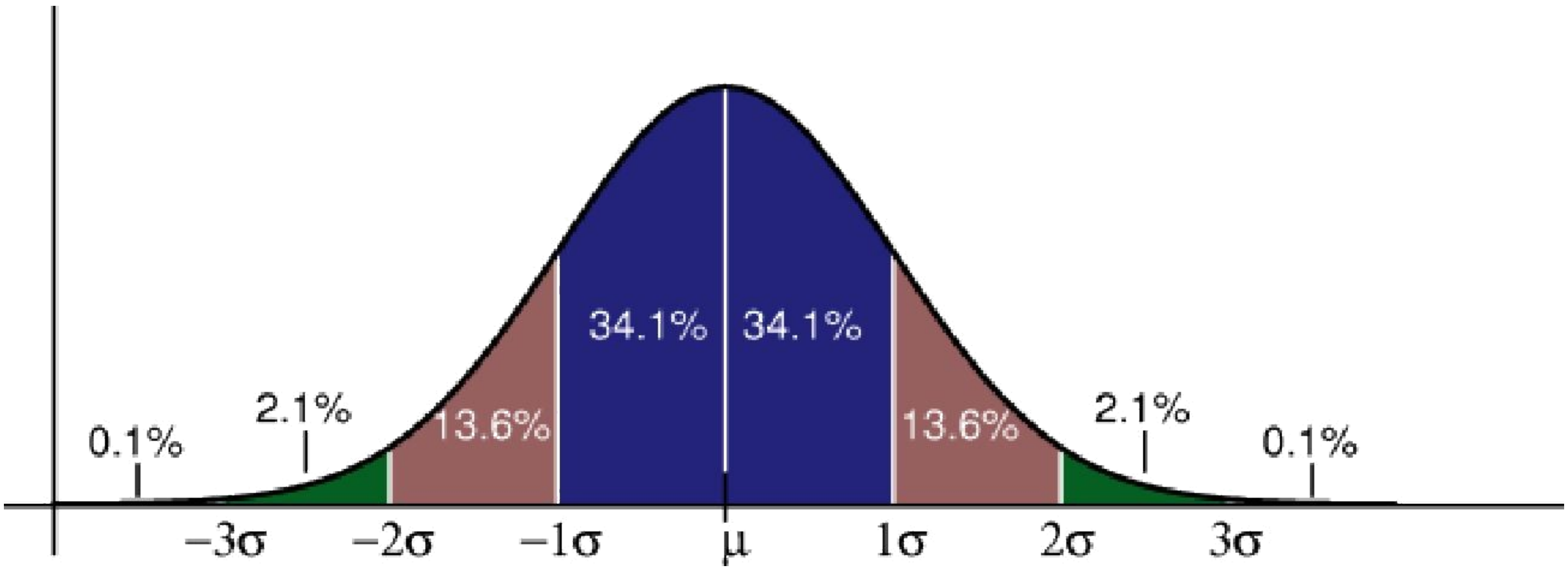
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

donde μ (mu) es la media y σ (sigma) es la desviación estándar

  - Al cuadrado de sigma (σ2) se le llama varianza.

Regla del 68–95–99.7

- Casi todos los datos se encuentran a menos de 3 desviaciones estándar (3σ) de la media(μ)

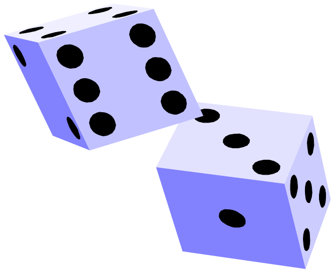


Tipificación de distr. normales

- Si μ=0 y σ=1, distribución normal estándar.
- Dada una variable aleatoria normal X, con media μ y desviación típica σ, si definimos otra variable aleatoria

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

entonces la variable aleatoria Z tendrá una distribución normal estándar



Dibujo de diversas distribuciones estadísticas

Dibujo de una distribución normal

```
close all
N = 500;         % Numero de puntos generados
m = 3;           % Media de la distribucion
s = 2;           % Desv. tipica de la distribucion

x=randn(1,N);    % Datos
x=s*x+m;

I = m-3*s:s/5:m+3*s;      % Curva pdf gaussiana
h=normpdf(I,m,s);

p=hist(x,I);              % Histograma obtenido a partir de x

p=p/N*sum(h);             % Cambio de escala

bar(I,p);hold on;
plot(I,h,'r','LineWidth',3);hold off;
```

Dibujo de la distribución normal N(0,1)

```
close all
suma = randn(1,10000);

intervalo = -5:0.1:5;
a1=hist(suma,intervalo); a1=a1/sum(a1);
a2=normpdf(intervalo,0,1); a2=a2/sum(a2);

plot(intervalo,a1);hold on;
plot(intervalo,a2,'r');hold on;
```

Dibujo de la distribución normal N(μ,σ)

```
close all
m = 3;
s = 2;
suma = m + s*randn(1,10000);

intervalo = -5*s:s/10:5*s;
a1=hist(suma,intervalo); a1=a1/sum(a1);
a2=normpdf(intervalo,m,s); a2=a2/sum(a2);

plot(intervalo,a1);hold on;
plot(intervalo,a2,'r');hold on;
```

Dibujo de una distribución $\chi2(N)$ a partir de N distribuciones N(0,1)

```
close all
% La suma de N normales N(0,1) al cuadrado
% es una chi2 con N grados de libertad
% ================================================
N = 10;
suma=sum(randn(N,10000).^2);

intervalo = 0:0.1:30;
a1=hist(suma,intervalo); a1=a1/sum(a1);
a2=chi2pdf(intervalo,N); a2=a2/sum(a2);

plot(intervalo,a1);hold on;
plot(intervalo,a2,'r');hold on;
```

Dibujo de una $\chi2$(N) a partir de N distribuciones N(0,σ)

```
% La suma de N normales N(0,v) es una chi2(N)
%            Y = (v*X1)^2 + (v*X2)^2 + (v*X3)^2
%            => Y/(v*v) es una chi2(3)
% ======================================================
close all
N = 10;
v = 2;
suma=sum ((v*randn(N,10000)).^2);
suma = suma/(v*v);

intervalo = 0:0.1:30;
a1=hist(suma,intervalo);
a1=a1/sum(a1);
a2=chi2pdf(intervalo,N);
a2=a2/sum(a2);

plot(intervalo,a1);hold on;
plot(intervalo,a2,'r');hold on;
```

Dibujo de una χ2(N) a partir de N distribuciones N(μ,σ)

```
% La suma de N normales N(mu,v) es una chi2(N) no centrada con
%         parametro delta = N*mu*mu/(v*v)
%             Y = (v*X1+mu)^2 + (v*X2+mu)^2 + (v*X3+mu)^2 =>
%             Y/(v*v) es una ncx2 con 3 grados de libertad y
%                                   con delta=3*mu*mu/v/v
% ===========================================================
N = 10;
mu = 4;
v  = 2;
suma = sum((mu+v*randn(N,10000)).^2);
suma=suma/(v*v);

intervalo = 0:400;
a1=hist(suma,intervalo);
a1=a1/sum(a1);
a2=ncx2pdf(intervalo,N,N*mu^2/(v*v));
a2=a2/sum(a2);

plot(intervalo,a1);hold on;
plot(intervalo,a2,'r');hold on;
```

Dibujo de una distribución F(m,n)

```
% La suma de :
%          (N normales al cuadrado dividido por N) /
%          (M normales al cuadrado dividido por M)
%          es una F con M y N grados de libertad
%
%              (X1^2 + X2^2 + X3^2)/3
%           Y= ----------------------  => Y es una F(3,2)
%                   (X4^2 + X5^2)/2
N = 50;
M = 5;
suma1=sum(randn(N,10000).^2);
suma2=sum(randn(M,10000).^2);
suma=(suma1/N)./(suma2./M);

intervalo = 0:0.1:10;
a1=hist(suma,intervalo);   a1=a1/sum(a1);
a2=fpdf(intervalo,N,M);   a2=a2/sum(a2);

plot(intervalo,a1);hold on;
plot(intervalo,a2,'r');hold on;
```

Dibujo de una t-Student

```
% Si x y s son la media y desv. standard de una muestra de tamaño n
% obtenida de una N(mu,sigmacuadrado=n), entonces (x-mu)/s tiene una
% distribucion t de Student con n-1 grados de libertad

close all
N = 20000;      % Numero de puntos generados
n = 5;       % Tamaño de la muestra

datos = sqrt(n) * randn(n,N);
x = mean(datos);
s = std(datos);

I = -5:0.1:5;            % Curva pdf

h1=tpdf(I,n-1);

h2=normpdf(I,0,1);

p=hist(x,I);             % Histograma t-Student obtenido a partir de x
p=p/N*sum(h);            % Cambio de escala

bar(I,p);hold on;
plot(I,h1,'r','LineWidth',3);
plot(I,h2,'g','LineWidth',3);
hold off;
title('Rojo = t-Student , Verde=Normal')
```

Dibujo de una χ2(N) por paso al límite (M→∞) de la F(N,M)

```
% El limite, cuando M tiende a infinito, de la suma de :
%          (N normales al cuadrado dividido por N) /
%          (M normales al cuadrado dividido por M),
%          todo ello multiplicado por N
%          cuando M tiende a infinito
%          es una chi2 con N grados de libertad
%
%              lim        (X1^2 + X2^2 + X3^2)/3
%           Y=           -------------------------  => Y es una chi2 con 3 grados
%              M->Inf    (X4^2 + X5^2 + ...XM^2)/M

N = 5;
M = 500;  % Aproximacion a Infinito
suma1=sum(randn(N,10000).^2);
suma2=sum(randn(M,10000).^2);
suma = N*(suma1/N)./(suma2./M);

intervalo = 0:0.1:20;
a1=hist(suma,intervalo);
a1=a1/sum(a1);
a2=chi2pdf(intervalo,N);
a2=a2/sum(a2);

plot(intervalo,a1);hold on;
plot(intervalo,a2,'r');hold on;
```

Dibujo de una $\chi 2(N,\delta)$ no centrada

```
% La suma de N normales al cuadrado, de media mu,
%         es una chi2 no central con N grados de libertad, y
%         parametro delta = N * mu * mu
%
%            Y = (X1+mu)^2 + (X2+mu)^2 + (X3+mu)^2 =>
%            => Y es una ncx2 con 3 grados de lib. y delta = 3*mu*mu
%===================================================================
N = 10;
mu = 4;
suma = sum((mu+randn(N,10000)).^2);

intervalo = 0:400;
a1=hist(suma,intervalo);
a1=a1/sum(a1);
a2=ncx2pdf(intervalo,N,N*mu^2);
a2=a2/sum(a2);

plot(intervalo,a1);hold on;
plot(intervalo,a2,'r');hold on;
```

Dibujo de una F(m,n, δ) no centrada

```
% La suma de :
%         (N normales al cuadrado de media mu1 dividido por N) /
%         (M normales al cuadrado dividido por M),
%         es una F no central con M y N grados de libertad, y delta=N*mu1^2
%
%                 (X1^2 + X2^2 + X3^2)/3
%          Y=  -------------------------
%               (Y1^2 + Y2^2 + ...YM^2)/M
%
%           => Y es una ncf con 3 y M grados de libertad, y delta = 3*M*M,
% ============================================================================
N = 3;
M = 10;
mu1=2;
suma1=sum((mu1+randn(N,10000)).^2);
suma2=sum(randn(M,10000).^2);
suma = (suma1/N)./(suma2./M);

intervalo = 0:0.1:25;
a1=hist(suma,intervalo);
a1=a1/sum(a1);
a2=ncfpdf(intervalo,N,M,N*mu1*mu1);
a2=a2/sum(a2);

plot(intervalo,a1);hold on;
plot(intervalo,a2,'r');hold on;
```

Dibujo de una $\chi2(N,\delta)$ no centrada por paso al límite ($M \to \infty$) de la $F(N,M, \delta)$ no centrada

```
% El limite, cuando M tiende a infinito, de la suma de :
%          (N normales al cuadrado de media mu1 y varia v1 dividido por N) /
%          (M normales al cuadrado dividido por M),
%          todo ello multiplicado por N
%          cuando M tiende a infinito
%          es una "ncx2" con M y N grados de libertad, y delta=N*mu1^2
%
%              lim         (X1^2 + X2^2 + X3^2)/3
%           Y=           -------------------------
%              M->Inf     (X4^2 + X5^2 + ...XM^2)/M
%
%            => Y es una ncx2 con 3 grados de libertad, y delta=3*mu1*mu1,
%
% ============================================================================
N = 3;
M = 500;  % Aproximacion a infinito
mu1=2;
suma1=sum((mu1+randn(N,10000)).^2 );
suma2=sum(randn(M,10000).^2);
suma = (suma1/N)./(suma2./M) * N;

intervalo = 0:50;
a1=hist(suma,intervalo);
a1=a1/sum(a1);
a2=ncx2pdf(intervalo,N,N*mu1*mu1);
a2=a2/sum(a2);

plot(intervalo,a1);hold on;
plot(intervalo,a2,'r');hold on;
```

Paso al limite de la ncf =< ncx

```
% demo 9: El limite, cuando M tiende a infinito, de la suma de :
  %          (N normales al cuadrado de media mu1 y varia v1 dividido por N) /
  %          (M normales al cuadrado de media mu2 y varia v2),
  %          todo ello multiplicado por N
  %          cuando M tiende a infinito
  %          es una "ncx2" con M y N grados de libertad, y delta=N*mu1^2
  %
  %              lim         ((mu1+v1*X1)^2 + (mu1+v1*X2)^2 + (mu1+v1*X3)^2)/3
  %           Y=            -----------------------------------------------------
  %              M->Inf     ((mu2+v2*X4)^2 + (mu2+v2*X5)^2 + ...+(mu2+v2*XM)^2)/M
  %
  %            => Y es una ncx2 con 3 grados de libertad, y delta=3*mu1*mu1,
  %
  % ============================================================================

%=======================================================
%    !!!!!! SOLO FUNCIONA SI mu2 ES IGUAL A 0 !!!!!!!!!
% =====================================================
  N = 3;
  M = 200;
  mu1=3;
  mu2=0;
  v1=2;
  v2=5;
  suma1=0;
  for i=1:N,
     suma1=suma1 + (mu1+v1*randn(1,10000)).^2;
  end;

  suma2=0;
  for i=1:M,
     suma2=suma2 + (mu2+v2*randn(1,10000)).^2;
  end;

  if mu2~=0,
     disp('Este caso aun no lo he resuelto');
  end;

  suma = (suma1/N)./(suma2./M) * N /v1/v1*v2*v2;

  intervalo = 0:0.25:25;
  a1=hist(suma,intervalo);
  a1=a1/sum(a1);
  a2=ncx2pdf(intervalo,N,N*mu1*mu1/v1/v1);
  a2=a2/sum(a2);

  plot(intervalo,a1);hold on;
  plot(intervalo,a2,'r');hold on;
```

Paso al limite de la ncf =< ncx

```
% demo 10: El limite, cuando M tiende a infinito, de la suma de :
   %         (N normales al cuadrado de media mu1 y varia v1 dividido por N) /
   %         (M normales al cuadrado de media mu2 y varia v2),
   %         todo ello multiplicado por N
   %         cuando M tiende a infinito
   %         es una "ncx2" con M y N grados de libertad, y delta=N*mu1^2
   %
   %             lim        ((mu1+v1*X1)^2 + (mu1+v1*X2)^2 + (mu1+v1*X3)^2)/3
   %          Y=           -----------------------------------------------------
   %             M->Inf    ((mu2+v2*X4)^2 + (mu2+v2*X5)^2 + ...+(mu2+v2*XM)^2)/M
   %
   %           => Y es una ncx2 con 3 grados de libertad, y delta=3*mu1*mu1,
   %
   % ============================================================================

   N = 3;
   M = 200;
   mu1=3;
   mu2=0.5;
   v1=12;
   v2=1;
   suma1=0;
   for i=1:N,
      suma1=suma1 + (mu1+v1*randn(1,10000)).^2;
   end;

   suma2=0;
   for i=1:M,
      suma2=suma2 + (mu2+v2*randn(1,10000)).^2;
   end;

   if mu2~=0,
      disp('Este caso aun no lo he resuelto');
   end;

   suma = (suma1/N)./(suma2./M) * N /v1/v1*v2*v2;

   intervalo = 0:0.25:25;
   a1=hist(suma,intervalo);
   a1=a1/sum(a1);
   a2=ncx2pdf(intervalo,N,N*((mu1/v1)^2+(v2/mu2)^2));
   a2=a2/sum(a2);

   plot(intervalo,a1);hold on;
   plot(intervalo,a2,'r');hold on;
```