Sociología

Estadística MUY aplicada

Teresa Villagarcía

Probabilidad

- ¿Por qué estudiamos probabilidad?
  - Proporción de piezas defectuosas producidas en un proceso industrial o a favor de una ley. Tomamos una muestra y obtenemos que el 3% es defectuoso. O el 40% está a favor

- ¿Cómo podremos extrapolar nuestras conclusiones al futuro?
  - Evidentemente si otro día se toma otra muestra, es difícil que obtengamos precisamente un 3% de piezas defectuosas. O el 40% a favor.

Probabilidad

- Sabemos que si nuestro servicio funciona bien recibimos en promedio 156 reclamaciones al mes.
  - Este mes hemos recibido 150. ¿Hemos mejorado?
  - El mes pasado recibimos 160. ¿Hemos empeorado?
  - Si recibimos 180 ¿qué debemos hacer?

- Cuando todo funciona bien nuestra máquina produce un 2% de piezas defectuosas.
  - Hoy hemos producido 1000 piezas y ha habido 23 piezas defectuosas. ¿Está mal la máquina?

La probabilidad es el arma que vamos a utilizar para poder generalizar nuestras conclusiones a toda la población de referencia.

Para introducir la Probabilidad vamos previamente a definir una serie de conceptos:

- Experimento aleatorio
- Sucesos
  - Elementales
  - Compuestos
- Espacio muestral
- Suceso seguro
- Suceso imposible

La Probabilidad se define sobre los sucesos y es una medida de la incertidumbre de un suceso

- Si tenemos un suceso, A
- La probabilidad se define sobre A
- Se escribe P(A)
- Y representa la probabilidad de que A ocurra

Propiedades de la probabilidad

- P(A) está comprendida entre 0 y 1

$$0 \leq P(A) \leq 1$$

Todos tenemos una idea intuitiva de lo que es la probabilidad:

Sirve para medir incertidumbres.

¿Es probable que mañana haga 40ºC?

¿Y que yo termine la clase afónica?

Propiedades de la probabilidad

No vamos a trabajar en clase la probabilidad y sus cálculos.

Variables Aleatorias

- Son las variables numéricas que pueden tomar valores al azar:
- Ejemplos:
  - Nº de reclamaciones al mes en un servicio
  - Nº de servicios prestados
  - Nº personas que hay en un hotel.
  - Retrasos promedio en los trenes

Variables Aleatorias

- Pueden ser
  - Discretas:
    - Toman un conjunto de valores:
      - Número de hijos de una familia: 1,2,3, pero no 4.34
  - Continuas
    - Pueden tomar cualquier valor
      - Retrasos del tren: puede ser 5.34 minutos

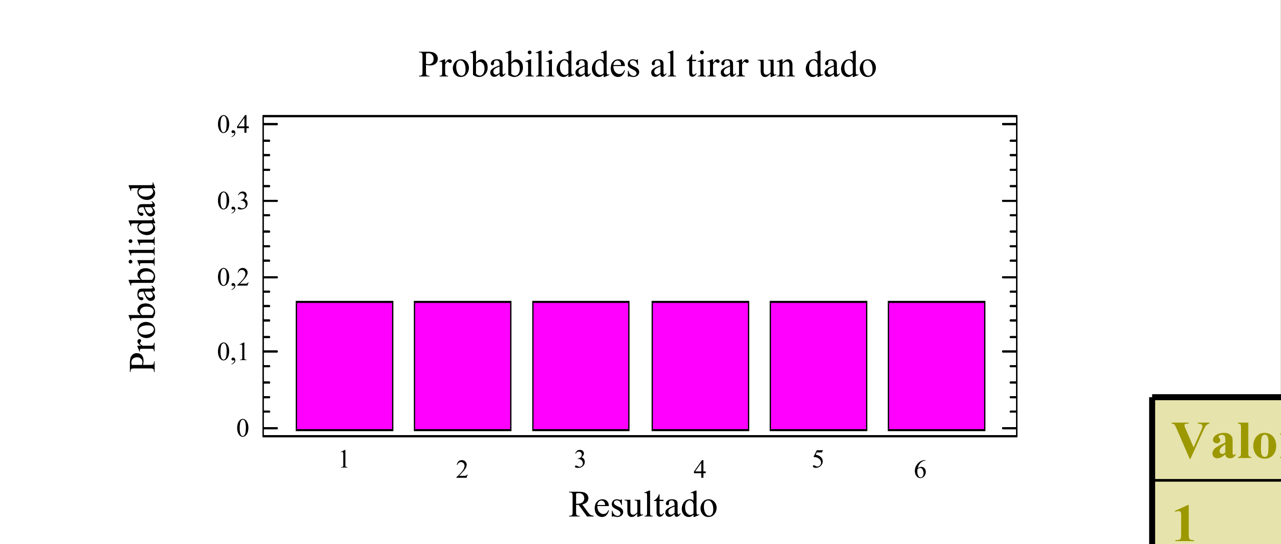
¿Cómo se estudian las Variables Aleatorias?

- Definiendo todos los valores que puede tomar la variable
- Y sus probabilidades.
- Ejemplo: Tirar un dado

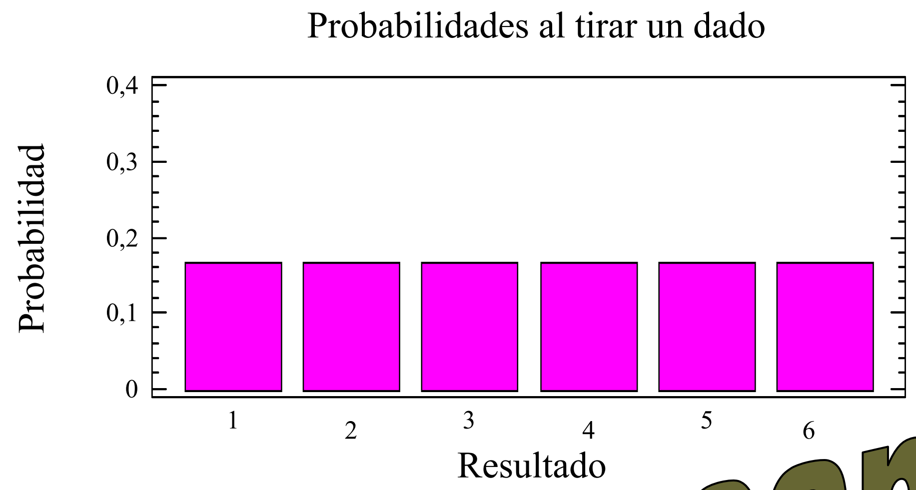
(discreta)

Valor	Probabilidad
1	1/6=0.166
2	1/6=0.166
3	1/6=0.166
4	1/6=0.166
5	1/6=0.166
6	1/6=0.166

¿Cómo se estudian las Variables Aleatorias? DADO (Cont)



Valor	Probabilidad
1	1/6=0.166
2	1/6=0.166
3	1/6=0.166
4	1/6=0.166
5	1/6=0.166
6	1/6=0.166



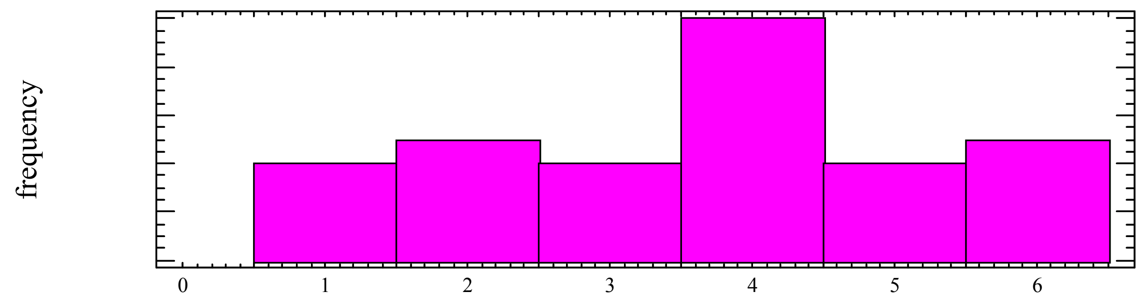
Valor	Probabilidad
1	1/6=0.166
2	1/6=0.166
3	1/6=0.166
4	1/6=0.166
5	1/6=0.166
6	1/6=0.166

¿Qué pensamos?

DADO correcto.

Hacemos un experimento: Tirar el dado

150 veces. Los datos se representan en el histograma:



Esta es la forma de trabajar:

1. Tenemos una idea de cómo será el proceso cuando todo vaya bien (Dado equilibrado)
2. Tomamos datos (Tiramos el dado)
3. Comprobamos si todo va bien

¿Cómo se estudian las Variables Aleatorias?

- Definir en cada problema los valores y las probabilidades puede ser muy complejo.
  - Ejemplo: Niños y niñas en familias de cuatro hijos

- Lo simplificamos mediante distribuciones de probabilidad
  - Discretas: Distribución Bernouilli/ Binomial
  - Continuas: Distribución normal

Proceso de Bernouilli

- Estudia fenómenos con las características:
  - Existen dos posibles resultados:
    - Aceptable-Defectuoso, Niño-Niña,....
  - La proporción (probabilidad) de aparición de cada tipo es constante:
    - *P(Aceptable)=constante*
    - *P(Defectuoso)=1-P(Aceptable)*
  - Las observaciones son independientes
- Ejemplos:

Proceso de Bernouilli: Ejemplos

- Una máquina produce piezas de dos tipos:
  - defectuosas en un 2% de los casos. Y aceptables en el 98%.
  - Probabilidades fijas $P(D)=0.02$ $P(A)=0.98$
  - Las observaciones son independientes
- Sexo de los niños nacidos en un hospital:
  - Dos tipos: niños y niñas
  - Probabilidades fijas $P(\text{Niño})=0.51$ $P(\text{Niña})=0.49$
  - En un hospital acaba de nacer un niño: No da información sobre el sexo del siguiente nacimiento.

Distribución Binomial
Para procesos de Bernouilli

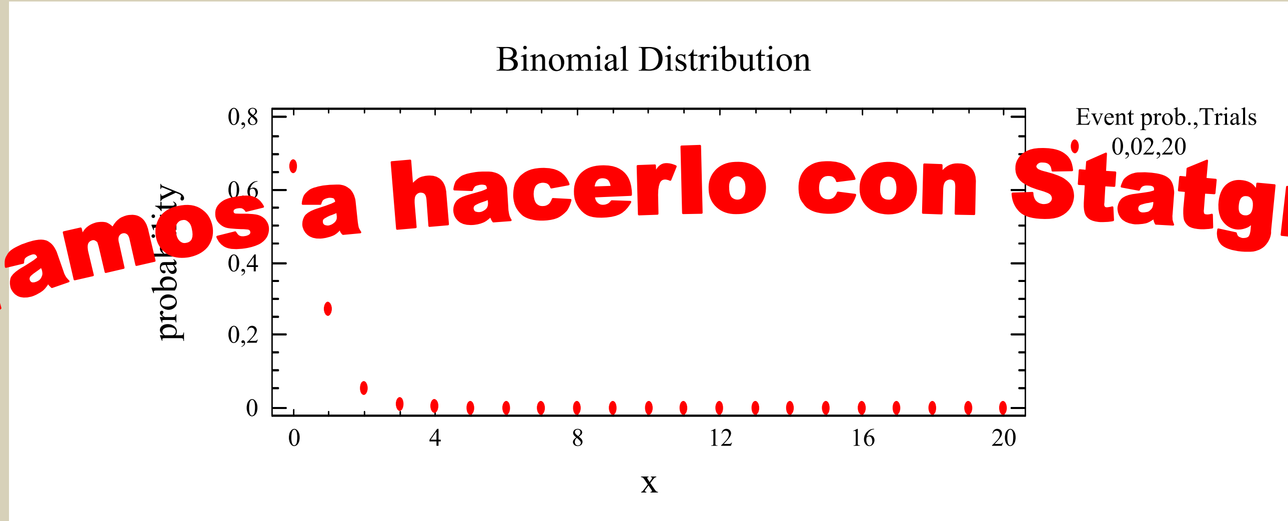
- En un proceso Bernouilli (dos resultados)
- Tomamos una muestra de *n* piezas y miramos cuantas hay de cada tipo.
- La distribución nos proporciona la probabilidad de obtener un número de piezas defectuosas

Distribución Binomial
Para procesos de Bernouilli

- Ejemplo:
  - Nuestro proceso produce un 2% de piezas defectuosas cuando *todo funciona bien.*
  - Tomamos muestras de tamaño 20.
  - La distribución binomial proporciona la probabilidad de obtener un número concreto de defectuosos.

  - Se dice que el número de defectuosos en la muestra, X, se distribuye como una binomial de probabilidad 0,02 y tamaño 20:)
    - B(0.02 ; 20)

Distribución binomial
B(20; 0.02)



Vamos a hacerlo con Statgraphics

Probabilidad de defectuoso 2% : P(D)=0.02

Muestras de tamaño n=20

X='Número de defectuosos en la muestra de 20 0bservaciones'

P(tener 0 defectuosos)=66,7%=P(X=0)

P(tener 1 defectuoso)=27,24%=P(X=1)

P(tener 2 defectuosos)=5,28%=P(X=2)

P(tener 3 defectuosos)=0,64%=P(X=3)

P(tener 4 defectuosos)=0,05%=P(X=4)

Distribución binomial

Probabilidad de que nazca un niño: P(Niño)=0.51

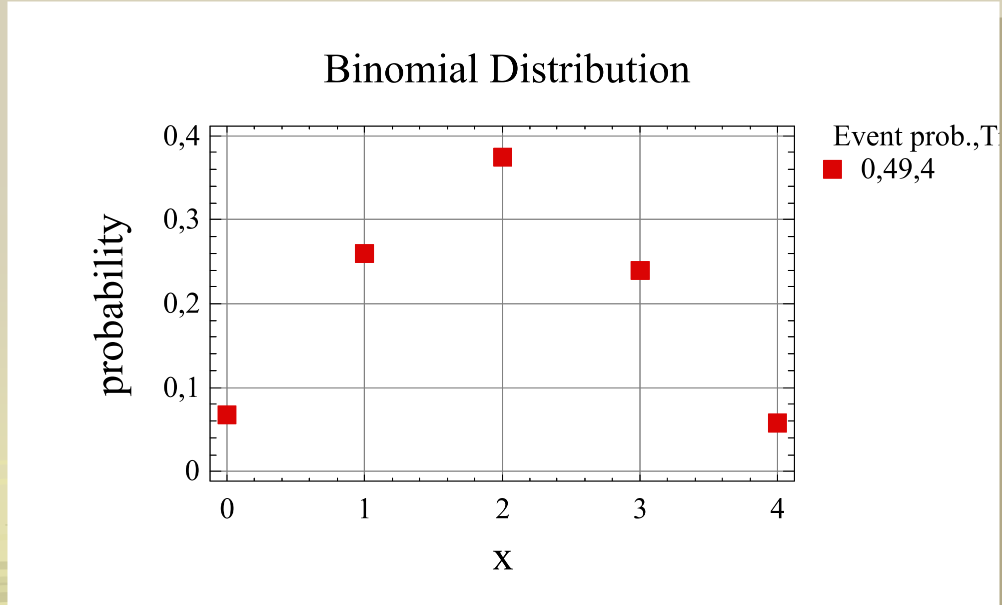
Probabilidad de que nazca una niña: P(Niña)=0.49

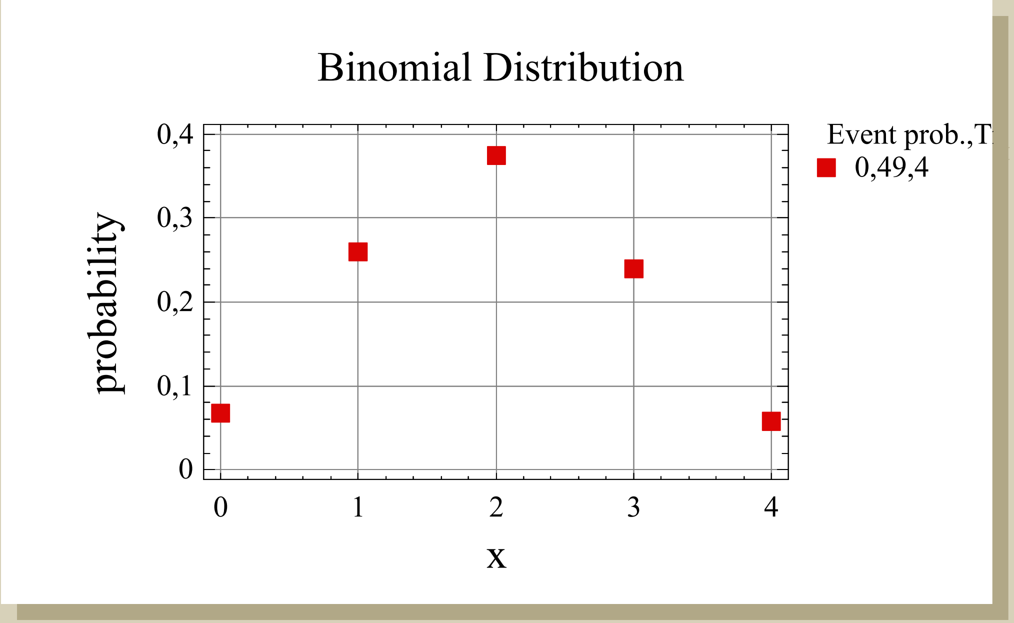
Familias con cuatro hijos.

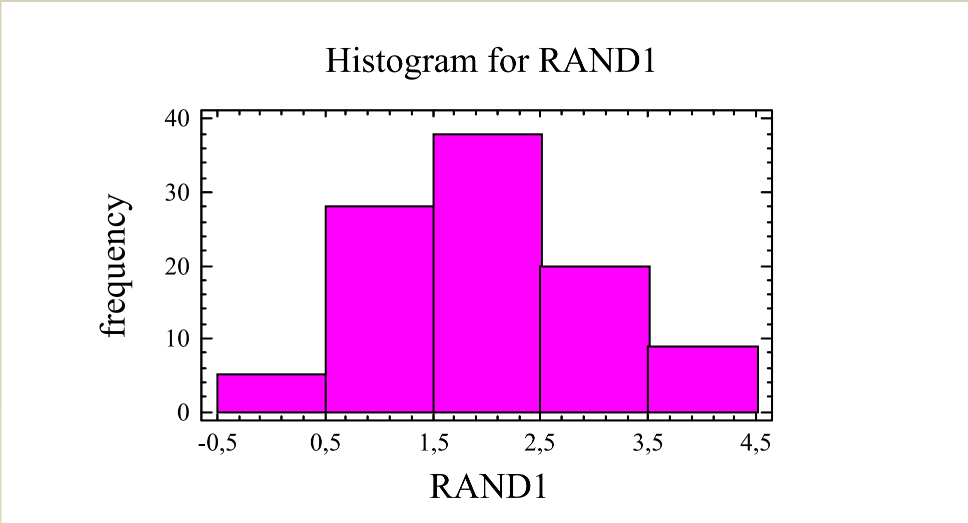
X='Número de niñas'

Obtenemos:

Número de niñas X	Probabilidad
0	**0,067**
1	**0,259**
2	**0,374**
3	**0,240**
4	**0,057**

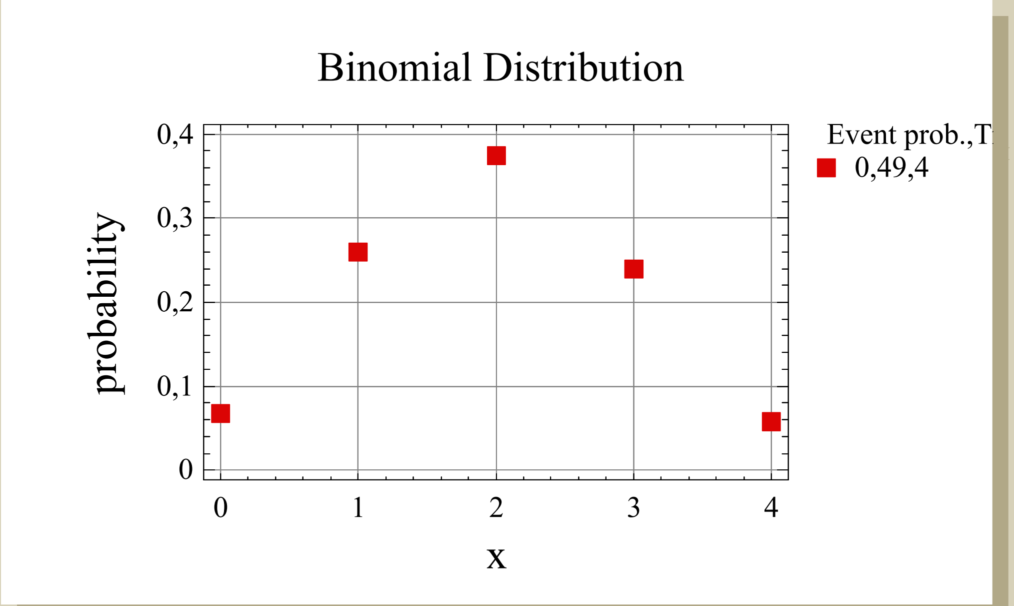


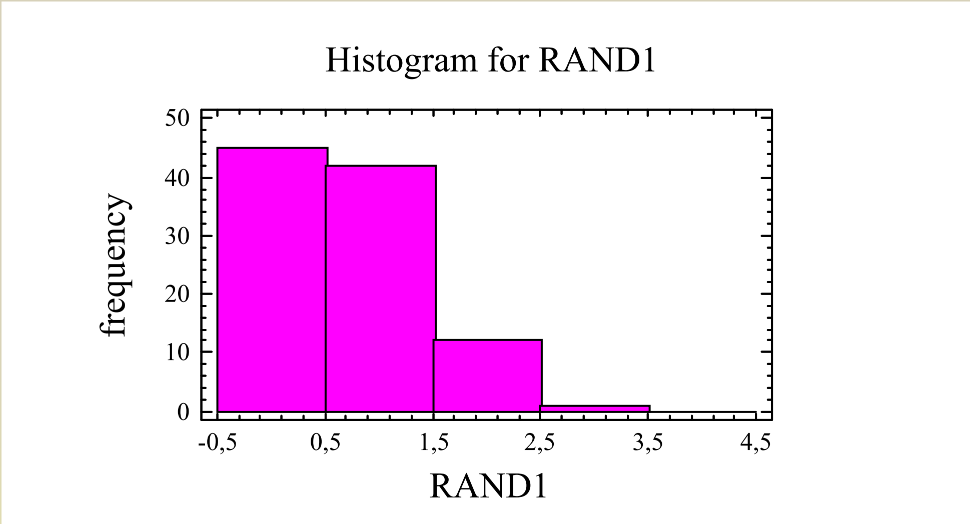




Hemos encontrado 134 familias con cuatro hijos y encontramos esos datos.

Si hubiéramos encontrado……….





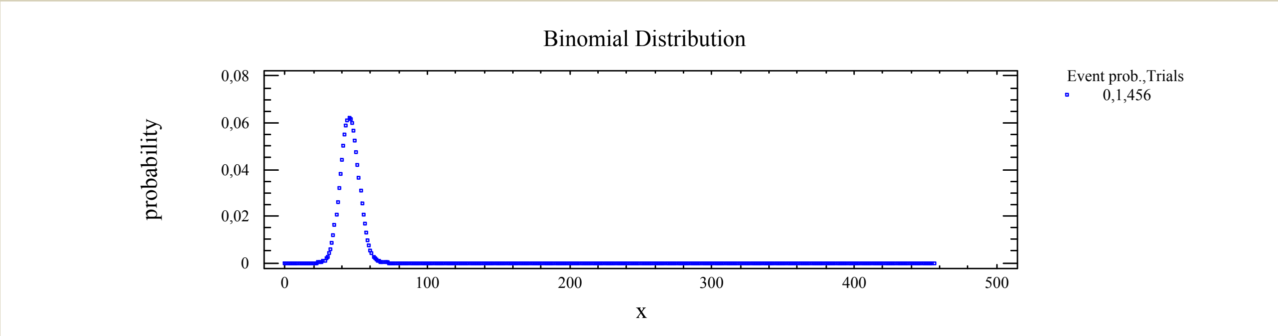
¿qué hay que pensar?

Repetimos el análisis para reclamaciones:

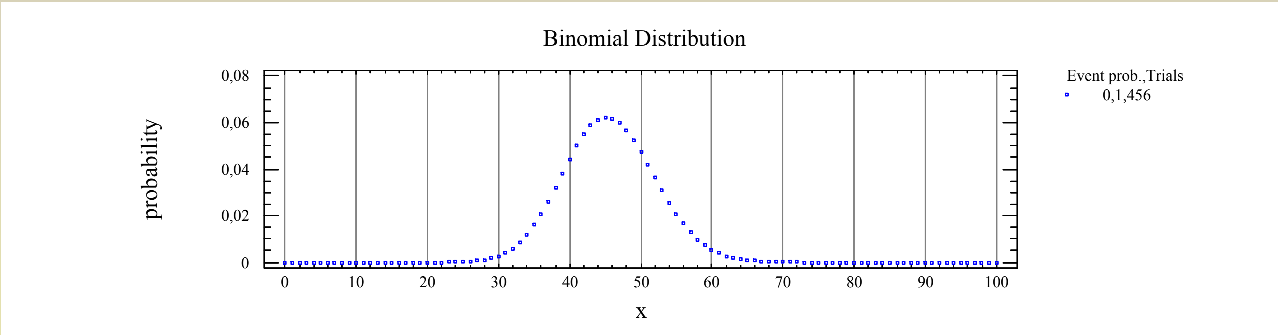
En nuestra empresa el 10% de los servicios acaban en una reclamación.

Este mes se han prestado 456 servicios. Esperaríamos 456x0,10=45 reclamaciones

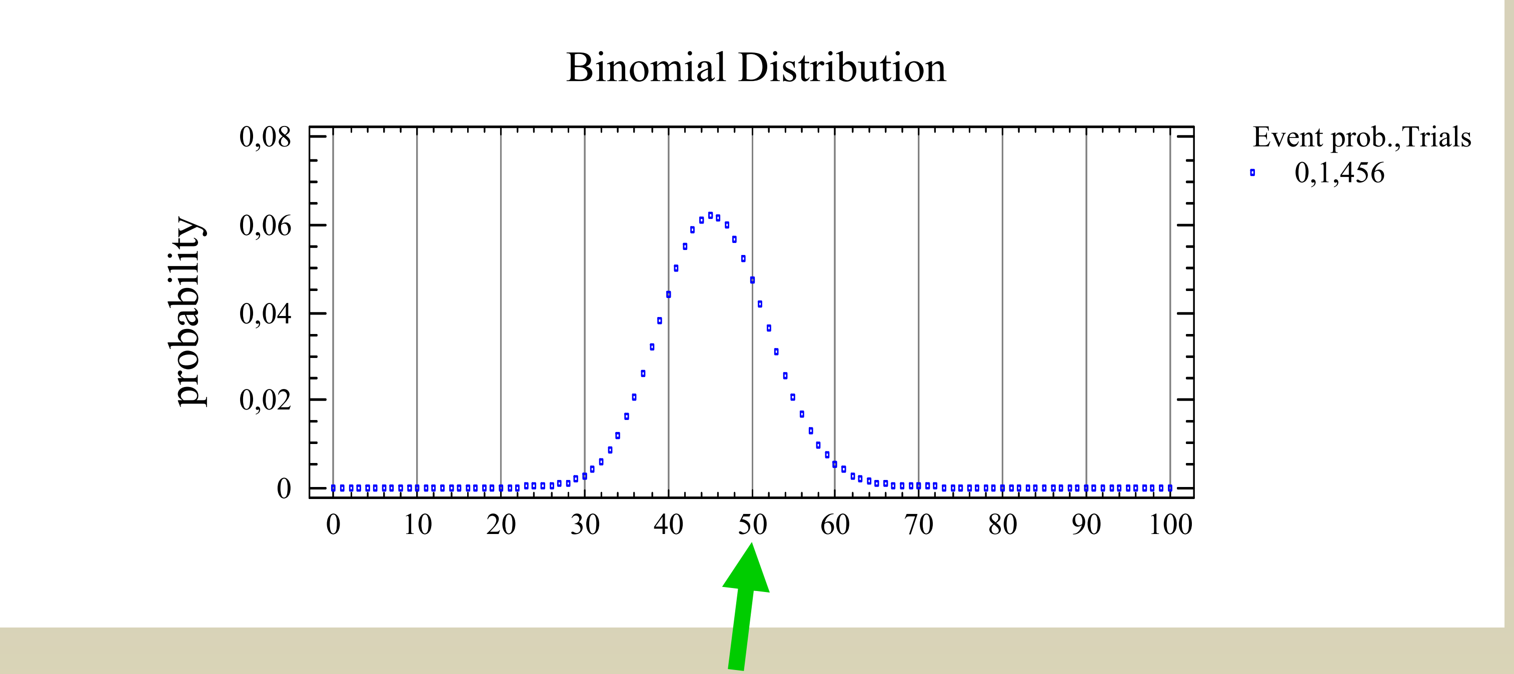
Y ha habido 50 reclamaciones. ¿Hemos trabajado mal?



Hacemos un zoom



Haciendo un zoom mayor



P(x>50)=21.9%

P(x mayor o igual que 50)=26,7%

¿Y si hubiéramos recibido 30 reclamaciones?

P(x<30)=0,04

Resumiendo:

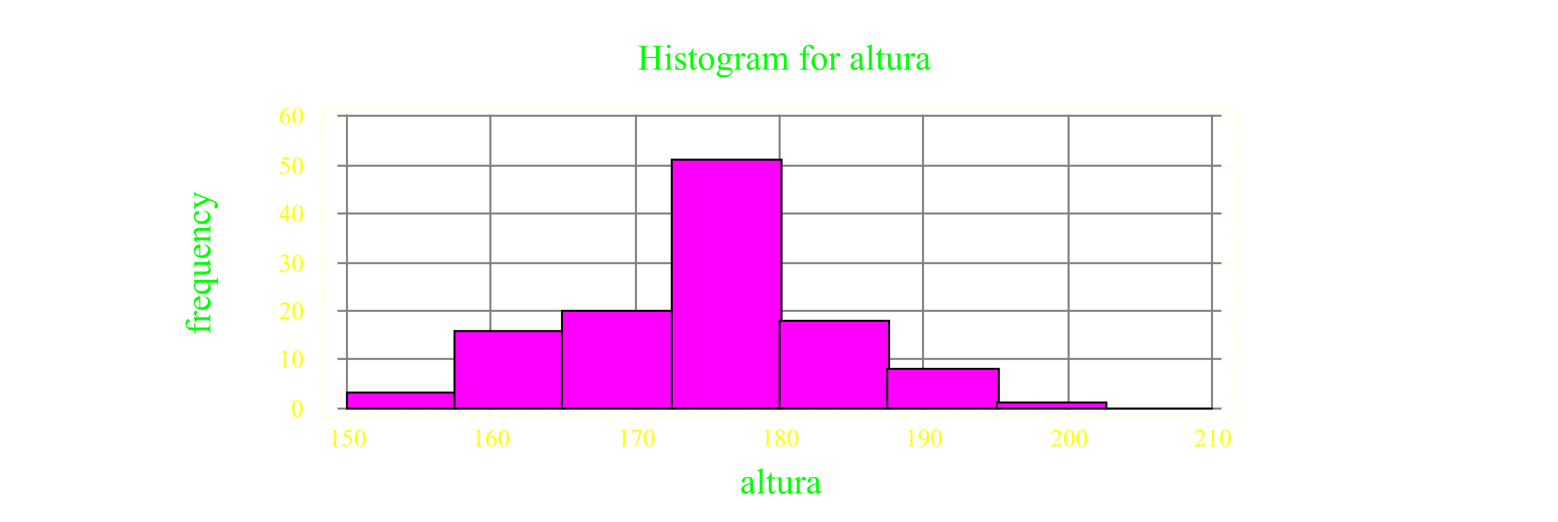
- La distribución binomial sirve para estudiar número de defectuosos de un total de n

- Es decir proporciones: estudiamos si la proporción de defectuosos permanece constante, aumenta o disminuye

- Ya tenemos una manera saber si mejoramos, nos mantenemos o empeoramos.

Distribución normal

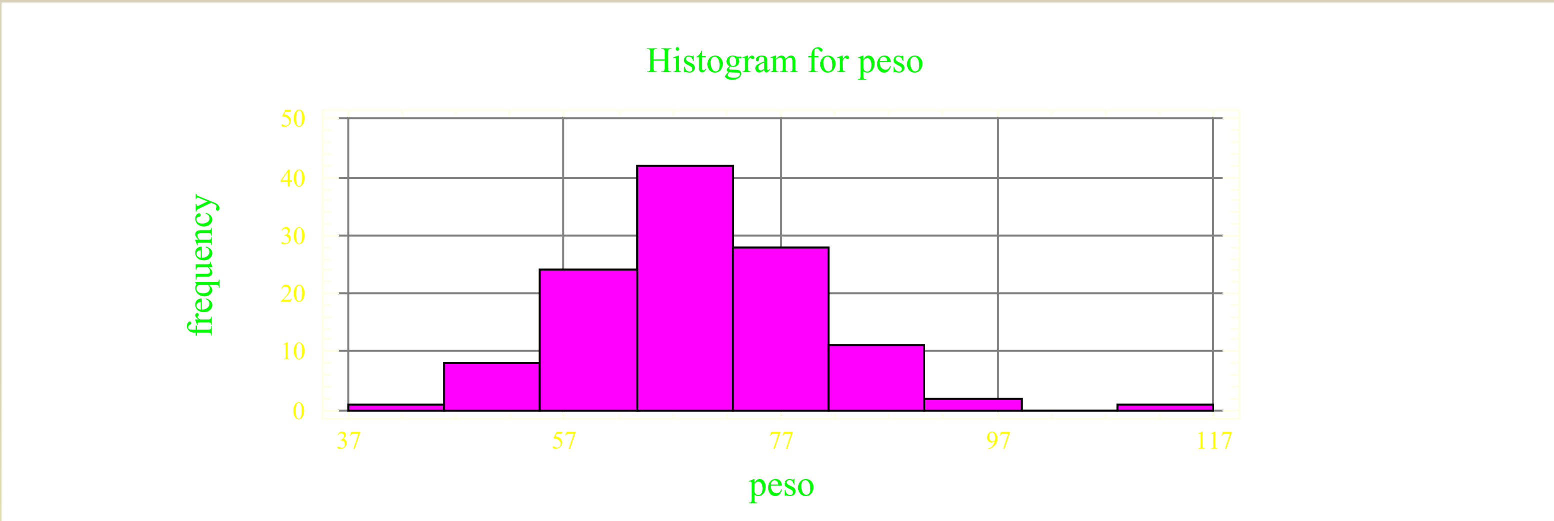
- Para variables continuas



¿Cómo será la altura en general?

Distribución normal

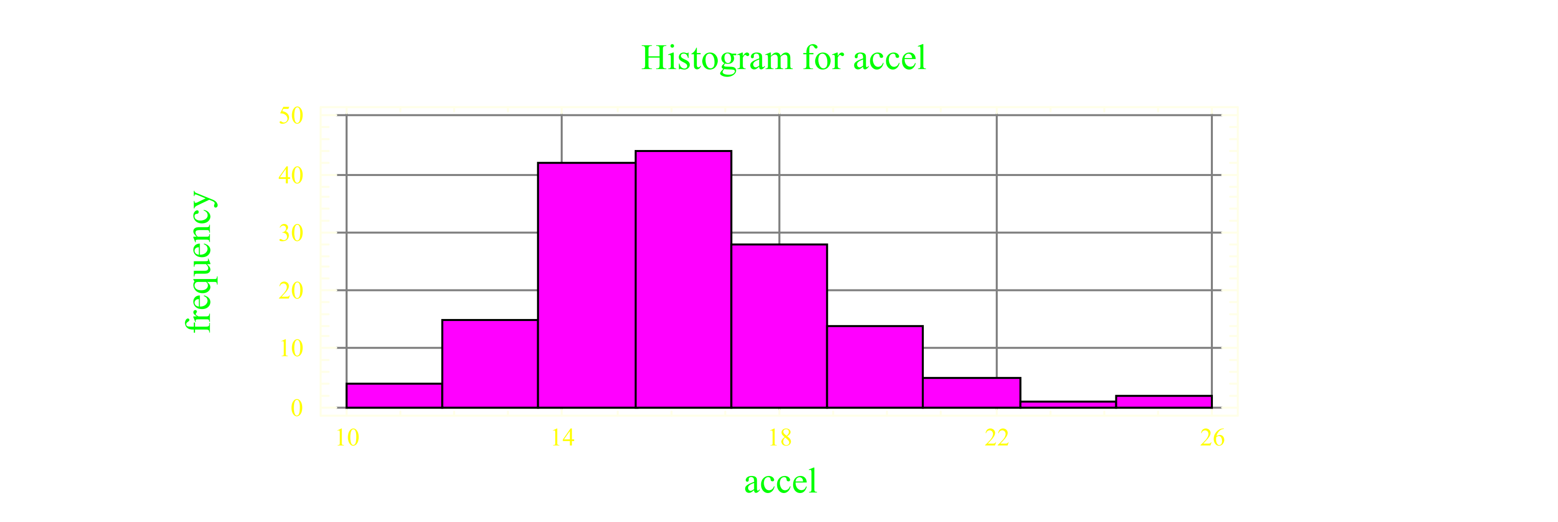
- Para variables continuas



¿Cómo será el peso en general?

Distribución normal

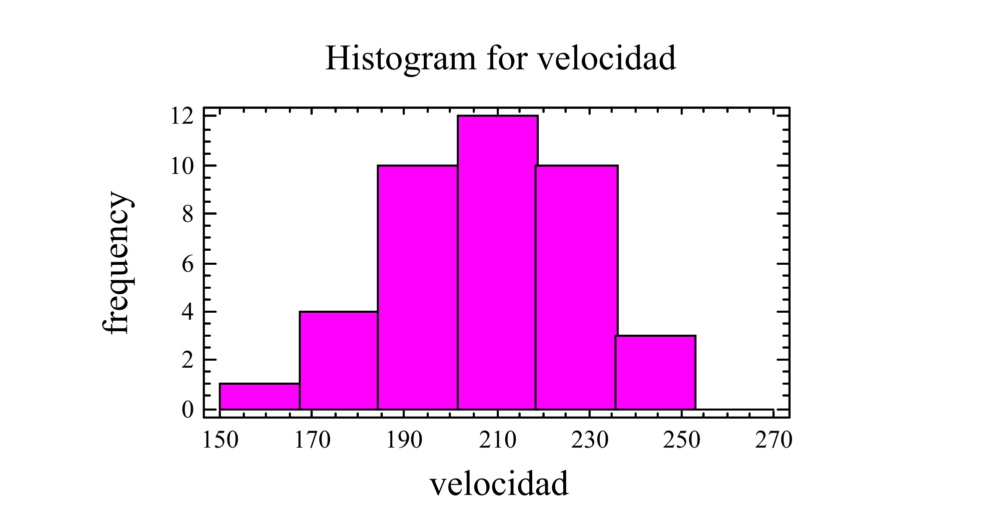
- Para variables continuas



¿Cómo será en general la aceleración?

Distribución normal

➢ Para variables continuas

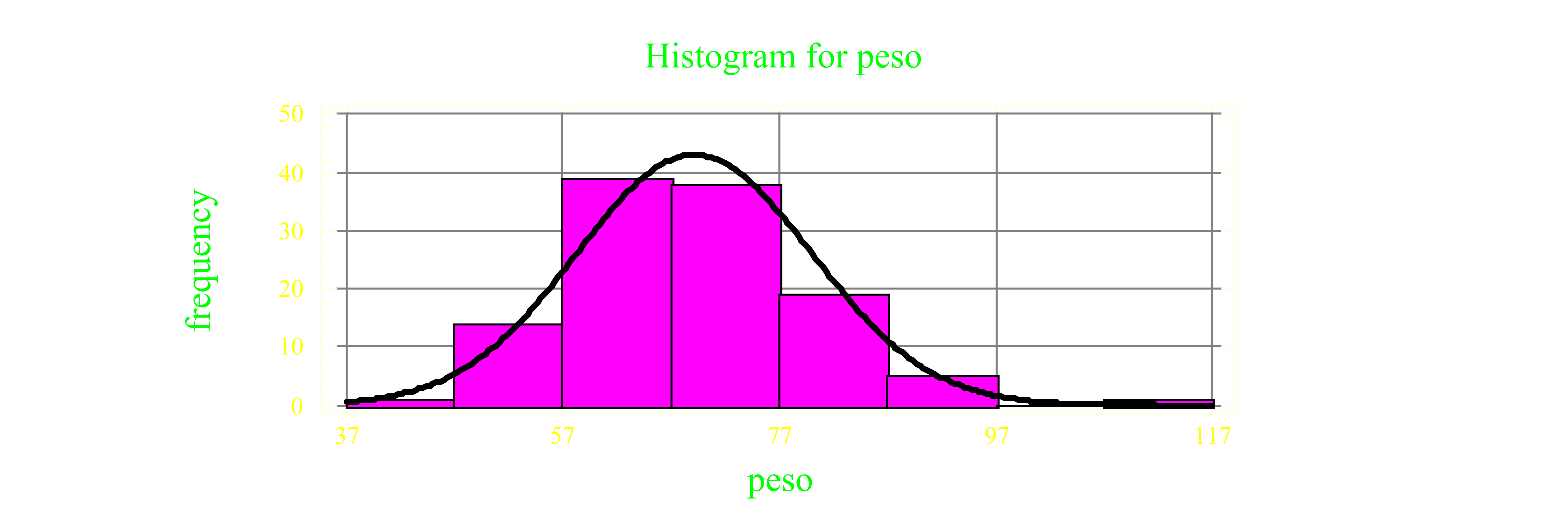


¿Cómo será en general al velocidad?

Distribución normal

Queremos encontrar una curva que represente los datos y

Usarla para generalizar nuestras conclusiones.

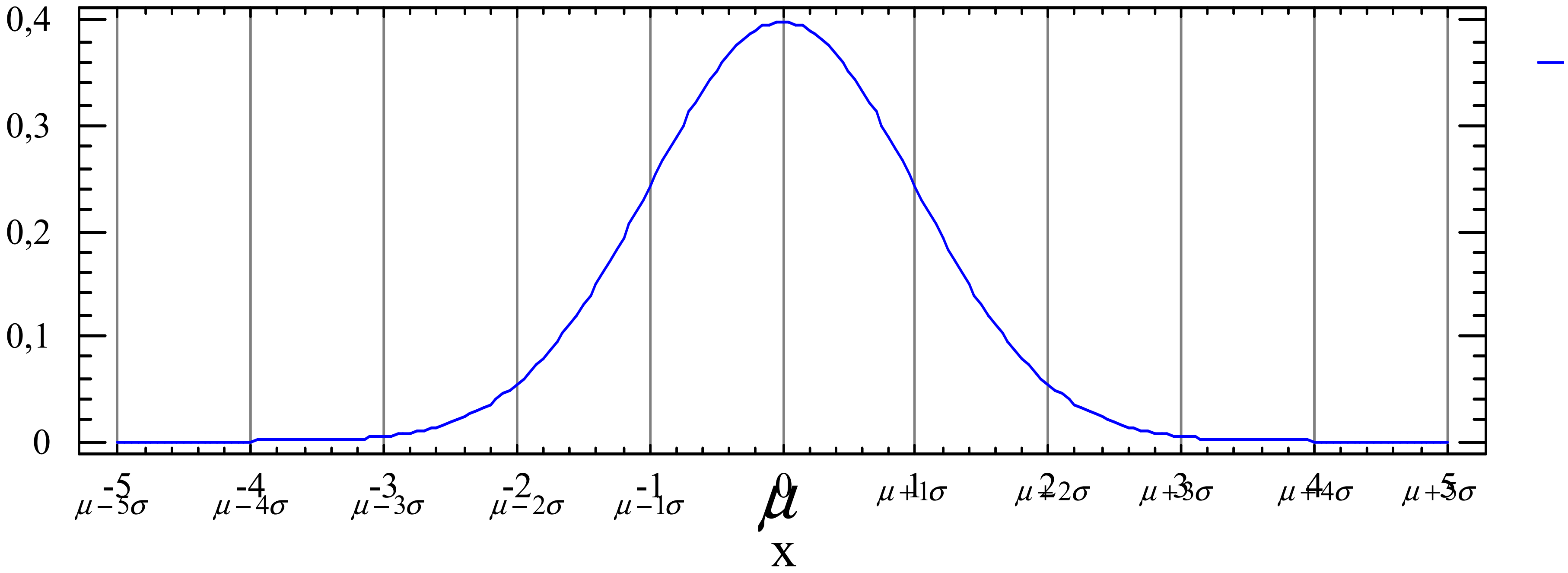


Distribución Normal

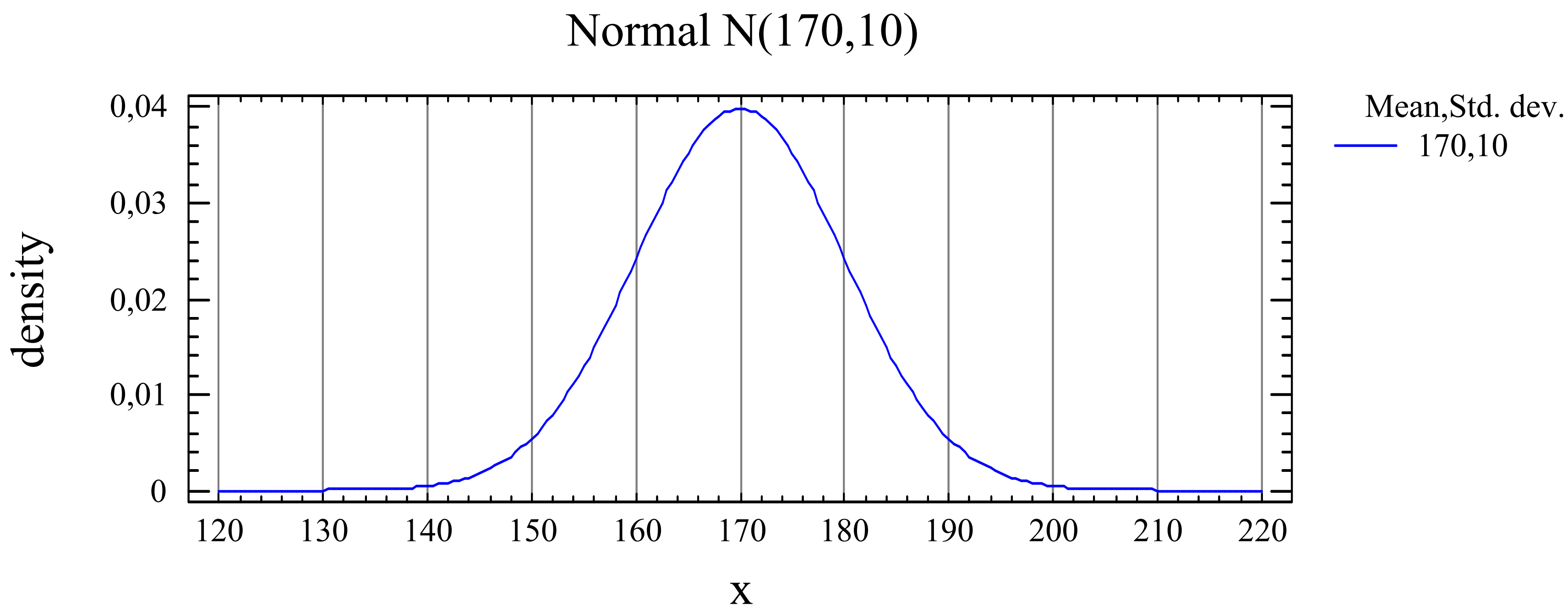
- La distribución normal se llama también CAMPANA DE GAUSS.
- El área que abarca bajo ella es 1
- Se caracteriza por dos valores: Media μ y la desviación típica σ
- La normal se escribe:

$$N(\mu,\sigma)$$

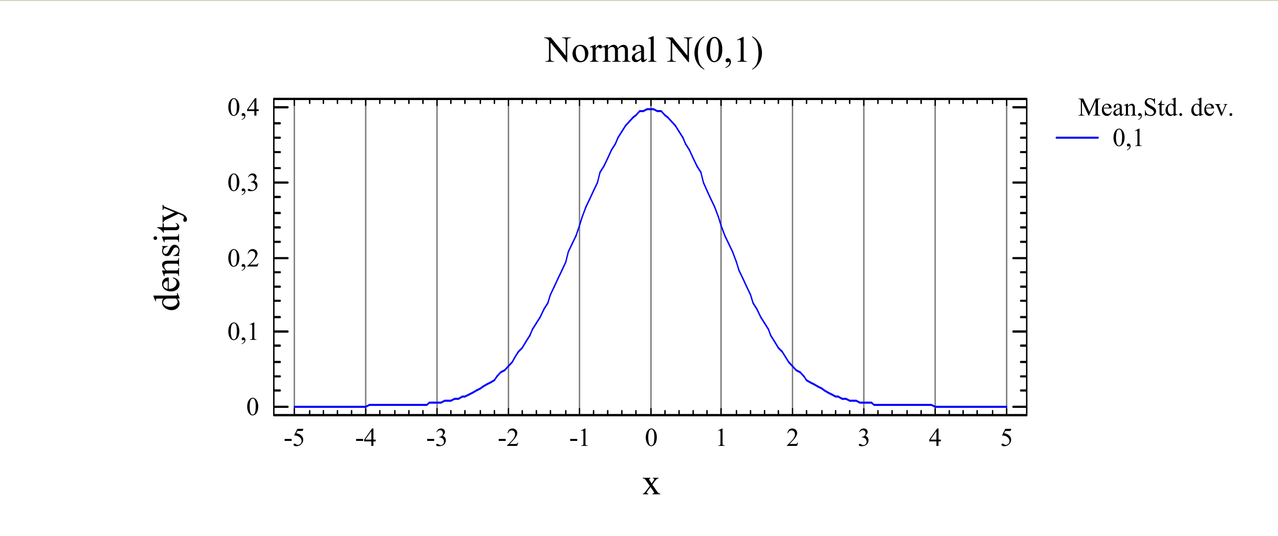
Distribución Normal



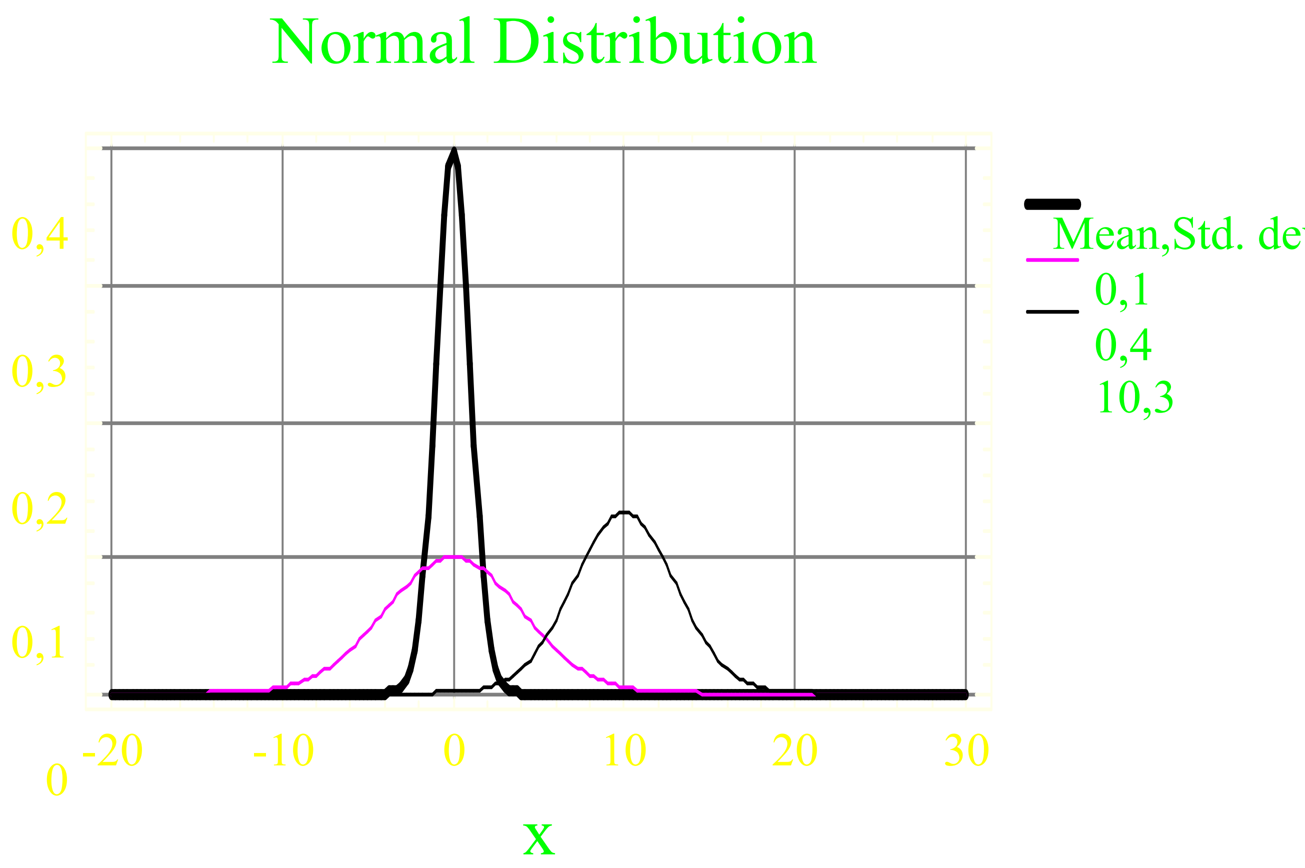
Distribución Normal



Distribución Normal

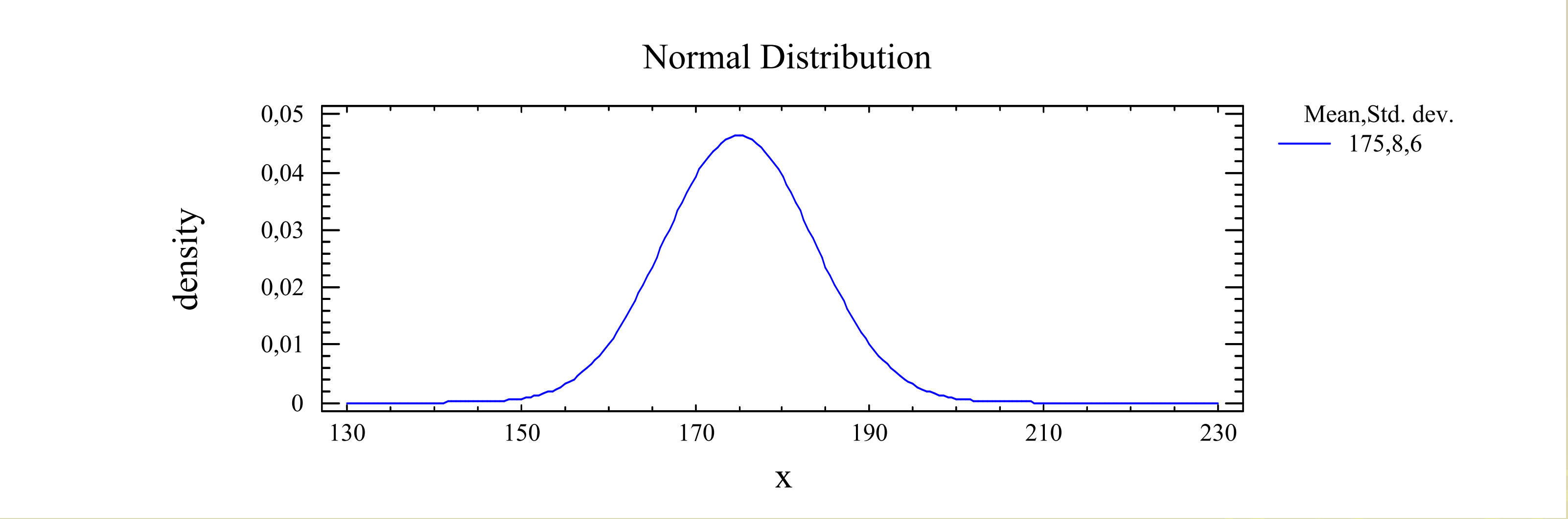


Distribución Normal

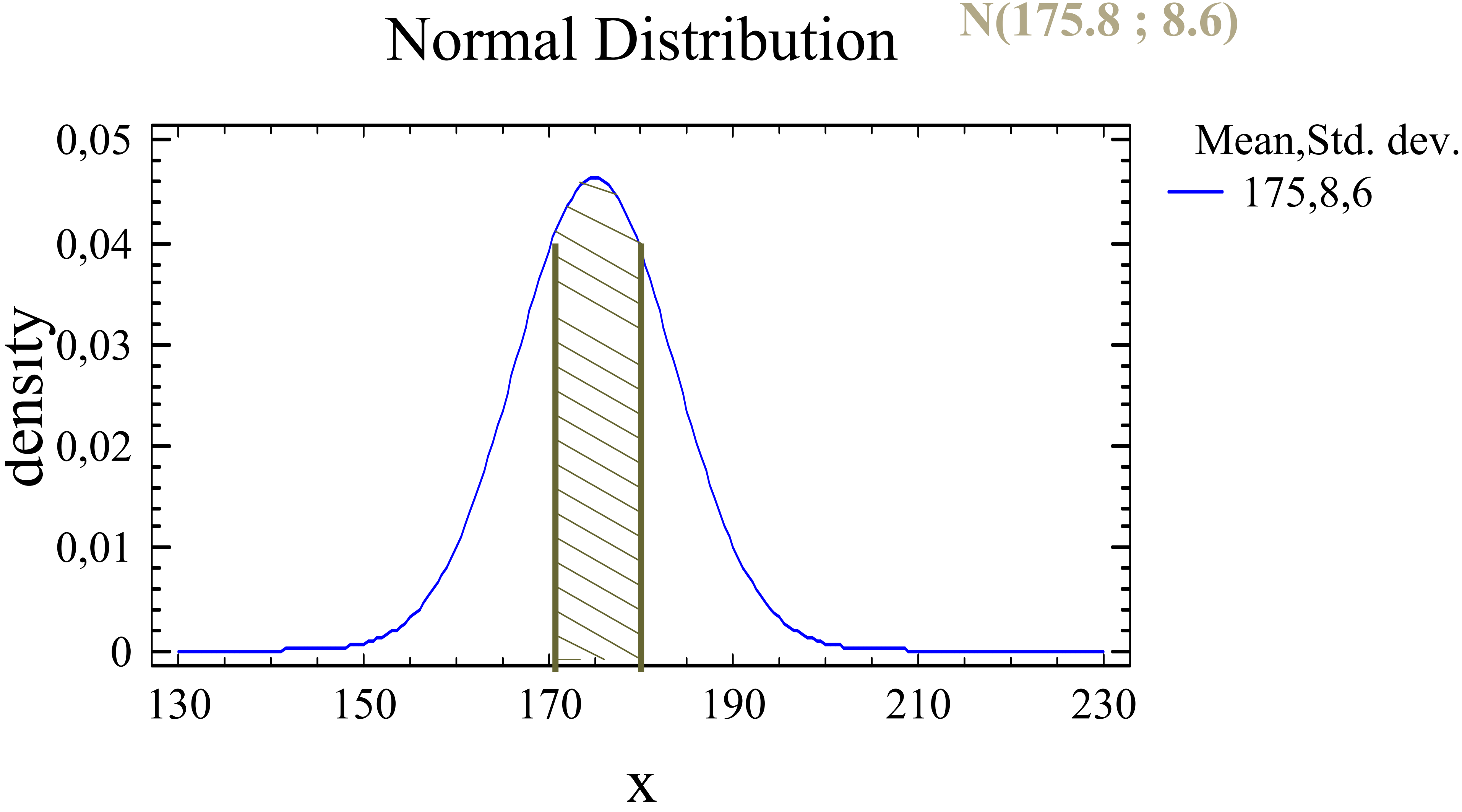


Distribución Normal

El área comprendida bajo la curva normal es la probabilidad de encontrar observaciones en ese intervalo.

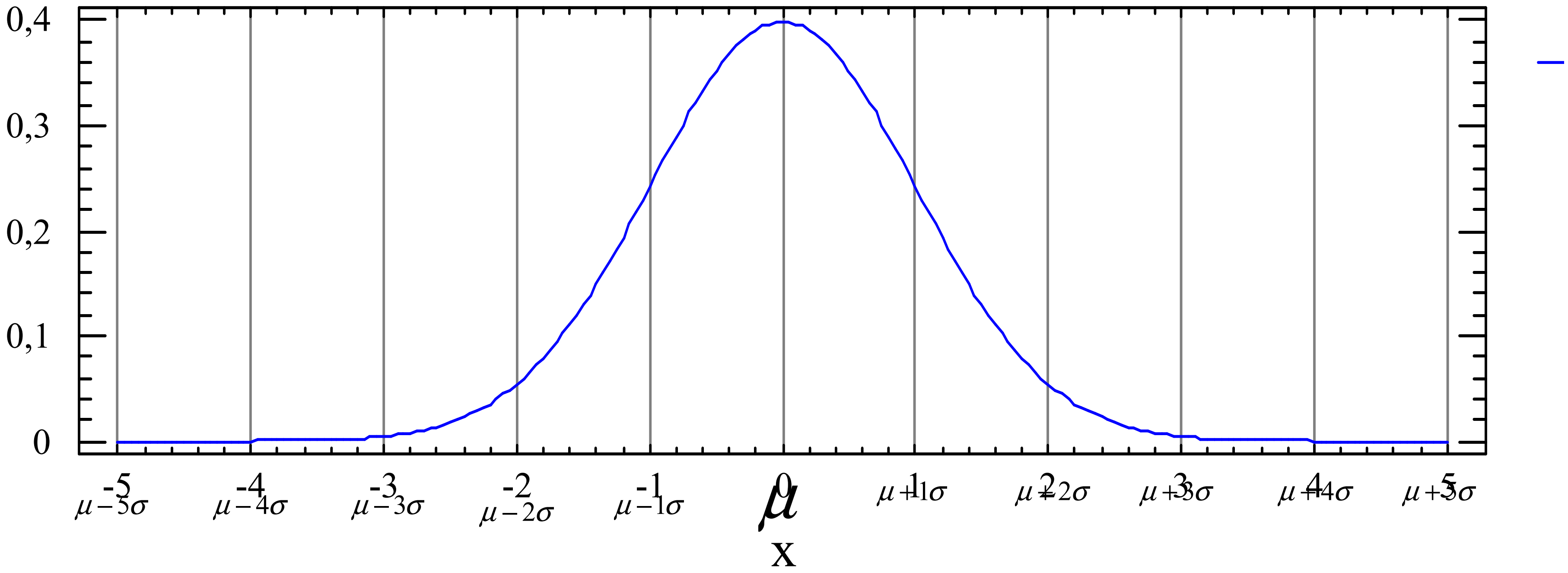


N(175.8 ; 8.6)



P(170<X<180)=Area bajo la curva. Lo calcula la máquina

Distribución Normal: Algunas probabilidades

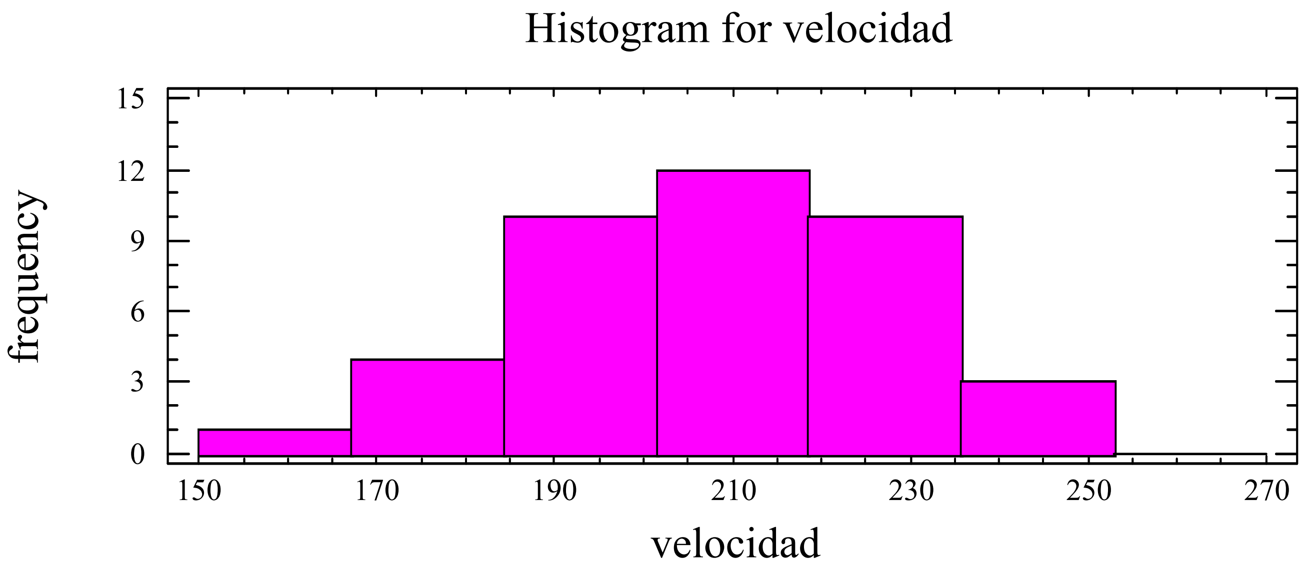


Entre $\mu-2\sigma$ y $\mu+2\sigma$ hay un 95% de probabilidad

Entre $\mu-3\sigma$ y $\mu+3\sigma$ hay un 99.7% de probabilidad

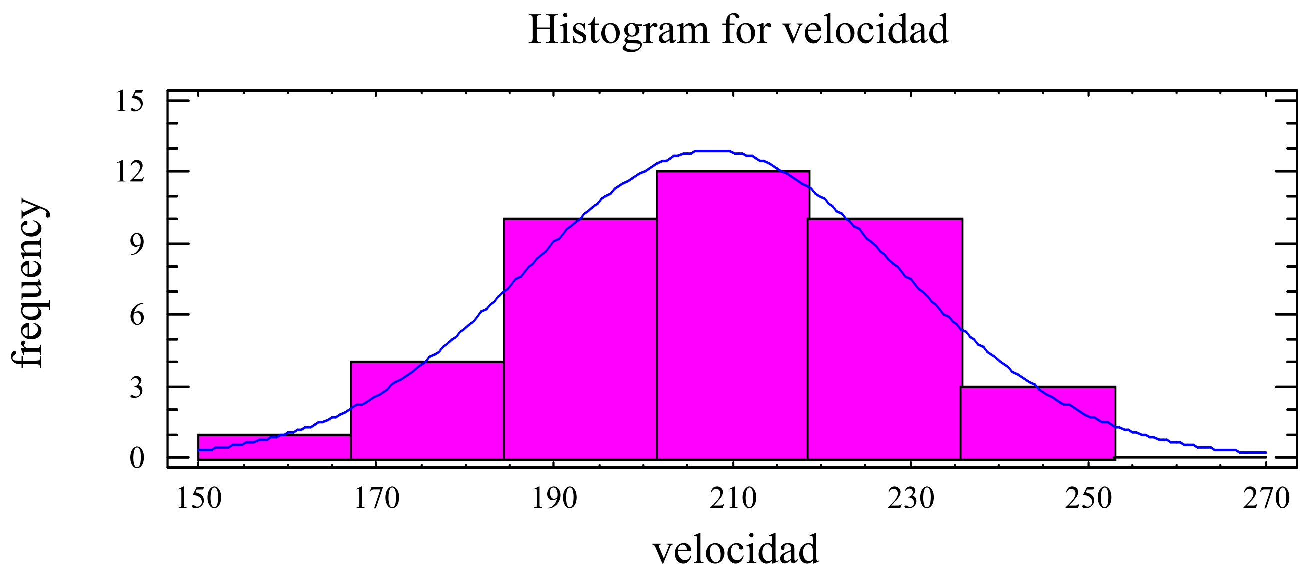
Cálculo de probabilidades con la normal

- El objetivo de las distribuciones es AJUSTARLAS a los datos.
- Este proceso se llama estimación. Y consiste en encontrar la normal que mejor ajusta a los datos:



Cálculo de probabilidades con la normal

- El objetivo de las distribuciones es AJUSTARLAS a los datos.
- Este proceso se llama estimación. Y consiste en encontrar la normal que mejor ajusta a los datos:



Se puede demostrar

- La normal que mejor ajusta a los datos es

$$N(\mu, \sigma) = N(Media, desviación)$$

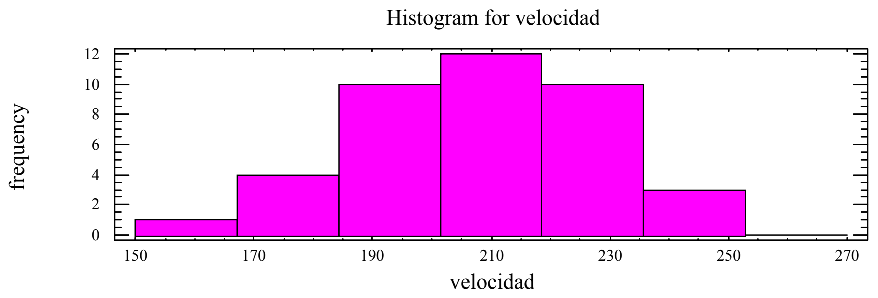
Ejemplo velocidad

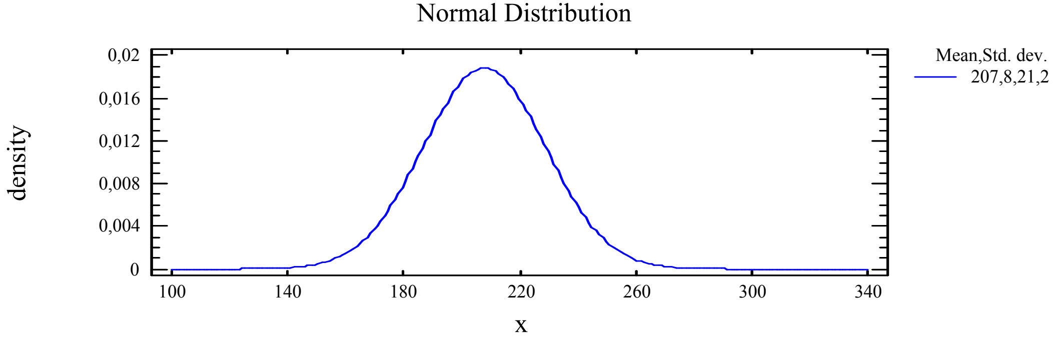
- La velocidad de una muestra de coches bastante deportivos, tiene un valor medio
- Media=207,82
- Desviación típica=21,219

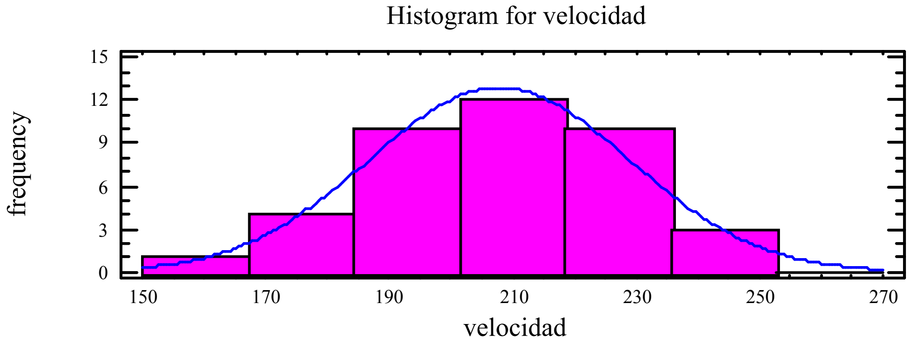
- POR TANTO LA NORMAL QUE AJUSTA ESTOS DATOS ES:

$$N(207.8 ; 21.2)$$

- Veámoslo gráficamente.







Summary Statistics for velocida

Count = 40
Average = 207,821
Variance = 450,246
Standard deviation = 21,219
Minimum = 163,24
Maximum = 252,49
Range = 89,25
Stnd. skewness = -0,0390404
Stnd. kurtosis = -0,443825

Una vez estimada la NORMAL Podemos calcular probabilidades N(207.8, 21.2)

- P(Velocidad < 200)=0.356
- P(Velocidad < 250)=0.977
- P(200<Velocidad<250)= 0.977 -0.356=0.621
- En STATGRAPHICS vamos a:
  - describe

    Distributions

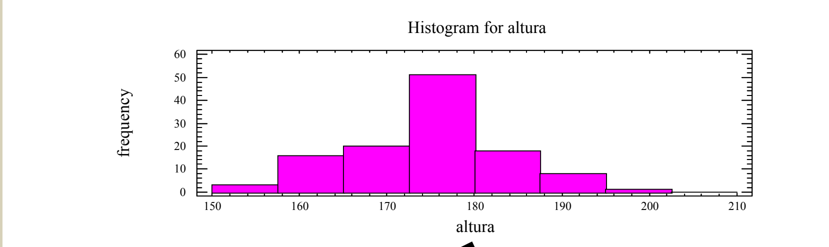
    Distribution fitting (uncensored)

    Ajustamos la normal

    Tail Areas

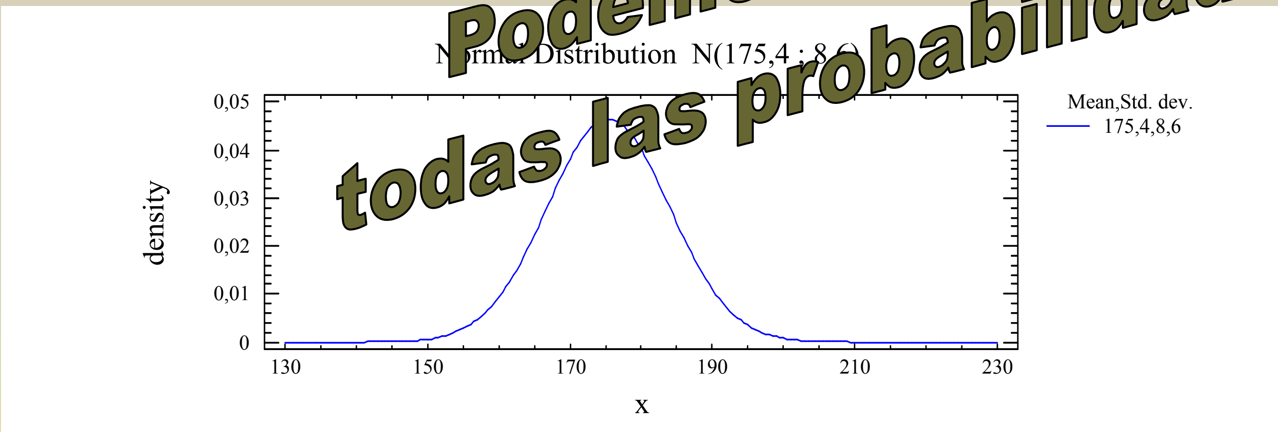
Otro ejemplo: Alturas de los estudiantes

```
Analysis Summary

Data variable: altura

117 values ranging from 155,0 to 199,0

Fitted normal distribution:
     mean = 175,427
     standard deviation = 8,63066
```



N(175.4 ; 8.6)

Podemos calcular todas las probabilidades



En Tail areas obtenemos:

P(Altura<170) = 0,265031

P(Altura<180) = 0,703636   P(170<Altura<180)=0,70-027=0,43

Proceso de estimación (resumen)

- Miramos los datos
- Vemos si parecen normales
- Les ajustamos la distribución normal N(media, desviación)
- Ya podemos calcular probabilidades y...

- TOMAR DECISIONES

Estimación de la binomial: Proporción de defectuosos

- Para estimar la proporción de defectuosos, se toma una muestra grande cuando el proceso se sabe que está bien y se estima

P(D)=número de defectuosos/número total

Ejemplo

- Hemos tomado 200 piezas y obtenido 4 defectuosas:

- P(D)=4/200=0.02 (El 2%)
- Con este dato ya podemos calcular probabilidades:
- Tomamos normalmente muestras de tamaño 20. Como P(D)=0.02

- X=Número de defectuosos en 20 con P(D)=0.02 es una Binomial.
  - B(20, 0.02)

INTERVALOS de confianza

- Los intervalos de confianza proporcionan una zona en la que previsiblemente estará el valor auténtico del parámetro.
- Se calculan con una confianza determinada
- Normalmente 95%

INTERVALOS de confianza

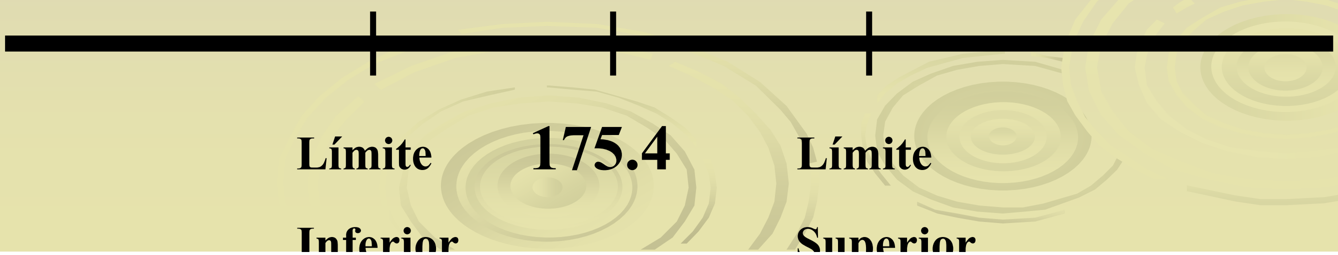
- Se calculan para:
  - Media
  - Varianza o desviación típica
  - Proporciones

INTERVALOS DE CONFIANZA MEDIA

- Hemos estimado la altura media de los estudiantes de ingeniería en

$$\boxed{\bar{x} = 175.4}$$

Un intervalo de confianza proporciona una zona en la que con una confianza predeterminada estará la altura media de verdad de esa población



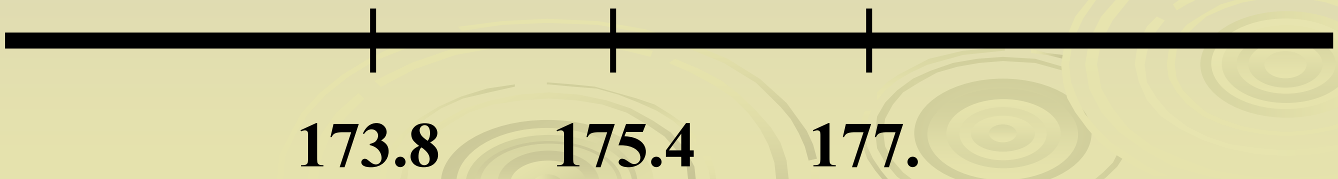
INTERVALOS DE CONFIANZA MEDIA

- Statgraphics:
  - Describe-One variable-Confidence intervals

Confidence Intervals for altura

95,0% confidence interval for mean: 175,427 +/- 1,58035 [173,847;177,008]

95,0% confidence interval for standard deviation: [7,64853;9,90463]



INTERVALOS DE CONFIANZA MEDIA

- La fórmula que sirve para calcularlo está en la documentación y es:

$$\bar{x} - t_{\alpha/2}\hat{s}/\sqrt{n} \qquad \bar{x} + t_{\alpha/2}\hat{s}/\sqrt{n}$$

n=número de datos

t es una distribución que viene en las tablas

s es la desviación típica

Si n>30 y la confianza es del 95% entonces t=2 y el intervalo es:

$$\bar{x} - 2\hat{s}/\sqrt{n} \qquad \bar{x} + 2\hat{s}/\sqrt{n}$$

Intervalo de confianza
Varianza o desviación típica

- Hemos estimado que la media de altura era 175.4
- La desviación es 8.63
- En describe-one variable-confidence intervals
- Fórmulas en documentación.

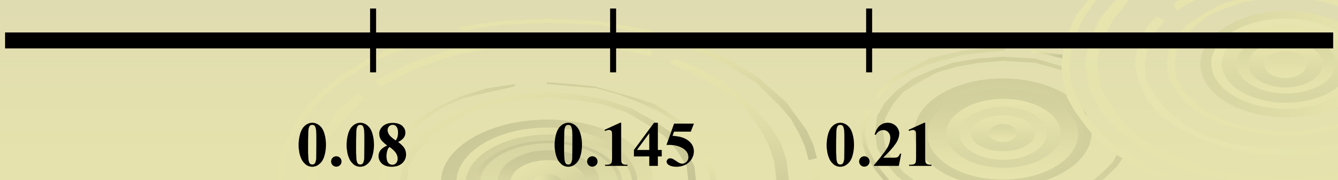
```
Confidence Intervals for altura

95,0% confidence interval for mean: 175,427 +/- 1,58035   [173,847;177,008]

95,0% confidence interval for standard deviation: [7,64853;9,90463]
```

INTERVALOS DE CONFIANZA
Binomial

- Statgraphics:
- Codificando la variable como defectuoso = 1 y aceptable = 0
- Describe-One variable-Confidence intervals
- Tenemos una muestra de tamaño 117 con una proporción de 1 del 14.52 %. El intervalo de confianza será

95,0% confidence interval for mean: 0,145299 +/- 0,0648057 [0,0804935;0,210105]

MEAN = PROPORCION



Contrastes de hipótesis

- Contraste de hipótesis en estadística es contestar a preguntas
- Siempre se contesta en términos de probabilidad.
- Es semejante a intervalos de confianza
- Un intervalo también puede resolver estos problemas.

Contrastes de hipótesis

Problema:

Tenemos un muestra de alturas de 2000 españoles.

Media=173 y desviación=10cm.

Hace 10 años la media era de 168.

¿Ha aumentado la altura de los españoles o es una casualidad de la muestra?

Contrastes de hipótesis

Problema:

Las bolsas de un fabricante soportan 16 Kg tomamos una muestra y, a la vista de su media y desviación ¿Debemos creer al fabricante?

Problema:

La duración de una enfermedad es de 15 días. Se estudia un nueva droga en 20 pacientes y se observa que, en promedio la duración son 13 días. ¿Es mejor la nueva droga? ¿Es una casualidad?

Problema:

Tenemos un muestra de alturas de 2000 españoles. Media=173 y desviación=10cm.

Hace 10 años la media era de 168.

¿Ha aumentado la altura de los españoles o es una casualidad de la muestra?

Metodología:

Definimos un Hipótesis nula H_0 que es lo que se quiere comprobar

y una hipótesis alternativa H_1 que es lo que observamos

H_0: Media=168

H_1: Media>168

H_0: Media=168

H_1: Media distinta 168

Problema:

Tenemos un muestra de alturas de 2000 españoles. Media=173 y desviación=10cm.

Hace 10 años la media era de 168.

¿Ha aumentado la altura de los españoles o es una casualidad de

la muestra?

La decisión se toma en términos de probabilidad de que ocurra H_0 con los datos que obtenemos.

La herramienta para decidir se llama p-valor

Describe

Hipotesis Test

H_0: Media=168

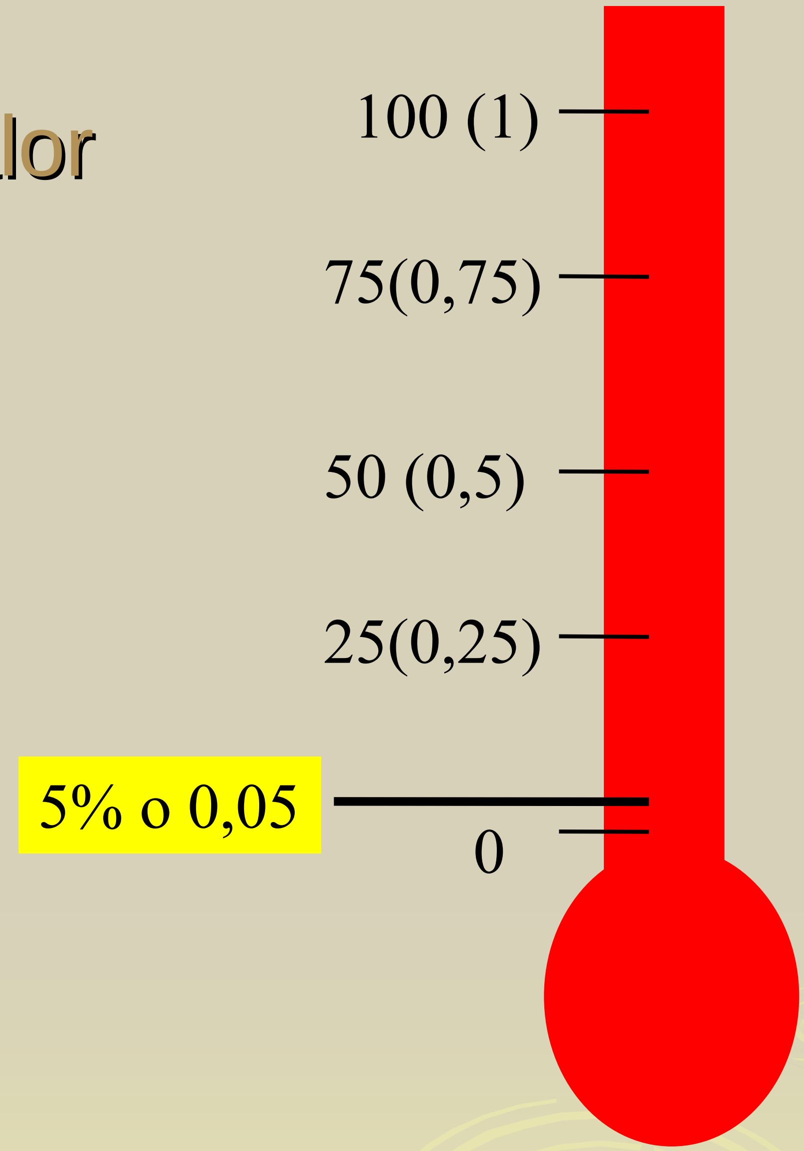
H_1: Media>168

El p-valor

- El p-valor da una idea de lo verosímil que es la hipótesis nula con los datos que tenemos
- P-valor es bajo….Poco razonable que Ho sea verdad
- P-valor alto……Es bastante posibe que Ho sea verdad

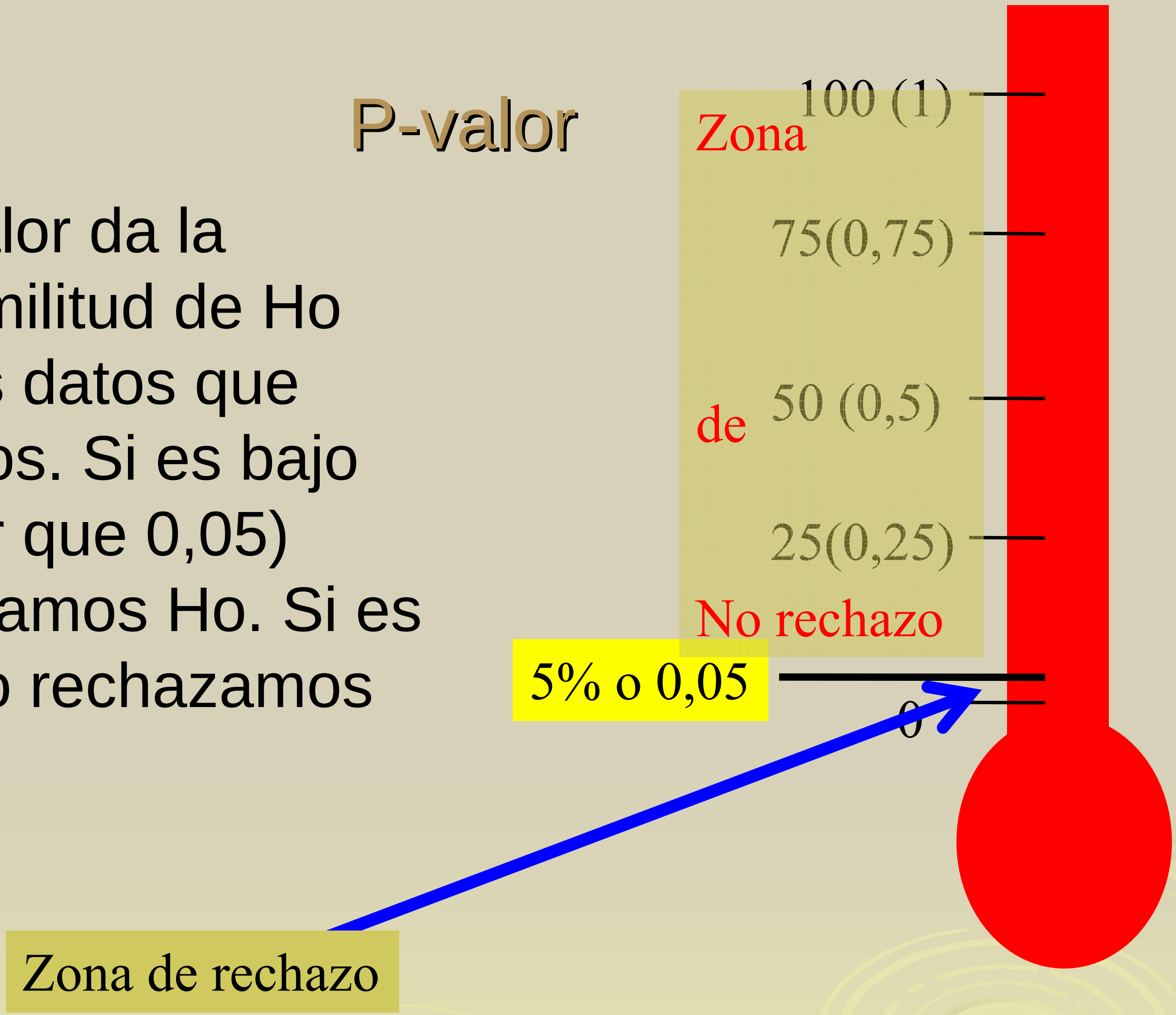
P-valor

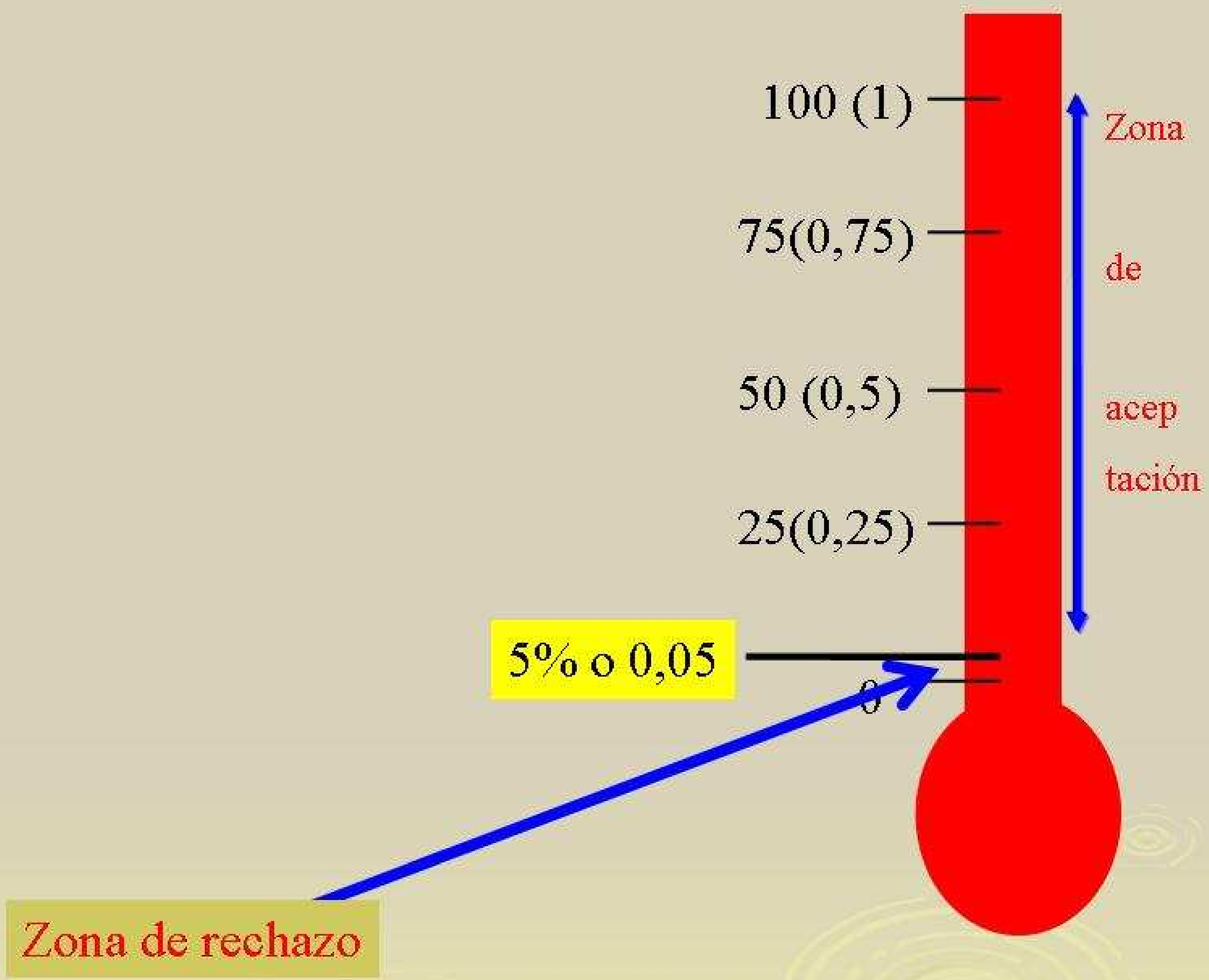
- El p-valor da la verosimilitud de Ho con los datos que tenemos. Si es bajo (menor que 0,05) rechazamos Ho. Si es alto, no rechazamos Ho



P-valor

- El p-valor da la verosimilitud de Ho con los datos que tenemos. Si es bajo (menor que 0,05) rechazamos Ho. Si es alto, no rechazamos Ho




100 (1)
75(0,75)
50 (0,5)
25(0,25)
5% o 0,05
0
Zona
de
acep
tación
Zona de rechazo

Problema:

Tenemos un muestra de alturas de 2000 españoles. Media=173 y desviación=10cm.

Hace 10 años la media era de 168.

¿Ha aumentado la altura de los españoles o es una casualidad de la muestra?

La decisión se toma en términos de probabilidad de que ocurra H_0 con los datos que obtenemos.

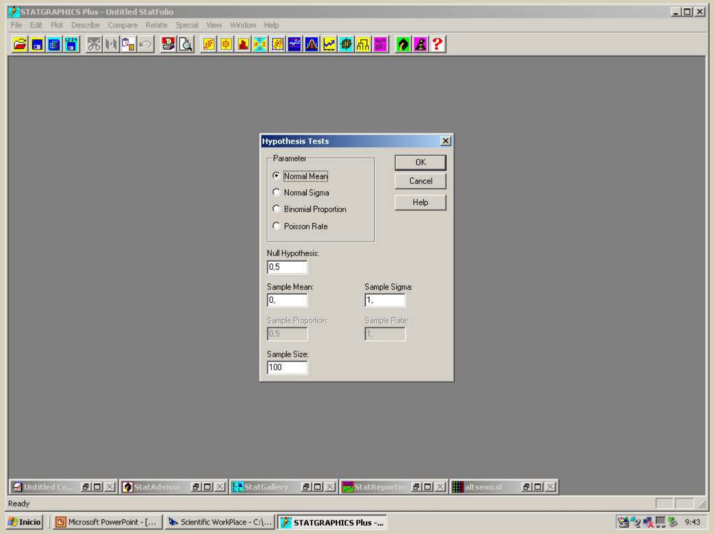
La herramienta para decidir se llama p-valor

Describe

Hipotesis Test

H_0: Media=168

H_1: Media>168

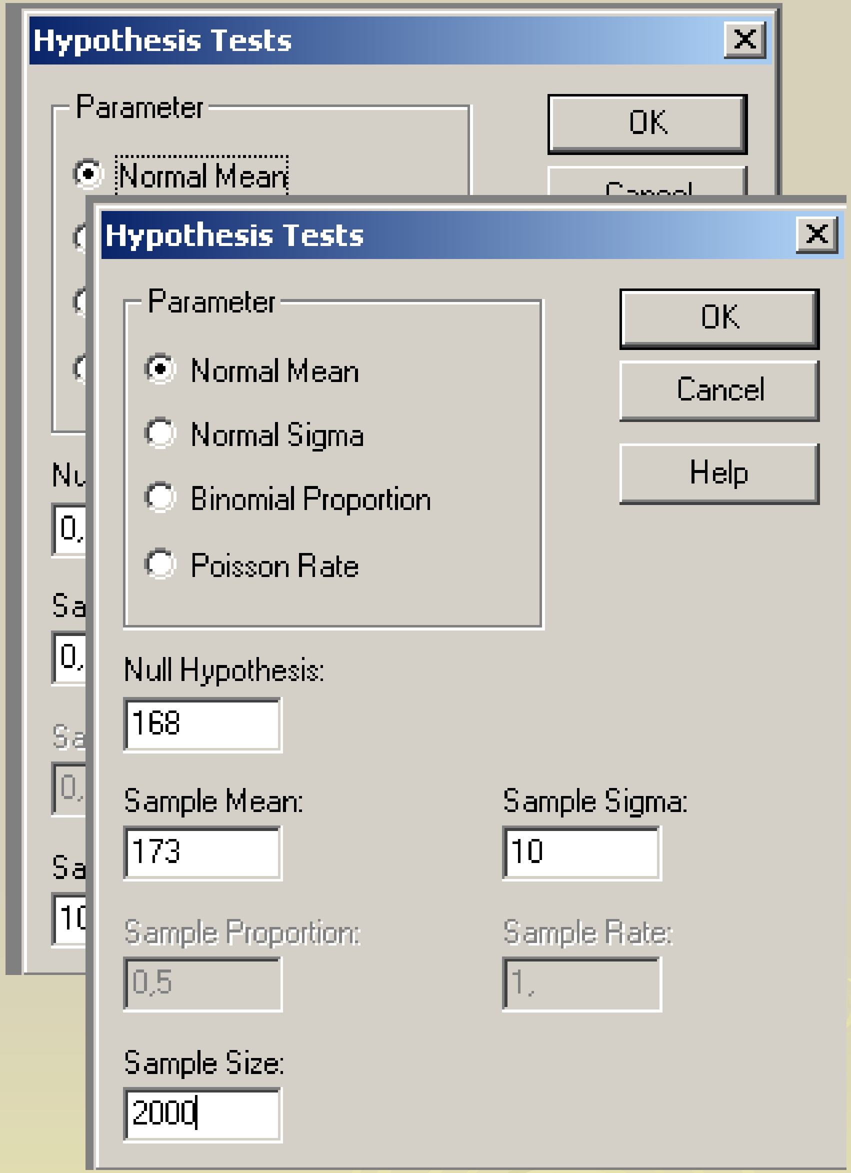


Pide: Null hypothesis. H_0 Media=168

Media de la muestra: 173

Desviación de la muestra: 10

Tamaño de la muestra: 2000



RESULTADOS

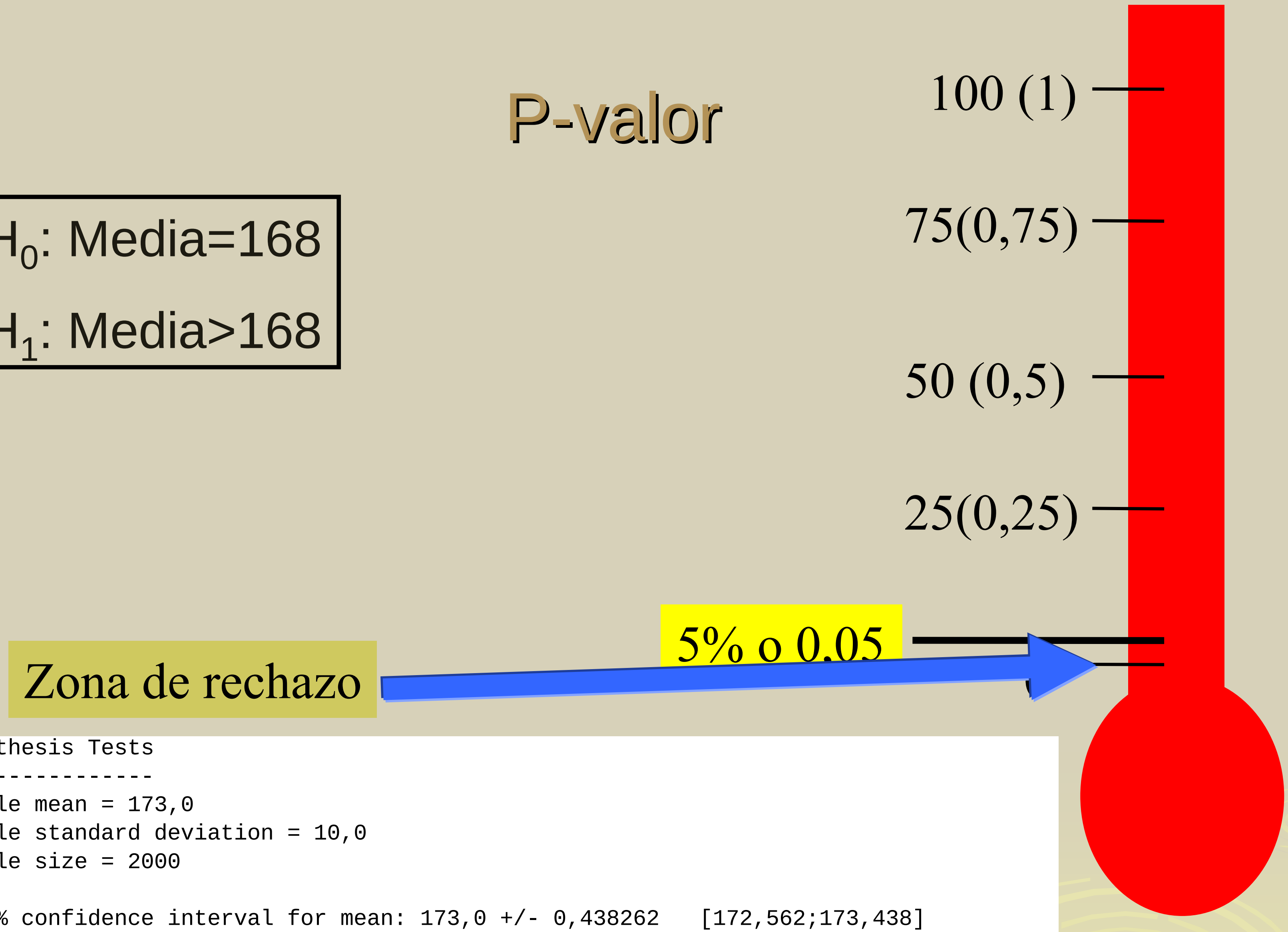
```
Hypothesis Tests
----------------
Sample mean = 173,0
Sample standard de[illegible]iation = 10,0
Sample si[illegible]e = [illegible]000

95,0% con[illegible] [illegible]3,[illegible] +/- 0,438262   [172,562;173,438]

Null Hypo[illegible]: me[illegible] 168,0
Alternati[illegible]e[illegible]t[illegible]qual
Computed [illegible] statistic = 22,3607
P-Value = 0,0
Reject the null hypothesis for alpha = 0,05.
```

Más altos????

P-valor

H_0: Media=168

H_1: Media>168



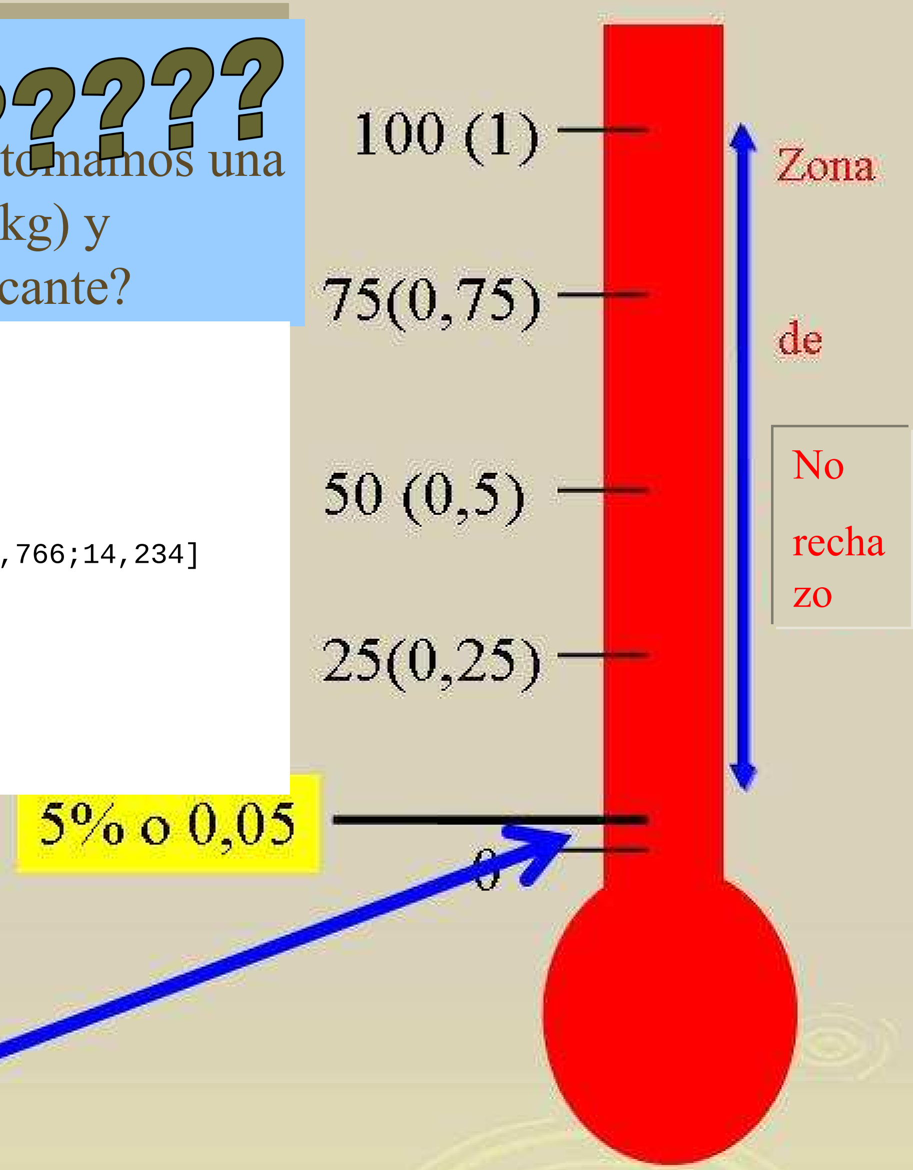

```
Hypothesis Tests
----------------
Sample mean = 173,0
Sample standard deviation = 10,0
Sample size = 2000

95,0% confidence interval for mean: 173,0 +/- 0,438262   [172,562;173,438]

Null Hypothesis: mean = 168,0
Alternative: not equal
Computed t statistic = 22,3607
P-Value = 0,0
Reject the null hypothesis for alpha = 0,05.
```

Problema:

Las bolsas de un fabricante s[illegible] 16Kg tomamos una mues[illegible] n=20 y a la vista de su media (14kg) y desviación (0.5Kg) ¿Debemos creer al fabricante?

Mejor o peor?????

```
Hypothesis Tests
----------------
Sample mean = 14,0
Sample standard deviation = 0,5
Sample size = 20

95,0% confidence interval for mean: 14,0 +/- 0,234008   [13,766;14,234]

Null Hypothesis: mean = 16,0
Alternative: not equal
Computed t statistic = -17,8885
P-Value = 2,39586E-13
Reject the null hypothesis for alpha = 0,05.
```



Zona de rechazo

H_0: Media=16

H_1: Media<16

Problema:

La duración de una enfermedad es de 15 días. Se estudia un nueva droga en 20 pacientes y se observa que, en promedio la duración son 13 días con desviación típica de 2 días.

¿Es mejor la nueva droga? ¿Es una casualidad?

```
Hypothesis Tests
----------------
Sample mean = 13,0
Sample standard deviation = 2,0
Sample size = 20,0

95,0% upper confidence bound for mean: 13,0 + 0,773293   [13,7733]

Null Hypothesis: mean = 15,0
Alternative: less than
Computed t statistic = -4,47214
P-Value = 0,000130597
Reject the null hypothesis for alpha = 0,05.
```

Mejor o peor?????

H_0: Media=15

H_1: Media<15

Problema:

La duración de una enfermedad es de 15 días. Se estudia un nueva droga en 20 pacientes y se observa que, en promedio la duración son 13 días con desviación típica de 2 días.

¿Es mejor la nueva droga? ¿Es una casualidad?

```
Hypothesis Tests
----------------
Sample mean = 13,0
Sample standard deviation = 2,0
Sample size = 20

95,0% upper confidence bound for mean: 13,0 + 0,773293   [13,7733]

Null Hypothesis: mean = 15,0
Alternative: less than
Computed t statistic = -4,47214
P-Value = 0,000130597
Reject the null hypothesis for alpha = 0,05.
```

Mejor o peor?????

H_0: Media=15

H_1: Media<15

Podemos contrastar

- Media de la muestra (Normal mean)
- Desviación de la muestra (Normal sigma)
- Proporción de defectuosos (Binomial proportion)
- Diferencia de medias
- Diferencia de proporciones

- **Siempre con la misma idea de p-valor**

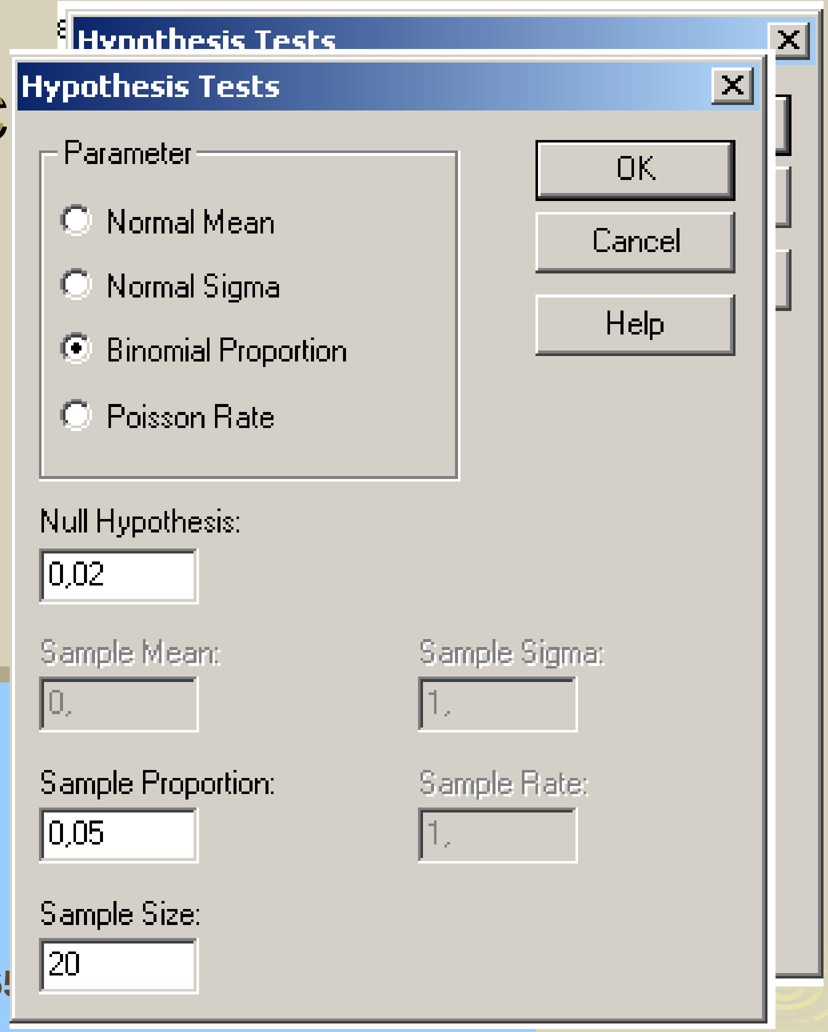
Proporción

- El proceso debe producir en buenas condiciones un 2% de defectuosos.
- ¿Cómo saber si funciona bien?
- Tomando muestras constantemente y contrastando si la proporción de verdad puede ser 2%.
- Tomamos una muestra:
  - n=20 obtenemos 1 defectuoso.

Proporc

H_0: prop= 0,02

H_1: prop distinta 0,02



Hypothesis Tests

Sample proportion = 0,05

Sample size = 20

Approximate 95,0% confidence interval for p: [0,00126

Null Hypothesis: proportion = 0,02

Alternative: not equal

P-Value = 0,664784

Do not reject the null hypothesis for alpha = 0,05.

Contraste para la diferencia de medias o proporciones

Problema:

Probamos dos marcas de neumáticos. El resultado es:

TIPO 1: Tamaño de la muestra: n_1

Media 1

Desviación típica 1

¿Son iguales o distintos?

TIPO 2: Tamaño de la muestra: n_2

Media 2

Desviación típica 2

H_0: media 1= media2 (Diferencia=0)

H_1: media 1 distinta de media 2

Compare-Two Samples-Hypoth

Problema:

Probamos dos marcas de neumáticos. El r

TIPO 1: Tamaño de la muestra: n_1=10

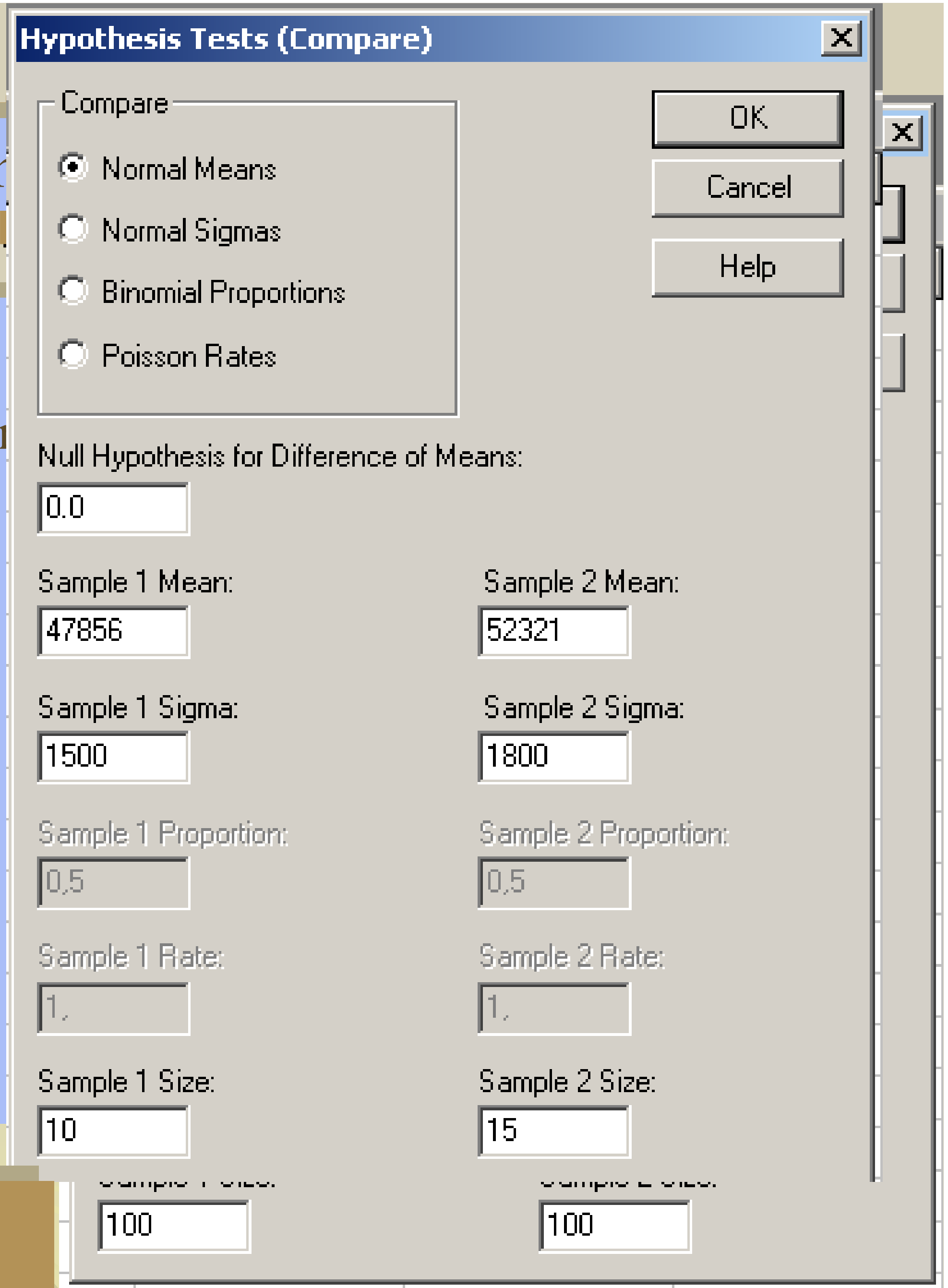
Media 1 =47.856Km

Desviación típica 1: 1500 Km

TIPO 2: Tamaño de la muestra: n_2 = 15

Media 2= 52.321

Desviación típica 2: 1800Km



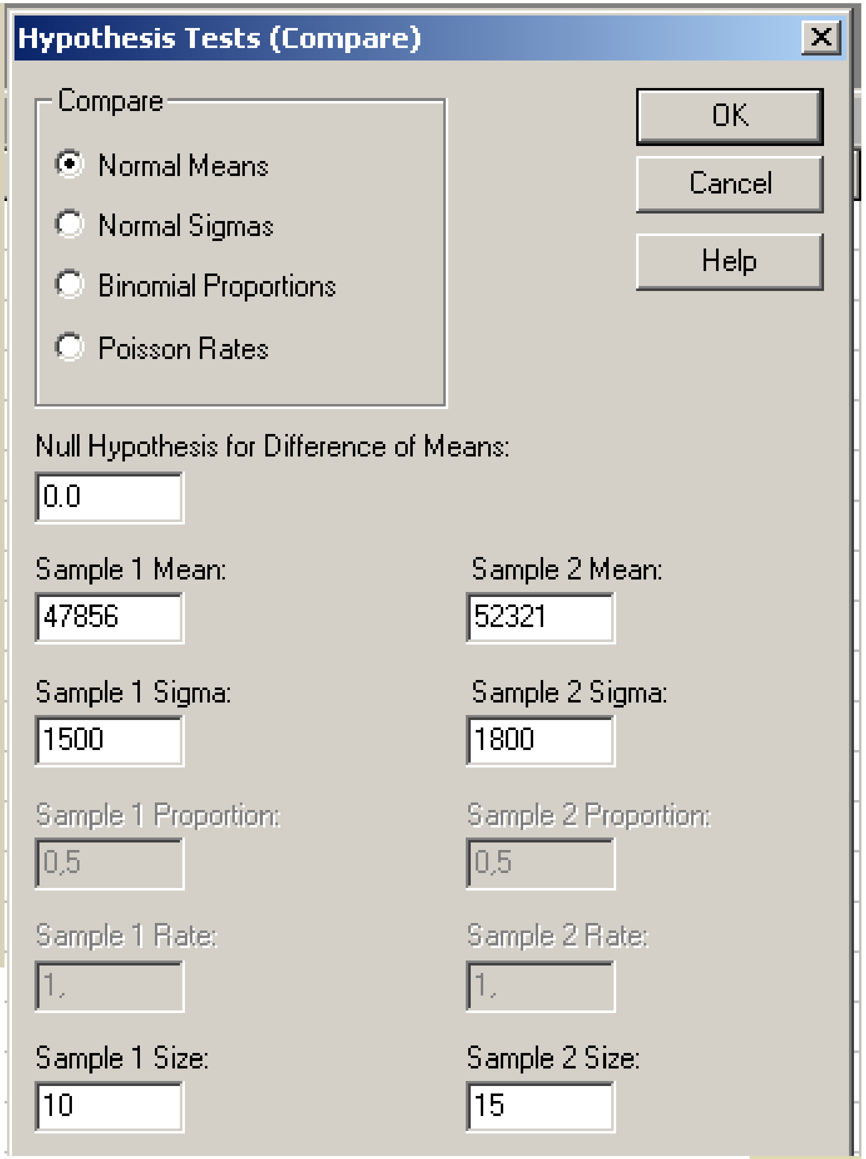
H_0: media 1= media2

H_1: media 1 distinta de media 2

RESULTADOS

H_0: media 1= media2

H_1: media 1 distinta de media 2




```
Hypothesis Tests
----------------
Sample means = 47856,0 and 52321,0
Sample standard deviations = 1500,0 and 1800,0
Sample sizes = 10 and 15

95,0% confidence interval for difference between means: -4465,0 +/- 1426,38   [-5891,38;-3038,62]

Null Hypothesis: difference between means = 0,0
Alternative: not equal
Computed t statistic = -6,47554
P-Value = 0,00000131452
Reject the null hypothesis for alpha = 0,05.
```

¿Son iguales o distintas?

Comparación de proporciones.

Problema:

Probamos dos máquinas y estudiamos el número de defectuosos. El resultado es:

MAQUINA 1: Tamaño de la muestra: n_1

Defectuosos: D_1

Proporción de defectuosos: $p_1=D_1/n_1$

MAQUINA 2: Tamaño de la muestra: n_2

Defectuosos: D_2

Proporción de defectuosos: $p_2=D_2/n_2$

H_0: $p_1=p_2$

H_1: p_1 distinta p_2

¿Son iguales o distintas?

Comparación de proporciones.

Problema:

Probamos dos máquinas y estudiamos el número de d

MAQUINA 1: Tamaño de la muestra: n_1

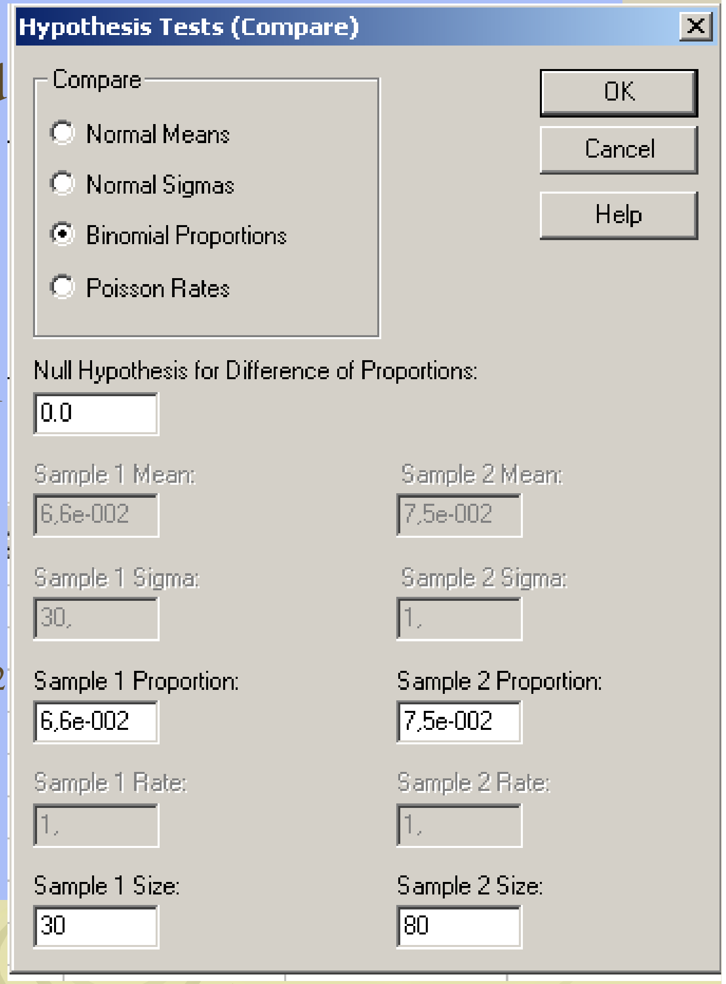
Defectuosos: D_1

Proporción de defectuosos: $p_1=D_1/n_1$

MAQUINA 2: Tamaño de la muestra: n_2

Defectuosos: D_2

Proporción de defectuosos: $p_2=D_2/n_2$



H_0: $p_1=p_2$

H_1: p_1 distinta p_2

ciones.

Hypothesis Tests

Sample proportions = 0,066 and 0,075

Sample sizes = 30 and 80

Approximate 95,0% confidence interval for difference between proportions: [-0,114947;0,0969469]

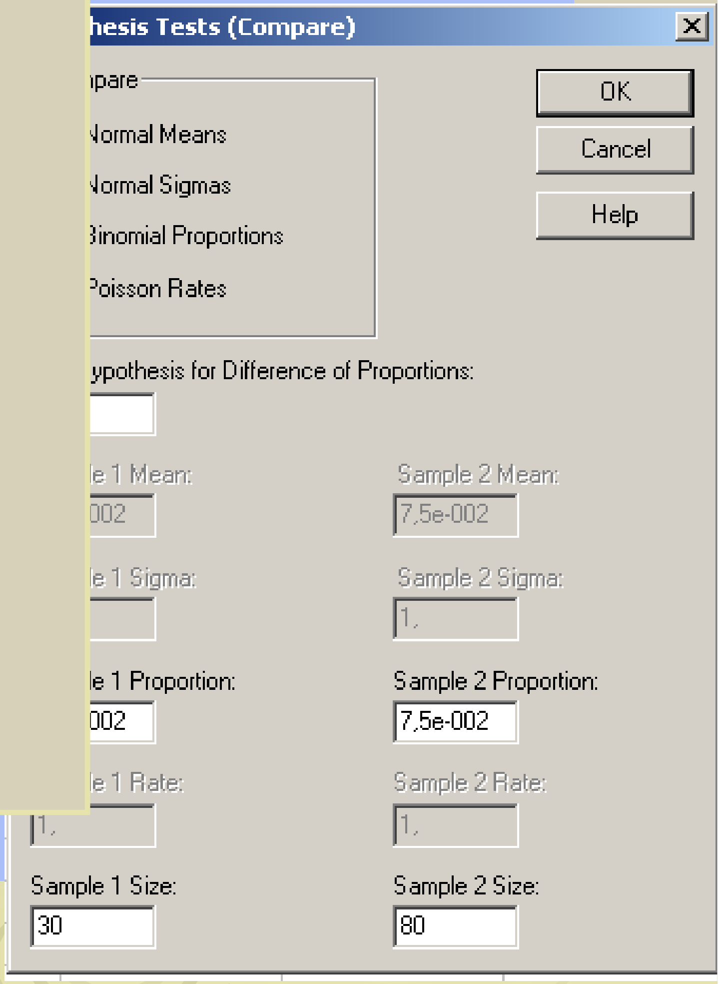
Null Hypothesis: difference between proportions = 0,0

Alternative: not equal

Computed z statistic = -0,162069

P-Value = 0,871247

Do not reject the null hypothesis for alpha = 0,05.



H_0: $p_1 = p_2$

H_1: p_1 distinta p_2

¿Para qué sirve la estadística?