

**ANEXO**

**TRATAMIENTO**

**ESTADÍSTICO DE DATOS**

**EXPERIMENTALES**



## 20. ANEXO. TRATAMIENTO ESTADÍSTICO DE DATOS EXPERIMENTALES.

### 20.1 Introducción.

La estadística es parte esencial en todas las ciencias y es la herramienta que nos permite tratar las incertidumbres inherentes a las medidas experimentales y obtener conclusiones a partir de los resultados. Permite diseñar y planificar los experimentos en vista a la precisión que se quiere alcanzar en los resultados.

El objetivo de todo experimento es medir magnitudes físicas. Por ejemplo: determinar el valor de un observable (determinación de un parámetro), probar la consistencia de una teoría con los datos obtenidos (test de una hipótesis).

En las ciencias cuantitativas las magnitudes vienen siempre determinadas por un número y una unidad. Pero las medidas que realicemos nunca nos proporcionarán el valor verdadero de la magnitud que queremos determinar. El resultado de una medida viene siempre afectado de un cierto grado de incertidumbre, lo que llamamos **error**. Algunas de las causas de estas incertidumbres son:

- imperfección de nuestros sentidos
- imperfección de los aparatos de medida
- imperfección del propio objeto
- carácter aleatorio de la naturaleza

Los errores pueden ser: **sistemáticos** y **aleatorios** o **accidentales**

### 20.2 Errores sistemáticos y accidentales.

Los **errores sistemáticos** son evitables o se pueden reducir si se identifican y eliminan las causas.

- Son debidos a un bias en el experimento, al método, a limitaciones de los instrumentos y a otras circunstancias que el físico tiene que evaluar, controlar y disminuir en lo posible. Estos errores afectan por igual a todas las medidas.
- No existen métodos ni procedimientos para calcularlos. Debe conocerse muy bien el detector y el aparato utilizado. Para estimarlos hay que recurrir a otro experimento independiente.

Los **errores aleatorios o accidentales** son incontrolables pero pueden ser tratados mediante técnicas estadísticas. Pueden deberse a incertidumbres experimentales, a la naturaleza estadística del fenómeno estudiado o a ambos. Podemos distinguir dos tipos:

- **Error instrumental:** Explica las fluctuaciones debidas a las imperfecciones de los aparatos o del observador. La precisión de la medida está limitada por el error instrumental. Suelen cumplir la ley de Gauss cuando se trata de magnitudes continuas.
- **Error estadístico:** es debido al hecho de utilizar una muestra finita para la determinación de los parámetros o porque el proceso sea intrínsecamente aleatorio. (Ej. Desintegraciones de un núcleo en un intervalo de tiempo).

En consecuencia,

- Medir es obtener una muestra finita de una población teórica infinita.
- La muestra permite estimar los verdaderos valores de los parámetros.
- La tarea del físico experimental es saber estimar y minimizar los errores aleatorios de su aparato y reducir los efectos de los sistemáticos al mínimo.
- La forma correcta de expresar una magnitud es:

$$\text{magnitud} = \text{valor medido} \pm \text{error accidental} \pm \text{error sistemático (unidades)}$$

### 20.3 Estimación del error de una magnitud.

Los errores aleatorios proceden de la acumulación de pequeñas alteraciones producidas al azar y que dan lugar a fluctuaciones de las medidas: lo que llamamos dispersión.

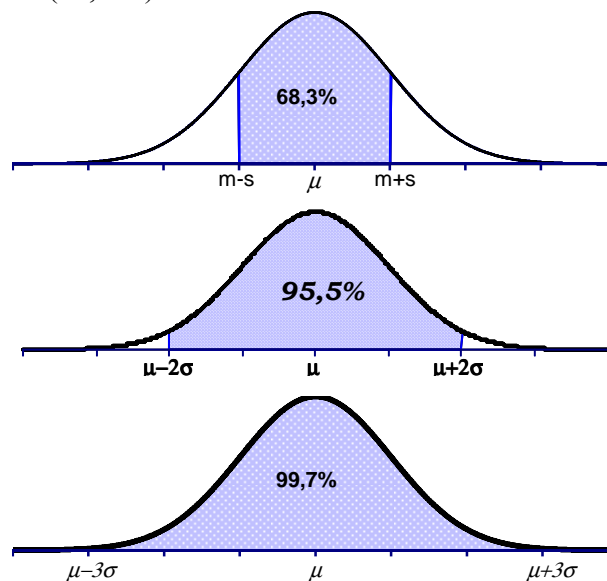
- Dadas  $n$  medidas  $x_i$  ( $i = 1, 2, \dots, n$ ) se pretende encontrar el valor  $m$  que haga mínima la dispersión de los datos, función que en principio denominamos  $u$  y que inmediatamente veremos que coincide con la varianza de los datos:

$$u = \frac{1}{n} \sum_{i=1}^n (x_i - m)^2 \rightarrow \frac{du}{dm} = 0 \rightarrow m = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}$$

El resultado anterior indica que el mejor valor de un conjunto de medidas dispersas es su valor medio, dado que es el que hace mínima la dispersión. Por otra parte, dicha dispersión viene dada por la varianza o desviación cuadrática media, siendo  $s$  la desviación típica o desviación estándar.

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

Si los errores son totalmente aleatorios, los datos siguen una distribución normal o de Gauss. El área sombreada es la correspondiente a los intervalos:  $\mu \pm \sigma$  (68,3%),  $\mu \pm 2\sigma$  (95,5%),  $\mu \pm 3\sigma$  (99,7%).



Si repitiéramos el experimento muchas veces obtendríamos diferentes estimaciones de  $\bar{x}$  y  $s$  (estimadores obtenidos a partir de la muestra).

- El estimador  $s$  es el error de dispersión de una medida individual, lo que significa que con un 68% de probabilidad una medida individual se encuentra en el intervalo:  $x_i \in [\bar{x} - s, \bar{x} + s]$  (análogamente para  $2s$  el 95.5% y para  $3s$  el 99.7%)
- Si el número de datos es pequeño, se suele definir la desviación típica como:

$$s_{n-1} = \left[ \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \right]^{1/2}$$

ya que  $n-1$  son las medidas independientes: **grados de libertad**, el punto restante se extrae de la definición del valor medio.

### 20.3.1 Propagación de desviaciones típicas.

Veamos cuál es el valor medio y la desviación típica de una magnitud  $q$  que depende de magnitudes medidas directamente.

- Supongamos que disponemos de dos conjuntos de datos  $x_i$ ,  $y_i$ , con valores medios y varianzas:

$$\bar{x} = \frac{1}{n} \sum_i x_i \quad ; \quad \bar{y} = \frac{1}{n} \sum_i y_i$$

$$s^2(x) = \frac{1}{n} \sum_i (x_i - \bar{x})^2 \quad ; \quad s^2(y) = \frac{1}{n} \sum_i (y_i - \bar{y})^2$$

y sea  $q(x, y)$  una función cualquiera. Desarrollando  $q_i = q(x_i, y_i)$  en serie de Taylor alrededor de  $\bar{x}$  e  $\bar{y}$ , hasta el primer término y suponiendo que las desviaciones  $(x_i - \bar{x})$ ,  $(y_i - \bar{y})$  son pequeñas:

$$q_i = q(x_i, y_i) = q(\bar{x}, \bar{y}) + \left. \frac{\partial q}{\partial x} \right|_{\bar{x}, \bar{y}} (x_i - \bar{x}) + \left. \frac{\partial q}{\partial y} \right|_{\bar{x}, \bar{y}} (y_i - \bar{y}) + \dots$$

$$\bar{q} = \frac{1}{n} \sum_i q(x_i, y_i) = \frac{1}{n} \sum_i q(\bar{x}, \bar{y}) + \frac{1}{n} \sum_i \left( \left. \frac{\partial q}{\partial x} \right|_{\bar{x}, \bar{y}} (x_i - \bar{x}) + \left. \frac{\partial q}{\partial y} \right|_{\bar{x}, \bar{y}} (y_i - \bar{y}) \right) + \dots = q(\bar{x}, \bar{y})$$

✓ El valor medio de una función es el valor de la función calculado para los valores medios de sus variables  $\rightarrow \bar{q} = q(\bar{x}, \bar{y})$

- Calculemos la varianza:

$$\begin{aligned} s^2(q) &= \frac{1}{n} \sum_i (q_i - \bar{q})^2 = \\ &= \left( \frac{\partial q}{\partial x} \right)^2 \frac{1}{n} \sum_i (x_i - \bar{x})^2 + \left( \frac{\partial q}{\partial y} \right)^2 \frac{1}{n} \sum_i (y_i - \bar{y})^2 + 2 \frac{\partial q}{\partial x} \frac{\partial q}{\partial y} \frac{1}{n} \sum_i (x_i - \bar{x})(y_i - \bar{y}) = \\ &= \left( \frac{\partial q}{\partial x} \right)^2 s^2(x) + \left( \frac{\partial q}{\partial y} \right)^2 s^2(y) + 2 \frac{\partial q}{\partial x} \frac{\partial q}{\partial y} s(x, y) \end{aligned}$$

donde  $s(x, y) = \frac{1}{n} \sum_i (x_i - \bar{x})(y_i - \bar{y})$  es la covarianza.

- La covarianza puede ser positiva, negativa o nula. Si las variables  $x$  e  $y$  son **independientes** la covarianza tiende a cero cuando el número de datos  $n$  crece.
- Supongamos que las variables son independientes y con covarianza nula, la fórmula de propagación se reduce a:

$$s^2(q) = \frac{1}{n} \sum_i (q_i - \bar{q})^2 = \left( \frac{\partial q}{\partial x} \right)^2 s^2(x) + \left( \frac{\partial q}{\partial y} \right)^2 s^2(y)$$

- Si las variables no son independientes la covarianza no es nula, por muchas medidas que realicemos, pero siempre se cumple la relación:  $|s(x,y)| \leq s(x)s(y)$ .

### 20.3.2 Desviación típica del valor medio.

Si un experimento se repite muchas veces se tendrán distintos valores de, dado que los valores  $x_i$  podrán ser distintos en las diferentes muestras. Sin embargo, todos los valores  $x_i$  vendrán caracterizados por la misma dispersión,  $s(x) \equiv s \equiv s_i$ , que medimos a partir de la desviación típica. A partir del resultado del apartado anterior se puede obtener la desviación típica,  $s(\bar{x})$ , del valor medio de un conjunto de datos.

$$s^2(\bar{x}) = \sum_i \left( \frac{\partial \bar{x}}{\partial x_i} s_i \right)^2 = \frac{1}{n^2} \sum_i (s_i)^2 = \frac{1}{n^2} (ns^2) = \frac{s^2(x)}{n} \quad \Rightarrow \quad \boxed{s(\bar{x}) = \frac{s(x)}{\sqrt{n}}}$$

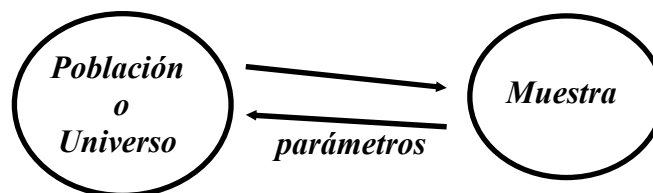
✓ La desviación típica de la media de  $n$  datos es la desviación típica de los datos dividida por la  $\sqrt{n}$ .

### 20.4 Función densidad de probabilidad.

En estadística no se puede predecir cuál será el resultado de un fenómeno aleatorio (una medida), pero es posible calcular la frecuencia de posibles resultados, lo que denominamos:

$f(x) \Rightarrow$  **función densidad de probabilidad (fdp)**

- La probabilidad de un suceso es la frecuencia relativa de aparición de éste, cuando el número de observaciones es muy grande.
- El objetivo de la estadística es inferir propiedades de la población analizando los datos de una muestra bien escogida y que se pueda considerar representativa de la población.



En la teoría de errores aleatorios la **población** está constituida por todas las medidas posibles de una o varias magnitudes, es por tanto infinita. La **muestra** es el conjunto de datos que proporcionan las medidas que realmente se han hecho. A partir de la muestra podemos obtener **parámetros** de toda la población mediante funciones de los datos que denominamos **estadísticos muestrales**.

- La **media** de un conjunto de datos y su **desviación típica** son ejemplos de estadísticos muestrales. Normalmente, se representa por  $\bar{x}$  y  $s$  a la media y

desviación típica de una muestra de datos. Cuando se trata de toda la población se denominan  $\mu$  y  $\sigma$ , respectivamente.

- Una característica de la población que puede variar al azar se denomina **variable estadística** o **aleatoria**, las variables pueden ser discretas o continuas.

#### 20.4.1 Probabilidad de variables aleatorias discretas y continuas.

Una **variable aleatoria  $x$  es discreta** cuando sus valores posibles son:  $x_1, x_2, \dots, x_n$  con probabilidades:  $f(x_1), f(x_2), \dots, f(x_n)$ . En tal caso, las **fdp** cumplen la condición de normalización:

$$\sum_{i=1}^n f(x_i) = 1 \rightarrow \text{La probabilidad total es 1}$$

Si la **variable aleatoria es continua** la probabilidad de encontrar uno de los infinitos valores posibles es nula. En cambio, es finita la probabilidad de encontrar para la variable un valor comprendido entre  $x$  y  $x + \Delta x$ , dado que entre  $x$  y  $x + \Delta x$  hay  $\infty$  valores. Si el intervalo  $\Delta x$  es tan pequeño que se puede sustituir por la diferencial,  $dx$ , la probabilidad de encontrar un valor entre  $x$  y  $x + \Delta x$  sólo será función de  $x$ , y podremos expresarla por  $f(x)dx$ .

Si  $f(x)$  es la densidad de probabilidad, la probabilidad de encontrar un valor entre  $x_1$  y  $x_2$  viene dada por:

$$\int_{x_1}^{x_2} f(x)dx \rightarrow \text{área de la función } f(x) \text{ en el intervalo } [x_1, x_2]$$

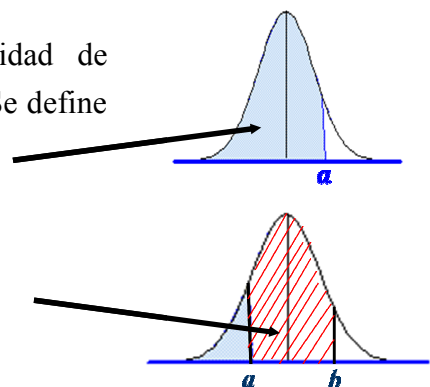
La **fdp** cumple la condición de normalización:

$$\int_{-\infty}^{\infty} f(x)dx = 1 \rightarrow \text{equivalente a } \sum_i f_i = 1$$

#### 20.4.2 Función acumulativa $F(a)$ .

La función  $F(a)$  se corresponde con la probabilidad de encontrar un valor para la variable en el intervalo  $[-\infty, a]$ . Se define como  $F(a) = \int_{-\infty}^a f(x)dx$  y da la probabilidad de que  $x \leq a$ .

- Es monótona creciente,  $0 \leq F(x) \leq 1$
- Verifica que  $\text{Prob}(a \leq x \leq b) = F(b) - F(a)$



#### 20.4.3 Valores esperados.

Se define el valor esperado de una función  $g(x)$  de la variable aleatoria  $x$  como:

$$E[g(x)] = \int_{-\infty}^{+\infty} g(x)f(x)dx$$

En particular, el valor esperado de la variable aleatoria  $x$  es  $E[x] = \mu$ , valor medio de  $x$  para la población o universo.

$$E[x] = \begin{cases} \sum_{i=1}^n x_i f(x_i) = \mu & \text{variables discretas } i = 1, 2, \dots, n \\ \int_{-\infty}^{+\infty} x f(x) dx = \mu & \text{variables continuas } x \in [-\infty, +\infty] \end{cases}$$

Además, se verifica que:  $E[ax + b] = a E[x] + b$

- Se mide la dispersión de  $x$  en la población por la varianza, la cual coincide con el valor esperado de  $(x - \mu)^2$ . La desviación típica  $\sigma$  es la raíz cuadrada positiva de la varianza.

$$\sigma^2 = E[(x - \mu)^2] = \begin{cases} \sum_{i=1}^n (x_i - \mu)^2 f(x_i) & \text{discretas} \\ \int_{-\infty}^{+\infty} (x - \mu)^2 f(x) dx & \text{continuas} \end{cases}$$

y dado que  $E[x] = \mu \rightarrow \sigma^2 = E[(x - \mu)^2] = E[x^2] + \mu^2 - 2\mu E[x] = \overline{x^2} - \mu^2$

#### 20.4.4 Propagación de errores cuadráticos

Consideremos dos variables aleatorias independientes  $x_1$  y  $x_2$  con densidades de probabilidad  $f_1(x_1)$  y  $f_2(x_2)$  y sea  $q(x_1, x_2)$  una función cualquiera de estas dos variables, supongamos que las dispersiones de ambas poblaciones son pequeñas. Haciendo un desarrollo en serie de Taylor:

$$q(x_1, x_2) = q(\mu_1, \mu_2) + \frac{\partial q}{\partial x_1}(x_1 - \mu_1) + \frac{\partial q}{\partial x_2}(x_2 - \mu_2) \quad ; \quad \begin{cases} \mu_1 = E[x_1] = \int_{-\infty}^{+\infty} x_1 f_1(x_1) dx_1 \\ \mu_2 = E[x_2] = \int_{-\infty}^{+\infty} x_2 f_2(x_2) dx_2 \end{cases}$$

- El valor esperado de  $q$ , o valor medio  $\bar{q}$  será:

$$\mu_q = \bar{q} = E[q] = \int_{-\infty}^{\infty} q(x_1, x_2) f_1(x_1) f_2(x_2) dx_1 dx_2 = q(\mu_1, \mu_2)$$

- La varianza  $\sigma^2(q)$  es el valor esperado de  $(q - \mu_q)^2$

$$\begin{aligned} \sigma^2(q) &= E[(q - \mu_q)^2] = \int (q(x_1, x_2) - q(\mu_1, \mu_2))^2 f_1(x_1) f_2(x_2) dx_1 dx_2 = \\ &= \int \left( \frac{\partial q}{\partial x_1}(x_1 - \mu_1) + \frac{\partial q}{\partial x_2}(x_2 - \mu_2) \right)^2 f_1(x_1) f_2(x_2) dx_1 dx_2 \end{aligned}$$

$$\boxed{\sigma^2(q) = \left( \frac{\partial q}{\partial x_1} \right)^2 \sigma^2(x_1) + \left( \frac{\partial q}{\partial x_2} \right)^2 \sigma^2(x_2)} \rightarrow \text{propagación de errores cuadráticos}$$

siendo:  $\sigma^2(x_1) = E[(x_1 - \mu_1)^2]$ ;  $\sigma^2(x_2) = E[(x_2 - \mu_2)^2]$  las varianzas de las variables.

Cuando nos referimos a variables aleatorias de una población, las covarianzas son totalmente nulas si las variables son independientes:  $E[(x_1 - \mu_1)(x_2 - \mu_2)] = 0$



Como aplicación de los resultados anteriores podemos calcular la varianza de la media de un conjunto de variables aleatorias  $x_i$  cuando todas ellas tienen la misma varianza  $\sigma^2(x)$ .

$$\bar{x} = \frac{1}{n} \sum x_i \Rightarrow \sigma^2(\bar{x}) = \sum \frac{1}{n^2} \sigma^2(x) = \frac{\sigma^2(x)}{n}$$

Cuando se trata de una muestra, la desviación típica de la media y por tanto su error se calcula como  $s(\bar{x}) = s(x)/\sqrt{n}$ .

#### 20.4.5 Matriz de covarianza.

Supongamos un proceso caracterizado por las variables aleatorias:  $x, y, z, \dots$ , el proceso vendrá descrito por la función de distribución multivariada  $f(x, y, z, \dots)$ . Por otra parte, además de los valores esperados para cada una de las variables, se puede definir la **covarianza** para cada par de variables:

$$\text{cov}(x, y) = E[(x - \mu_x)(y - \mu_y)]$$

estas cantidades son los elementos de la matriz de covarianzas y proporcionan una medida de la correlación entre dos variables.

Se define el coeficiente de correlación lineal como:

$$\rho_{xy} = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y} \text{ donde } -1 \leq \rho_{xy} \leq 1$$

indicando el signo el sentido de la correlación. Si la correlación es total:  $|\rho_{xy}| = 1$ , mientras que si  $\rho_{xy} = 0$  sólo podemos decir que las variables son linealmente independientes. (Ej.  $y = x^2$  con  $\rho_{xy} = 0$ , donde la correlación es total pero no lineal).

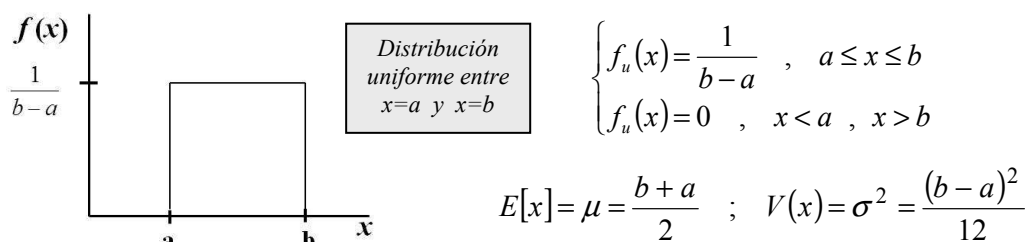
Dos variables aleatorias son independientes si  $f(x, y) = f_1(x) \cdot f_2(y)$  y  $\rho_{xy} = 0$

$$\text{Verificándose: } \begin{cases} E[f(x, y)] = E[f_1(x) \cdot f_2(y)] = E[f_1(x)] \cdot E[f_2(y)] \\ V(x + y) = V(x) + V(y) \end{cases}$$

### 20.5 Distribuciones de probabilidad.

A pesar de las numerosas distribuciones de probabilidad que se pueden definir, la mayoría de los procesos en física pueden ser descritos mediante un número reducido de distribuciones teóricas.

#### 20.5.1 Distribución uniforme.



### 20.5.2 Distribución binomial.

Estudia procesos cuyo resultado es dual: sí → éxito    no → fracaso

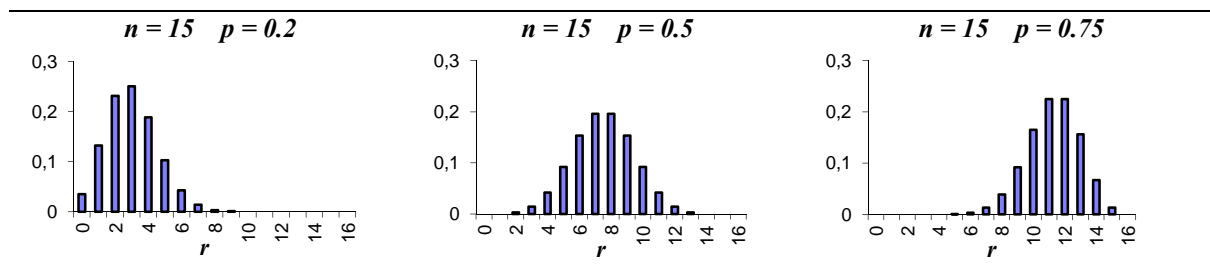
Si se realizan  $n$  pruebas independientes y  $p$  es la probabilidad de éxito ( $q=1-p$ , la de fracaso), la distribución **binomial** nos da la probabilidad de obtener  $r$  ( $r = 0, 1, 2, \dots, n$ ) éxitos. Se trata por tanto de variables aleatorias discretas. La función densidad de probabilidad:

$$f_b(r; n, p) = \frac{n!}{r!(n-r)!} p^r (1-p)^{n-r}$$

la media de la distribución es  $E[r] = \mu = np$ , para demostrarlo basta realizar los cambios:  $r' \rightarrow r-1$  y  $n' \rightarrow n-1$ , también se puede comprobar que la distribución está normalizada a 1 y que el valor de la varianza se calcula como:  $V(x) = \sigma^2 = np(1-p) = npq$ .

$$\mu = \sum_0^n r f_b = np \sum_0^n p^{r-1} (1-p)^{n-r} \frac{(n-1)!}{(r-1)!(n-r)!}$$

A continuación se dan algunos ejemplos para distintos valores de  $n$  y  $p$ ; como se observa, a mayor valor de  $p$  la distribución es más simétrica y tiende a la de Gauss.



Se puede afirmar que la distribución binomial tiende a las funciones de Gauss y Poisson en los siguientes casos:

si  $n \geq 30$  y  $p \geq 0.05$ ,  $f_b \rightarrow f_g(\mu, \sigma)$

si  $p \leq 0.05$  y  $np$  finito,  $f_b \rightarrow f_p(\mu)$

### 20.5.3 Distribución de Poisson.

Las observaciones científicas no se limitan a medir variables aleatorias continuas. A veces consiste en contar objetos o sucesos que tienen lugar en un intervalo espacial, temporal, etc., se trata de variables aleatorias discretas donde las probabilidades siguen la denominada distribución de Poisson. Ej.: *Desintegraciones de una sustancia radioactiva en un determinado intervalo de tiempo.*

- Sea  $t$  el intervalo de referencia donde se realizan las medidas (por ej.: un intervalo temporal) y  $\lambda$  la probabilidad de un suceso por unidad de tiempo, donde supondremos que  $\lambda$  es muy pequeña.
  - El promedio de sucesos en el intervalo  $t$  será:  $\mu = \lambda t$ .
  - La probabilidad de que tenga lugar un suceso en  $dt$  será:  $\lambda dt$  y supondremos nula la probabilidad de que se produzcan dos o más, por ser  $\lambda$  muy pequeña.
  - La probabilidad de no tener ningún suceso en  $dt$  será:  $(1 - \lambda dt)$ .
  - La probabilidad de que en  $t+dt$  ocurran  $x$  sucesos:

$$p_x(t+dt) = p_x(t)(1-\lambda dt) + p_{x-1}(t)\lambda dt$$

operando obtenemos:  $\frac{dp_x}{dt} = \lambda[p_{x-1} - p_x] \Rightarrow p_x = \frac{(\lambda t)^x}{x!} e^{-\lambda t}, (x=0,1,2,3,\dots)$

Expresión de la distribución de Poisson y que da la probabilidad de que se produzcan  $x$  sucesos en un intervalo  $t$ , donde el valor medio de sucesos esperados es  $E[x] = \mu = \lambda t$ .

### Características de la distribución de Poisson<sup>2</sup>:

- Sólo depende del valor medio  $\mu$ , único parámetro que define la distribución:

$$f_p(x; \mu) = \frac{\mu^x}{x!} e^{-\mu}, \quad x = 0, 1, 2, \dots$$

- Se verifica la condición de normalización:  $\sum_{x=0}^{\infty} f_p(x; \mu) = 1$

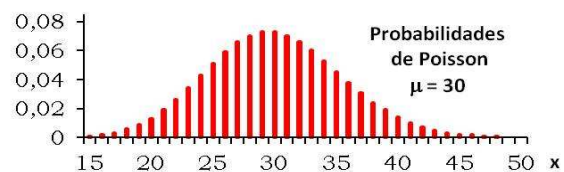
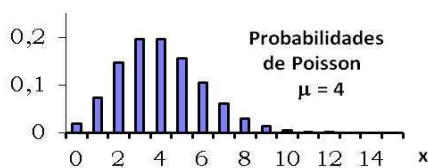
- El valor medio de  $x$ , único parámetro, es:  $E[x] = \bar{x} = \sum_{x=0}^{\infty} x f_p(x; \mu) = \mu$

- La varianza:  $V(x) = \sigma^2 = E[(x - \bar{x})^2] = E[x^2 + \bar{x}^2 - 2x\bar{x}] = \overline{x^2} - \bar{x}^2 = \overline{x^2} - \mu^2$ , donde

$$\overline{x^2} = \sum x^2 f_p(x; \mu) = \mu^2 + \mu, \text{ y por tanto } \boxed{V(x) = \sigma^2 = \mu \Rightarrow \sigma = \sqrt{\mu}}$$

Este resultado se refiere a un universo que satisfaga la distribución de Poisson y nos indica que el estimador que hemos de utilizar para la desviación típica de una muestra de  $x$  sucesos es:  $s = \sqrt{x}$ . Por tanto, si el resultado de las observaciones son  $x$  sucesos, éste vendrá expresado en la forma:  $x \pm \sqrt{x}$ .

- Se verifica que para  $p \rightarrow 0$ ,  $n \rightarrow \infty$  y  $np$  finito,  $f_p$  es el límite de la distribución binomial con  $p = \mu/n$ . Asimismo, cuando los números son grandes, la distribución de Poisson tiende a la de Gauss con  $\sigma = \sqrt{\mu}$ .



En el caso de procesos independientes, como por ejemplo señal y ruido en desintegraciones radioactivas, los procesos contribuyen a la medida con  $x_s$  y  $x_r$ , dando como resultado  $x = x_s + x_r$ , la media de la distribución es  $\mu = \mu_s + \mu_r$ , y la probabilidad total:

$$f_p(x; \mu) = \sum_{x_s=0}^x f_p(x_s; \mu_s) \cdot f_p(x - x_s; \mu_r)$$

<sup>2</sup> Para demostrar las siguientes relaciones hay que tener en cuenta que:  $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$

### 20.5.4 Distribución de Gauss o normal.

Cuando se efectúan medidas de una magnitud que viene afectada por variaciones aleatorias se puede demostrar estadísticamente, por el teorema del límite central, que las medidas se distribuyen de acuerdo con la denominada distribución de Gauss o normal.

El **teorema del límite central** establece: si  $x_i$  es un conjunto de  $n$  variables independientes de media  $\mu$  y varianza  $\sigma^2$ , entonces para  $n$  grande ( $n \rightarrow \infty$ ), la variable  $y = (1/n)\sum x_i$  tiende a una distribución de Gauss de media  $\mu$  y varianza  $\sigma^2/n$ .

Aún cuando las  $x_i$  procedan de distribuciones con diferentes  $\mu_i$  y  $\sigma_i^2$ , la variable  $y$  tiende a una distribución de Gauss, pero ahora con media  $\mu = (1/n)\sum \mu_i$  y varianza  $\sigma^2 = \sum \sigma_i^2$ . La única restricción importante es que la varianza  $\sigma^2$  sea una cantidad finita.

- La densidad de probabilidad *normal* o de Gauss es  $f_g(x; \mu, \sigma) = A e^{-(x-\mu)^2/2\sigma^2}$ ,  $\mu$  y  $\sigma$  son los dos parámetros de la distribución y  $A$  una constante de proporcionalidad, tal que:

$$\int_{-\infty}^{+\infty} f_g(x; \mu, \sigma) dx = 1 \Rightarrow A = \frac{1}{\sigma\sqrt{2\pi}}$$

por tanto la densidad de probabilidad normal:

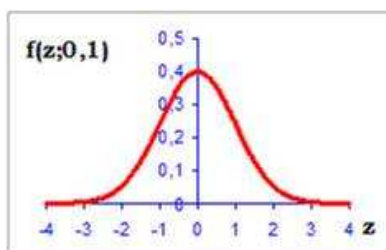
$$f_g(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$

▪ Un universo al cual corresponde esta densidad de probabilidad sigue una distribución  $N(\mu, \sigma)$ .

#### 20.5.4.1 La distribución normal tipificada.

La integral acumulativa de la distribución normal  $N(\mu, \sigma)$  no es analítica y para su cálculo se recurre a tablas de la denominada **función reducida** o **tipificada**  $N(0,1)$ .

- Dada una distribución  $N(\mu, \sigma)$  se define la variable tipificada:  $z = \frac{x-\mu}{\sigma} \rightarrow dz = \frac{dx}{\sigma}$
- La densidad de probabilidad:  $f_g(z; 0,1) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$  corresponde a la distribución de Gauss tipificada  $N(0,1)$ , independiente de  $\mu$  y  $\sigma$ . Se verifica que:



- Máximo:  $\frac{df}{dz} = 0 \Rightarrow z = 0$ ,  $f(0) = 0.3989$
- Puntos de inflexión:  $\frac{d^2f}{dz^2} = 0 \Rightarrow z = \pm 1$
- Con el cambio de variable:  $x = \mu + \sigma z$  se pasa a la distribución  $N(\mu, \sigma)$ 
  - ⇒ Máximo:  $x = \mu$
  - ⇒ Puntos de inflexión:  $x = \mu \pm \sigma$

- Las áreas para distintos intervalos de  $\sigma$  son:

$$\mu \pm \sigma \rightarrow 68.3\% \quad \mu \pm 2\sigma \rightarrow 95.5\% \quad \mu \pm 3\sigma \rightarrow 99.7\%$$

Por tanto, cualquier integral de una distribución de Gauss,  $N(\mu, \sigma)$ , se puede reducir a una integral de la función tipificada  $N(0,1)$ . Asimismo, a partir de los valores esperados de  $z$  y  $z^2$  se pueden obtener el valor medio y la varianza de la variable no tipificada  $x$ :

$$E[z] = \int_{-\infty}^{+\infty} z f_g(z; 0, 1) dz = 0 \rightarrow E[z] = E\left[\frac{x - \mu}{\sigma}\right] = \frac{1}{\sigma}(E[x] - \mu) = 0 \rightarrow E[z] = \mu$$

$$E[z^2] = \int_{-\infty}^{+\infty} z^2 f_g(z; 0, 1) dz = 1 \rightarrow E[z^2] = \frac{1}{\sigma^2} E[(x - \mu)^2] = 1 \rightarrow E[(x - \mu)^2] = E[x^2] - \mu^2 = \sigma^2$$

En los resultados anteriores,  $\mu$  es el valor medio de la variable aleatoria de un universo normal,  $\sigma^2$  la varianza y  $\sigma$  la desviación típica.

#### 20.5.4.2 Intervalos de probabilidad.

Un problema muy común es el de determinar la probabilidad de que la variable aleatoria continua  $x$  tome valores en el intervalo  $[x_1, x_2]$ , es decir, calcular la integral:

$$\int_{x_1}^{x_2} f(x) dx \quad \blacksquare \text{ Cuando se trata de la distribución de Gauss, la función no tiene integral analítica y hay que recurrir al cálculo numérico o a las tablas de la variable tipificada } z.$$

Cualquier integral definida de la función  $f_g(x; \mu, \sigma)$  se puede evaluar en términos de la función  $f_g(z; 0, 1)$  realizando el cambio de variable  $x \rightarrow z$ , es decir pasando a la variable tipificada.

$$z = \frac{x - \mu}{\sigma} \rightarrow \int_{x_1}^{x_2} \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} dx = \int_{z_1}^{z_2} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$$

Otra posibilidad para poder evaluar la integral, y por tanto la probabilidad, es a partir de la **función error**, definida como:

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \approx \frac{2}{\sqrt{\pi}} \left( x - \frac{x^3}{3} + \frac{1}{2!} \frac{x^5}{5} - \frac{1}{3!} \frac{x^7}{7} + \dots \right) ; \quad \operatorname{erf}(x) = -\operatorname{erf}(-x)$$

mediante un cambio de variable se puede demostrar que la relación entre dicha función y la distribución normal es:

$$\int_0^x f_g(z) dz = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-z^2/2} dz = \frac{1}{2} \operatorname{erf}\left(\frac{x}{\sqrt{2}}\right)$$

además, teniendo en cuenta que la distribución normal es simétrica y que el área total es 1, se puede evaluar la integral para cualquier valor de los límites.

La integral de la distribución  $N(\mu, \sigma)$  en el intervalo  $[x_1, x_2]$  se corresponde con la probabilidad de que al efectuar una observación sobre una población gobernada por dicha distribución, el valor  $x$  obtenido esté comprendido en el intervalo  $x_1 < x < x_2$ , lo que también se puede expresar como:  $\mu + z_1\sigma < x < \mu + z_2\sigma$ . Si el intervalo es simétrico alrededor del valor medio,  $(x_2 - \mu) = -(x_1 - \mu) \rightarrow -z_1 = z_2 = z_c$ , podemos construir la siguiente tabla, que corresponde a las probabilidades **P** de que  $\mu - z_c\sigma < x < \mu + z_c\sigma$ .

Probabilidad  $P$  de que  $x \in [\mu - z_c \sigma, \mu + z_c \sigma]$  para distintos valores de  $z_c$ .

$z_c$	$P$
0,2	0,158
0,4	0,331
0,6	0,451
0,8	0,576
1,0	0,683

$z_c$	$P$
1,2	0,770
1,4	0,838
1,6	0,890
1,8	0,928
2,0	0,954

$z_c$	$P$
2,2	0,972
2,4	0,984
2,6	0,991
2,8	0,995
3,0	0,997

La última fila corresponde a los casos:  $\mu \pm \sigma$  (68.3%),  $\mu \pm 2\sigma$  (95.5%),  $\mu \pm 3\sigma$  (99,7%). Al final de este Anexo se proporcionan tablas completas de las ordenadas y las áreas de la curva normal tipificada.

#### Puntos a considerar:

- El error que se asocia a una medida individual es la desviación típica de la muestra:  $x \pm s$
- Al valor medio de la muestra se le asocia la desviación típica de la media:  $\bar{x} \pm s/\sqrt{n-1}$ , el significado estadístico de este criterio es: **con una probabilidad del 68%, la observación  $x_i$  se encuentra en el intervalo:  $\mu - \sigma < x < \mu + \sigma$ .**
- Se puede afirmar que dada una medida  $x_i$ , el valor verdadero  $\mu$  de la población se encuentra en el intervalo  $x_i - \sigma < \mu < x_i + \sigma$ , con una probabilidad del 68%. Lo mismo se verifica para el valor medio, pero ahora su desviación típica es  $\sigma(\bar{x}) = \sigma(x)/\sqrt{n-1}$ .

#### 20.5.4.3 Niveles de confianza y significación.

La integral de la distribución normal permite calcular la probabilidad de que el valor  $\mu$  de la población se encuentre en intervalo  $x - z_c \sigma < \mu < x + z_c \sigma$ .

Probabilidad en % de que  $\mu \in [x - z_c \sigma, x + z_c \sigma]$  para distintos valores de  $z_c$ .

$P$ (%)	$z_c$	$P$ (%)	$z_c$	$P$ (%)	$z_c$
99,9	3,29	95	1,96	70	1,04
99,5	2,81	90	1,64	60	0,84
99	2,58	80	1,28	50	0,674

La probabilidad dependerá de la anchura del intervalo, es decir del valor crítico  $z_c$ . A mayor  $z_c$  mayor probabilidad, pero afinamos menos en el valor de la magnitud. La probabilidad será la relación del área de la gaussiana entre los valores críticos y el área total, que vale 1.

Al valor de  $P$  se le denomina “**nivel de confianza**”.  
 $Q = 1 - P$  es el “**nivel de significación**”

No tiene sentido hablar de “un nivel de confianza del 100 %”, pues esto significaría que el valor de la magnitud se encuentra ente  $+\infty$  y  $-\infty$ .

#### 20.5.4.4 Diferencias significativas.

Supongamos que se han determinado los valores medios  $\bar{x}$ ,  $\bar{y}$  de dos conjuntos de medidas independientes de la misma magnitud, se trata de decidir si las medidas son compatibles o si la diferencia  $\bar{x} - \bar{y}$  es significativa,

- Si se verifica:  $|\bar{x} - \bar{y}| \leq z_c \sigma(\bar{x} - \bar{y})$ , la diferencia no es significativa con un nivel de confianza  $P$  o al nivel de significación  $Q$ , definidos a partir del valor crítico  $z_c$ , donde:

$$\sigma^2(\bar{x} - \bar{y}) = \sigma^2(\bar{x}) + \sigma^2(\bar{y}) = \sigma^2(x)/n_x + \sigma^2(y)/n_y$$

#### 20.5.5 Distribución de $\chi^2$ .

Dadas  $n$  variables aleatorias e independientes  $x_1, x_2, \dots, x_n$ , todas ellas pertenecientes a distribuciones normales de medias  $\mu_1, \mu_2, \dots, \mu_n$  y desviaciones típicas  $\sigma_1, \sigma_2, \dots, \sigma_n$ , se define la variable estadística  $\chi^2$ , con  $v = n$  grados de libertad como:

$$\chi^2 = \sum_{i=1}^n \frac{(x_i - \mu_i)^2}{\sigma_i^2}$$

El objetivo principal es determinar la función densidad de probabilidad  $f(\chi^2, n)$ , es decir, la función tal que  $f(\chi^2, n)d\chi^2$  sea la probabilidad de que  $\chi^2$  se encuentre entre  $\chi^2$  y  $\chi^2 + d\chi^2$ . La probabilidad conjunta de que las variables  $x_i$  tomen valores entre  $x_i$  y  $x_i + dx_i$  es:

$$\begin{aligned} f_g(x_1, \dots, x_i, \dots, x_n) dx_1 \dots dx_i \dots dx_n = \\ = f_g(x_1; \mu_1, \sigma_1) \dots f_g(x_i; \mu_i, \sigma_i) \dots f_g(x_n; \mu_n, \sigma_n) dx_1 \dots dx_i \dots dx_n \end{aligned}$$

$$f_g(x_1, \dots, x_i, \dots, x_n) = \frac{1}{\sigma_1 \sigma_2 \dots \sigma_n (2\pi)^{n/2}} e^{-\chi^2/2}$$

pasando a variables tipificadas,  $z_i = \frac{x_i - \mu_i}{\sigma_i}$ ,  $dz_i = \frac{dx_i}{\sigma_i}$ , la densidad de probabilidad queda:

$$q(z_1, \dots, z_i, \dots, z_n) = \frac{1}{(2\pi)^{n/2}} e^{-\chi^2/2} \rightarrow \boxed{\chi^2 = \sum_{i=1}^n z_i^2}$$

#### Consideraciones:

- $\chi^2$  es el cuadrado de la distancia entre el origen y el punto  $(z_1, \dots, z_i, \dots, z_n)$ , en un espacio de  $n$  dimensiones. La probabilidad de que  $\chi^2$  se encuentre entre  $\chi^2$  y  $\chi^2 + d\chi^2$  es  $f(\chi^2, n) \propto e^{-\chi^2/2} \chi^{n-1} d\chi$ . Donde  $e^{-\chi^2/2}$  es la densidad de probabilidad y  $\chi^{n-1} d\chi$  es proporcional al elemento de volumen o área, según la dimensión del espacio

$$\chi^{n-1} d\chi = \chi^{n-2} \chi d\chi = (\chi^2)^{\frac{n}{2}-1} d\chi^2 / 2$$

con lo cual,  $f(\chi^2, n)d\chi^2 = Ke^{-\chi^2/2}(\chi^2)^{\frac{n}{2}-1}d\chi^2$ , donde  $K$  es una constante para que se verifique:  $\int_0^\infty f(\chi^2, n)d\chi^2 = 1$ .

- Con el cambio de variable  $y = \chi^2/2 \rightarrow \frac{1}{K} = 2^{n/2} \int_0^\infty e^{-y} y^{(n/2-1)} dy = 2^{n/2} \Gamma(n/2)$ , donde

$\Gamma(\xi) = \int_0^\infty e^{-u} u^{\xi-1} du$  es la función **Gamma de Euler** definida como:

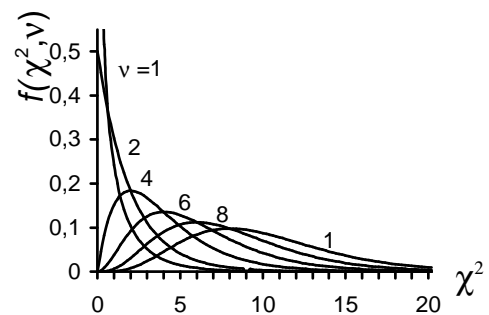
$$\Gamma(\xi) = (\xi-1)\Gamma(\xi-1), \quad \Gamma(1/2) = \sqrt{\pi}, \quad \Gamma(1) = 1, \quad \begin{cases} \Gamma(n) = (n-1)! & \text{para } n \text{ par} \\ \Gamma(n/2) = \sqrt{\pi} (1/2)(3/2)\cdots(n/2-1) & n \text{ impar} \end{cases}$$

La distribución  $\chi^2$  depende del parámetro  $n$ , número de variables aleatorias, si los parámetros poblacionales  $\mu_i$ ,  $\sigma_i$  son conocidos. Si alguno de los parámetros se desconoce, éste se ha de determinar a partir de la muestra.

Si  $k$  es el número de parámetros estimados, el estadístico  $\chi^2$  tiene  $\nu = n - k$  grados de libertad, cada parámetro que se estima es una ligadura entre los datos de la muestra y sólo hay  $n - k$  parámetros independientes, con lo cual la distribución de  $\chi^2$ :

$$f(\chi^2, \nu) = \frac{e^{-\chi^2/2} (\chi^2)^{\nu/2-1}}{2^{\nu/2} \Gamma(\nu/2)}$$

- Grados de libertad:  $\nu = n - k$
- Máximo:  $\chi^2 = \nu - 2$
- Valor medio:  $E[\chi^2] = \nu$
- Desviación típica:  $\sigma(\chi^2) = \sqrt{2\nu}$



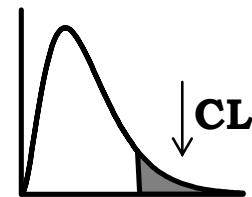
Conocida la densidad de probabilidad del estadístico  $\chi^2$ , podemos establecer criterios para juzgar la precisión de un ajuste utilizando el **nivel de confianza**. Téngase en cuenta que:

- La variable  $\chi^2$  caracteriza las fluctuaciones de  $x_i$ . Si las fluctuaciones son gaussianas se espera que  $\overline{\chi^2} = \nu$  y para cada ajuste el valor de  $\chi^2$  fluctuará según  $f(\chi^2, \nu)$ .
- Dado un valor fijo  $\chi_o^2$ , podemos calcular la probabilidad de que  $\chi^2$  sea inferior o superior a este valor utilizando la función acumulativa  $F(\chi_o^2)$ .

$$F(\chi_o^2) = P(\chi^2 < \chi_o^2) = \int_0^{\chi_o^2} f(\chi^2, \nu) d\chi^2$$

Se define el **nivel de confianza (CL)** como<sup>3</sup>:

$$CL(\chi_o^2) = 1 - F(\chi_o^2) = \int_{\chi_o^2}^\infty f(\chi^2, \nu) d\chi^2$$



<sup>3</sup> Nótese que el significado, y por tanto la definición que aquí se hace, de *nivel de confianza* no se corresponde con el que previamente se introdujo para el caso de la distribución de Gauss.



CL representa la probabilidad de obtener un valor  $\chi^2 \geq \chi_o^2$ , es decir la probabilidad de obtener un resultado peor, caso de repetir el experimento. Consideremos, por ejemplo, un valor  $\chi_o^2$  correspondiente a  $P = 0.95$ ,  $CL = 5\%$ .

- Si el  $\chi^2$  observado está entre 0 y  $\chi_o^2$  diremos que el ajuste es bueno a un nivel de confianza del 5%.
- Si está entre  $\chi_o^2$  e  $\infty$  diremos que no lo es, ya que la probabilidad de encontrar un resultado peor es pequeña.

Para saber si un ajuste es razonable, basta con ver si el  $\chi^2$  reducido:  $\chi^2/\nu \approx 1$ . Para que el test sea válido los errores han de estar bien estimados. Si los errores están infravalorados se observará un aumento del  $\chi^2$  a pesar del aparente acuerdo entre los datos y la función. Por el contrario si los errores se han tomado por exceso se obtendrán valores de  $\chi^2$  demasiado pequeños. **En términos de probabilidad, se admite que un buen ajuste es aquél que da una probabilidad de  $\chi^2 \geq 5\%$ .** Al final de este Anexo se proporcionan tablas y gráficas que permiten obtener los niveles de confianza para distintos grados de libertad.

## 20.6 Estimación de parámetros.

### 20.6.1 Método de máxima verosimilitud.

La finalidad del método es determinar parámetros de una población a partir de los datos de una muestra. El método sólo es aplicable si se conoce la forma teórica de la distribución a la cual pertenece la muestra. Si por el contrario se desconoce la forma de la distribución habrá que hacer diferentes hipótesis y ver cual es la función que proporciona mejores resultados.

**Hipótesis:** El mejor estimador de un parámetro es aquél para el cual es máxima la probabilidad de encontrar la muestra que de hecho se ha encontrado.

**Aplicación del método:** Sea una muestra finita de  $n$  medidas  $\{x_i, i=1, \dots, n\}$  y  $f(x|\theta)$  la *fdp* teórica que describe los datos, es decir  $f(x_i|\theta)$  es la probabilidad correspondiente al valor  $x_i$ . El mejor estimador  $\hat{\theta}$  del parámetro  $\theta$ , se obtiene haciendo máxima la denominada función de verosimilitud. Si las  $x_i$  son independientes, entonces la función que describe la probabilidad conjunta de encontrar los  $n$  valores de la muestra se escribe como:

$$\mathcal{L}(x|\theta) = \prod_{i=1}^n f(x_i|\theta)$$

**Ejemplo:** Se quiere calcular el parámetro  $\tau$  que corresponde a la vida media de una partícula que se desintegra. Se han hecho  $n$  medidas, encontrando los valores  $t_i$  ( $i=1, 2, \dots, n$ ). Conocidas las probabilidades de que aparezca cada uno de los valores  $t_i$  podemos estimar el parámetro  $\tau$  haciendo máxima la probabilidad total.

$$f(t|\tau) \rightarrow f(t_i|\tau) = \frac{1}{\tau} \exp(-t_i/\tau) \rightarrow \prod_{i=1}^n f(t_i|\tau) \rightarrow \text{máxima}$$

- El mejor estimador del parámetro se obtiene derivando la función  $\mathcal{L}(x|\theta)$ , o lo que es equivalente su logaritmo, e igualando a cero  $\rightarrow \boxed{\frac{\partial (\ln \mathcal{L})}{\partial \theta} \Big|_{\hat{\theta}} = 0}$
- El error se obtiene a partir de la varianza del estimador:  $\sigma^2(\hat{\theta}) = \int (\theta - \hat{\theta})^2 \mathcal{L}(x|\theta) \prod_{i=1}^n dx_i$ , lo cual no es fácil de calcular, pero en el límite  $n \rightarrow \infty$  en el que  $\mathcal{L} \rightarrow f_g$  se cumple  $\sigma^2(\hat{\theta}) = -\left(\frac{\partial^2 (\ln \mathcal{L})}{\partial \theta^2}\right)^{-1}$ ; fácil de comprobar si  $\mathcal{L}(\theta) \approx e^{-(\theta - \theta^*)^2 / 2\sigma^2}$ .
- Si la función es de varios parámetros se debe calcular la matriz  $V_{ij} = -\frac{\partial^2 (\ln \mathcal{L})}{\partial \theta_i \partial \theta_j}$ , en este caso las varianzas son los elementos de la diagonal de la matriz inversa  $\sigma^2(\hat{\theta}_i) = (V^{-1})_{ii}$ .

### 20.6.1.1 Ejemplo 1: Estimador de la distribución de Poisson

Vamos a aplicar el principio de máxima verosimilitud para encontrar la mejor estimación del parámetro  $\mu$  a partir de un conjunto de datos  $\{x_1, \dots, x_n\}$  que proceden de una distribución de Poisson.

- La función de máxima verosimilitud será:

$$\mathcal{L}(\mu|x) = \prod_{i=1}^n \frac{\mu^{x_i} e^{-\mu}}{x_i!} = e^{-n\mu} \prod_{i=1}^n \frac{\mu^{x_i}}{x_i!}$$

$$\ln \mathcal{L} = -n\mu + \sum_{i=1}^n x_i \ln \mu - \sum_{i=1}^n \ln x_i!$$

- Diferenciando e igualando a cero:

$$\frac{d\mathcal{L}^*}{d\mu} = \frac{d(\ln \mathcal{L})}{d\mu} = -n + \frac{1}{\mu} \sum_{i=1}^n x_i = 0 \rightarrow \boxed{\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}}$$

- Calcularemos la varianza utilizando el resultado anterior:

$$\sigma^2(\hat{\mu}) = \sigma^2(\bar{x}) = E[(\bar{x} - \mu)^2] = E\left[\left(\frac{1}{n} \sum x_i - \mu\right)^2\right] =$$

$$\frac{1}{n^2} E\left[\left(\sum x_i - n\mu\right)^2\right] = \frac{1}{n^2} E\left[\left\{\sum (x_i - \mu)\right\}^2\right]$$

desarrollando el cuadrado de la suma:

$$\left\{\sum_i (x_i - \mu)\right\}^2 = \sum_i (x_i - \mu)^2 + \sum_i \sum_{j \neq i} (x_i - \mu)(x_j - \mu)$$

al tomar valores esperados cancelan los términos cruzados, quedando finalmente:

$$\sigma^2(\bar{x}) = \sigma^2(\hat{\mu}) = \frac{1}{n^2} E\left[\sum_i (x_i - \mu)^2\right] = \frac{1}{n^2} \sum_i E[(x_i - \mu)^2] = \frac{\sigma^2}{n}$$

- En particular para la distribución de Poisson, donde  $\sigma^2 = \mu$ , el error del parámetro estimado  $\hat{\mu}$  es:  $\sigma(\hat{\mu}) = \sqrt{\frac{\mu}{n}} \cong \sqrt{\frac{\hat{\mu}}{n}} = \sqrt{\frac{\bar{x}}{n}}$

### 20.6.1.2 Ejemplo 2: Estimadores de la distribución de Gauss

- Para un conjunto de  $n$  puntos, todos pertenecientes a la misma distribución de Gauss, la función de verosimilitud:

$$\mathcal{L}(\mu, \sigma^2 | x) = \prod_{i=1}^n \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{(x_i - \mu)^2}{2\sigma^2}\right)$$

$$\ln \mathcal{L} = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2} \sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma}\right)^2$$

- Derivando respecto de  $\mu$ :

$$\frac{\partial(\ln \mathcal{L})}{\partial \mu} = \sum \frac{x_i - \mu}{\sigma^2} = 0 \rightarrow \hat{\mu} = \frac{1}{n} \sum x_i = \bar{x}$$

y por tanto,  $\sigma^2(\bar{x}) = \sigma^2(\hat{\mu}) = \frac{\sigma^2}{n} \rightarrow \sigma(\hat{\mu}) = \frac{\sigma}{\sqrt{n}}$

- El error de la media disminuye con  $n$ .
- La  $\sigma$  de la población debe ser conocida

- Derivando respecto de  $\sigma^2$ :

$$\frac{\partial(\ln \mathcal{L})}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2} \sum \left(\frac{x_i - \mu}{\sigma^2}\right)^2 \frac{1}{\sigma^2} = 0 \rightarrow \hat{\sigma}^2 = \frac{1}{n} \sum (x_i - \mu)^2 \approx \frac{1}{n} \sum (x_i - \bar{x})^2 = s^2$$

El estimador sólo puede calcularse si se conoce  $\mu$ . Además, para  $n$  finito su valor estimado está sesgado. Se puede ver que:

$$E[s^2] = E\left[\frac{1}{n} \sum (x_i - \bar{x})^2\right] = \sigma^2 - \sigma^2(\bar{x}) = \sigma^2 - \frac{\sigma^2}{n} = \frac{n-1}{n} \sigma^2$$

y por tanto, la expresión correcta y no sesgada del estimador de la varianza de una distribución de medidas gaussianas es

$$\hat{\sigma}^2 = s_{n-1}^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2$$

ya que su valor esperado coincide con  $\sigma^2$ .

- Asimismo se demuestra que:

$$\sigma^2(\hat{\sigma}^2) = \frac{2(n-1)\hat{\sigma}^4}{n^2} \rightarrow \sigma^2(\hat{\sigma}^2) \approx \frac{2\hat{\sigma}^4}{n} ; (n \rightarrow \infty) \rightarrow \sigma(\hat{\sigma}) = \hat{\sigma}/\sqrt{2(n-1)}$$

### 20.6.2 Media ponderada.

En el apartado anterior hemos considerado que los  $n$  valores de  $x_i$  pertenecían a la misma distribución  $N(\mu, \sigma)$ . Vamos ahora a tratar el problema de combinar medidas de la misma magnitud, pero de precisiones distintas, es decir cada una de ellas procedente de una distribución  $N(\mu, \sigma_i)$  y obtener su media ponderada.

- El mejor estimador de  $\mu$  será aquel que haga máxima la función de verosimilitud, o lo que es lo mismo, mínimo el *chi-cuadrado*

$$\chi^2 = \sum_{i=1}^n \left( \frac{x_i - \mu}{\sigma_i} \right)^2 \rightarrow \frac{d\chi^2}{d\mu} = 0 \rightarrow \boxed{\hat{\mu} = \frac{\sum_i x_i / \sigma_i^2}{\sum_i (1/\sigma_i^2)}}$$

La expresión anterior es la media ponderada de  $n$  valores  $x_i$  con pesos  $w_i = 1/\sigma_i^2$ , los cuales son proporcionales a las inversas de las varianzas.

- La desviación típica de la población combinada se obtiene por la regla de propagación de varianzas

$$\sigma^2(\hat{\mu}) = \sum_i \left( \frac{\partial \hat{\mu}}{\partial x_i} \sigma_i \right)^2 = \frac{1}{\sum_i 1/\sigma_i^2} \rightarrow \boxed{\sigma(\hat{\mu}) = \sqrt{\frac{1}{\sum_i 1/\sigma_i^2}} = \sqrt{\frac{1}{\sum_i w_i}}}$$

### 20.6.3 Ejemplos y aplicaciones

#### 20.6.3.1 Límites de confianza de procesos no observados

Muchos experimentos en física consisten en verificar leyes teóricas de conservación buscando reacciones o desintegraciones prohibidas por dichas leyes. Por ejemplo, experimentos doble  $\beta$  decay que violan la conservación del número leptónico.

- Si se detecta algún suceso la ley es rechazada.
- Si no se observa ningún proceso no puede asegurarse que la ley sea cierta, pero se pueden establecer límites, por ejemplo, para la vida media de la reacción o de la desintegración.

Se trata de cuantificar una ley de conservación a partir de resultados con cuentas nulas.

Supongamos que  $\lambda$  es la probabilidad de un suceso por unidad de tiempo.

- Según la distribución de Poisson, el promedio de sucesos en un intervalo de observación  $T$  será  $\mu = \lambda T$ .
- La probabilidad de no observar ningún suceso será:

$$f_p(x; \mu) = f_p(0; \lambda T) = \frac{\mu^0}{0!} e^{-\lambda T} = e^{-\lambda T}$$

La expresión se puede interpretar como una distribución de probabilidad, donde  $\lambda$  es la variable y no se observan cuentas en un período  $T$ .

- Multiplicando por  $T$  para que la  $fdp$  esté normalizada, se puede calcular, mediante la función acumulativa, la probabilidad de que  $\lambda$  sea menor que un cierto  $\lambda_o$

$$P(\lambda \leq \lambda_o) = \int_0^{\lambda_o} T e^{-\lambda T} d\lambda = 1 - e^{-\lambda_o T} = CL$$

lo cual se conoce como el **nivel de confianza** para el intervalo entre 0 y  $\lambda_o$ .

Elegido un determinado CL (ejemplo: CL=90%) se obtiene el límite superior de la probabilidad por unidad de tiempo de que ocurra el suceso.

$$1 - e^{-\lambda_o T} = CL \rightarrow \lambda_o = -\frac{1}{T} \ln(1 - CL)$$

En general, para un determinado CL, el intervalo no es único. Distintos límites en la integral pueden proporcionar el mismo valor del área bajo la curva de probabilidad y dar lugar al mismo CL. Como regla general, conviene tomar límites que cubran el menor intervalo posible de  $\lambda$ .

### 20.6.3.2 Distribución de intervalos temporales entre cuentas

Supongamos una fuente radioactiva cuya probabilidad de desintegración es  $\lambda$ . La fuente emite de forma aleatoria y queremos saber como es la distribución de los intervalos de tiempo en los que no se produce ninguna desintegración.

- La probabilidad de observar 0 cuentas en un período  $T$ :

$$f_p(0; \lambda T) = e^{-\lambda T}$$

Interesa interpretar la expresión anterior como una  $fdp$  donde la variable es el intervalo  $T$  en el cual no se registran cuentas, es decir el tiempo transcurrido entre desintegraciones sucesivas

- Normalizando la expresión:  $f_p(T; 0) = \lambda e^{-\lambda T}$

## 20.7 Ajuste de curvas

Una situación habitual en física es aquella en que disponemos de un conjunto de valores experimentales  $(x_i, y_i, \sigma_i)$ , con  $i = 1, 2, \dots, n$ , donde las  $x_i$  representan las abscisas,  $y_i$  son las ordenadas de los puntos medidos experimentalmente y las  $\sigma_i$  son los errores de medida que asociaremos únicamente a las ordenadas.

- Si las  $x_i$  también tienen error, éstos se pueden asociar por propagación a la variable  $y_i$ .

$$\left. \begin{matrix} x_i \pm \varepsilon(x_i) \\ y_i \pm \varepsilon(y_i) \end{matrix} \right\} \rightarrow \left\{ \begin{matrix} \varepsilon_1(y_i) = \varepsilon(y_i) \\ \varepsilon_2(y_i) = \frac{\partial y}{\partial x_i} \varepsilon(x_i) \end{matrix} \right\} \rightarrow \sigma_i = \sqrt{\varepsilon_1^2(y_i) + \varepsilon_2^2(y_i)}$$

### 20.7.1 Método de mínimos cuadrados.

Utilizando un lenguaje estadístico, podemos decir que cada uno de los valores experimentales  $y_i$  corresponde a una extracción muestral de un universo con densidad de probabilidad:

$$f_g(y_i; \mu_i, \sigma_i) = \frac{1}{\sigma_i \sqrt{2\pi}} e^{-(y_i - \mu_i)^2 / 2\sigma_i^2}$$

donde  $\mu_i$  es el valor verdadero de la magnitud medida para cada valor de la abscisa  $x_i$ . Los valores  $\mu_i$  suelen ser desconocidos y hay que estimarlos a partir de los datos experimentales.

- Supongamos que existe una función matemática, que depende de una serie de parámetros  $\vec{p} \equiv \{p_1, p_2, \dots, p_m\}$ , y que relaciona los valores de las abscisas y las ordenadas:  $y = f(x, \vec{p})$ .

Si la función es realmente una representación perfecta del experimento, una vez determinados los parámetros, el mejor estimador de  $\mu_i$  será aquel que proporcione la propia función, es decir:

$$\mu_i \approx y(x_i) = f(x_i, p_1, p_2, \dots, p_m) = f(x_i, \vec{p})$$

- La distribución de probabilidad de cada medida  $y_i$  :

$$f_g(y_i; \mu_i, \sigma_i) = \frac{1}{\sigma_i \sqrt{2\pi}} e^{-(y_i - f(x_i, \vec{p}))^2 / 2\sigma_i^2}$$

- La distribución de probabilidad conjunta si los  $n$  valores  $y_i$  son independientes vendrá dada por:

$$P(y_1, y_2, \dots, y_n) = \prod_{i=1}^n f_g(y_i; \mu_i, \sigma_i) = \frac{1}{(2\pi)^{n/2}} \prod_{i=1}^n \left( \frac{1}{\sigma_i} \right) e^{-\chi^2/2}$$

donde se ha definido:  $\chi^2 = \sum_{i=1}^n \frac{1}{\sigma_i^2} [y_i - f(x_i, \vec{p})]^2$

Por el principio de máxima verosimilitud, la mejor estimación de los parámetros es la que hace máxima la función distribución de probabilidad conjunta.

**Totalmente equivalente a que la función  $\chi^2$  sea mínima.**

- El método de mínimos cuadrados, implica hacer mínimos los cuadrados de las distancias entre los valores experimentales y la función modelo, afectados éstos por un peso  $w_i = 1/\sigma_i^2$ .
- Las condiciones matemáticas de mínimo corresponden a la anulación de todas las derivadas:

$$\frac{\partial \chi^2}{\partial p_k} = 0 \quad ; \quad k = 1, 2, \dots, m$$

- Obtener los parámetros consistirá en resolver el sistema de ecuaciones acopladas.

- Los errores se obtendrán a partir de la matriz de covarianzas calculada en el mínimo:

$$V_{ls} = \frac{1}{2} \frac{\partial^2 (\chi^2)}{\partial p_l \partial p_s}$$

- La matriz,  $V^{-1}$ , se corresponde con la matriz de errores. Las varianzas de los parámetros son los elementos de la diagonal  $\rightarrow \sigma_j^2 = (V^{-1})_{jj}$
- Si los errores son gaussianos  $\chi^2$  se distribuirá como una función  $f(\chi^2, \nu)$  con  $\nu = n - m$  grados de libertad, lo que permitirá calcular la bondad del ajuste.

### Casos particulares:

- Si la **función modelo**  $f(x, \vec{p})$  depende linealmente de los parámetros, es decir se puede escribir como:

$$f(x, \vec{p}) = \sum_{k=1}^m p_k f_k(x), \text{ donde las } f_k \text{ son funciones arbitrarias, el problema tiene solución única y fácil de determinar.}$$

- En general, cuando la dependencia no es lineal conviene obtener los parámetros mediante una búsqueda directa del mínimo de  $\chi^2$ , en lugar de resolver el sistema de ecuaciones acopladas.
- Si la función  $\chi^2$  no corresponde a un modelo lineal, en la mayoría de los casos se puede llevar a cabo un desarrollo en serie de Taylor alrededor del mínimo, hasta segundo orden en los parámetros y utilizar métodos semejantes al caso de *funciones modelo lineales*, pero esto sólo en el entorno del mínimo.

#### 20.7.1.1 Ajustes lineales

Cuando la función depende linealmente de los parámetros, o es linealizabile, se puede utilizar la siguiente representación:

$$f(x, \vec{p}) = p_1 f_1(x) + p_2 f_2(x) + \dots + p_m f_m(x)$$

donde las  $f_k(x)$  son cualquier tipo de función que no depende de los parámetros. Las ecuaciones generales se convierten en:

$$\frac{\partial \chi^2}{\partial p_k} = -2 \sum_{i=1}^n [y_i - p_1 f_1(x_i) - \dots - p_m f_m(x_i)] \frac{f_k(x_i)}{\sigma_i^2} = 0$$

$$\sum_{i=1}^n \frac{f_k(x_i) y_i}{\sigma_i^2} = \sum_{j=1}^m \sum_{i=1}^n \frac{f_k(x_i) f_j(x_i)}{\sigma_i^2} p_j, \quad k = 1, 2, \dots, m$$

- Sistema de  $m$  ecuaciones lineales que hay que resolver para obtener los parámetros  $p_k$ .

Definiendo un vector columna,  $\vec{v}$ , y una matriz simétrica  $F$ , las ecuaciones anteriores se pueden escribir en forma matricial.

$$v_k = \sum_{i=1}^n \frac{f_k(x_i)}{\sigma_i^2} y_i \quad ; \quad F_{kj} = \sum_{i=1}^n \frac{f_k(x_i) f_j(x_i)}{\sigma_i^2} \quad \rightarrow \quad v_k = \sum_{j=1}^m F_{kj} p_j$$

$$\vec{v} = F \vec{p} \quad \rightarrow \quad \vec{p} = F^{-1} \vec{v} \quad \rightarrow \quad \boxed{p_k = \sum_{i=1}^m C_{ki} v_i}$$

El método permite ajustar cualquier polinomio, en particular líneas rectas y parábolas, e incluso funciones más complicadas siempre que sean lineales en los parámetros (ej.: fórmula semiempírica de la masa). La matriz  $C = F^{-1}$  es la **matriz de covarianzas** y los elementos diagonales son las varianzas de los parámetros. Los elementos no diagonales son las covarianzas y nos proporcionan las correlaciones entre los parámetros. Las covarianzas deben ser nulas para la población si los parámetros son independientes. Sin embargo, en una muestra finita las covarianzas nunca llegan a ser nulas, porque los parámetros determinados por el método de mínimos cuadrados no son estadísticamente independientes, pero han de ser compatibles con el valor cero y tender a él conforme aumenta la estadística.

$$\varepsilon^2(p_k) = \sigma^2(p_k) = (C)_{kk} = (F^{-1})_{kk} \quad ; \quad \sigma_{ij}^2 = C_{ij} \rightarrow \text{corr}(p_i, p_j) = \frac{C_{ij}}{\sqrt{C_{ii} C_{jj}}}$$

### Ejemplo 1: Ajuste de rectas.

Supongamos que para determinar los dos parámetros ( $a_1, a_2$ ) de la relación lineal  $y = a_1 x + a_2$  se dispone de la siguiente tabla de datos, donde las  $\sigma_i$  son las desviaciones típicas de las  $y_i$ ,

$x_i$	1	2	3	4	5	6
$y_i$	1,5	2,5	4,0	3,6	5,9	6,1
$\sigma_i$	0,2	0,1	0,2	0,4	0,1	0,2

Siguiendo los pasos descritos en el apartado anterior calcularemos, en primer lugar, el vector columna y los elementos de la matriz  $F$ , a partir de estos valores obtendremos la matriz  $C$ , el vector de parámetros y las desviaciones típicas.

$$v_k = \sum_{i=1}^n \frac{f_k(x_i)}{\sigma_i^2} y_i \quad \rightarrow \quad \begin{cases} f_1(x) = x \\ f_2(x) = 1 \end{cases} \quad \rightarrow \quad \begin{cases} v_1 = \sum_{i=1}^n \frac{x_i y_i}{\sigma_i^2} = 4792,5 \\ v_2 = \sum_{i=1}^n \frac{y_i}{\sigma_i^2} = 1152,5 \end{cases}$$

$$F_{kj} = \sum_{i=1}^n \frac{f_k(x_i) f_j(x_i)}{\sigma_i^2} \quad \rightarrow \quad \begin{cases} F_{11} = \sum_{i=1}^n \frac{x_i^2}{\sigma_i^2} = 4150 \\ F_{12} = F_{21} = \sum_{i=1}^n \frac{x_i}{\sigma_i^2} = 975 \\ F_{22} = \sum_{i=1}^n \frac{1}{\sigma_i^2} = 281,25 \end{cases} \quad \rightarrow \quad F = \begin{pmatrix} 4150 & 975 \\ 975 & 281,25 \end{pmatrix}$$



$$C = F^{-1} = 10^{-3} \begin{pmatrix} 1,299 & -4,502 \\ -4,502 & 19,163 \end{pmatrix} \Rightarrow \boxed{\vec{a} = C\vec{v} = \begin{pmatrix} 1,037 \\ 0,510 \end{pmatrix}} ; \sigma(a_k) = \sqrt{C_{kk}} \Rightarrow \boxed{\sigma_1 = 0,036} \\ \boxed{\sigma_2 = 0,138}$$

El mejor ajuste corresponde a la ecuación:  $y = (1,04 \pm 0,04)x + (0,51 \pm 0,14)$  y la correlación entre los parámetros se calcula a partir de la relación siguiente, este valor debe ser próximo a  $\pm 1$  si la correlación es lineal.

$$\text{corr}(a_1, a_2) = \frac{C_{12}}{\sqrt{C_{11}C_{22}}} = 0,9023$$

### Ejemplo 2: Ajuste de parábolas.

Pasemos ahora a tratar el ajuste de una relación cuadrática, por ejemplo la parábola de ecuación  $y = a_1x^2 + a_2x + a_3$ , para ello seguiremos los pasos efectuados para el ajuste de la recta aunque sin asignar valores numéricos.

$f_1(x) = x^2$	$f_2(x) = x$	$f_3(x) = 1$
$v_1 = \sum_{i=1}^n \frac{x_i^2 y_i}{\sigma_i^2}$	$v_2 = \sum_{i=1}^n \frac{x_i y_i}{\sigma_i^2}$	$v_3 = \sum_{i=1}^n \frac{y_i}{\sigma_i^2}$
$F_{11} = \sum_{i=1}^n \frac{x_i^4}{\sigma_i^2}$	$F_{12} = F_{21} = \sum_{i=1}^n \frac{x_i^3}{\sigma_i^2}$	$F_{13} = F_{31} = \sum_{i=1}^n \frac{x_i^2}{\sigma_i^2}$
$F_{23} = F_{32} = \sum_{i=1}^n \frac{x_i}{\sigma_i^2}$	$F_{22} = \sum_{i=1}^n \frac{x_i^2}{\sigma_i^2}$	$F_{33} = \sum_{i=1}^n \frac{1}{\sigma_i^2}$

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \Rightarrow \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = F^{-1} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \text{ con } \sigma(a_i) = \sqrt{(F^{-1})_{ii}} = \sqrt{C_{ii}}$$

El método es aplicable a polinómios de orden superior, pero los cálculos se complican al aumentar el número de parámetros. Recurriendo al cálculo numérico es posible invertir matrices de dimensión elevada.

#### 20.7.1.2 Ajustes no lineales

Cuando la función modelo depende de los parámetros en forma no lineal, la estrategia de determinación del mínimo de  $\chi^2$  se basa en el uso de procedimientos generales de minimización de funciones de muchas variables. Dado que la resolución del sistema de ecuaciones no lineales y acopladas comporta serias dificultades, el procedimiento básico consiste en escribir la función  $\chi^2$  adecuada y aplicar algoritmos de minimización, como por

ejemplo el *Simplex*. El problema es que estos algoritmos no proporcionan la matriz de covarianzas. Sin embargo, en la mayoría de los casos,  $\chi^2$  es una función "educada" de los parámetros, lo cual significa que cerca del mínimo se comporta como una forma cuadrática. Aunque, estrictamente, el carácter cuadrático corresponde a los desplazamientos de los parámetros respecto al punto del mínimo.

- Si  $\vec{p}^0$  es el vector de parámetros en el mínimo y  $\vec{p}$  el de los parámetros genéricos, hablar de comportamiento cuadrático quiere decir que se verifica, en una región apreciablemente grande alrededor del mínimo:

$$\chi^2(\vec{p}) \approx \chi^2(\vec{p}^0) + [\vec{p} - \vec{p}^0]^T F [\vec{p} - \vec{p}^0]$$

donde F es, en principio, una matriz cuadrada del mismo tipo que la utilizada en el estudio de modelos lineales, pero que además se corresponde con la matriz de derivadas segundas.

- Para efectuar un desarrollo en serie de Taylor alrededor del mínimo hay que conocer las derivadas primera y segunda de  $\chi^2$  respecto de los parámetros, teniendo en cuenta que cuando las derivadas se calculan en el punto  $\vec{p}^0$  del mínimo la derivada primera es nula.

$$\frac{\partial \chi^2}{\partial p_k} = -2 \sum_i \left( \frac{y_i - f(x_i, \vec{p})}{\sigma_i} \right) \frac{1}{\sigma_i} \frac{\partial f(x_i, \vec{p})}{\partial p_k}$$

$$\frac{\partial^2 \chi^2}{\partial p_k \partial p_l} = \sum_i \frac{2}{\sigma_i^2} \frac{\partial f(x_i, \vec{p})}{\partial p_k} \frac{\partial f(x_i, \vec{p})}{\partial p_l} - \sum_i \left( \frac{y_i - f(x_i, \vec{p})}{\sigma_i} \right) \frac{2}{\sigma_i} \frac{\partial^2 f(x_i, \vec{p})}{\partial p_k \partial p_l}$$

- De este modo, el desarrollo en serie de Taylor corresponde a la forma cuadrática mencionada anteriormente, donde  $\mathbf{F}$  es la matriz de derivadas segundas también denominada *hessiano*.

$$F_{kl} = \frac{1}{2} \frac{\partial^2 \chi^2}{\partial p_k \partial p_l}$$

- Cerca del mínimo el desarrollo en serie alrededor de un punto  $\vec{p}$  proporciona la ecuación que permite obtener el desplazamiento  $\Delta \vec{p}$  que nos lleva al mínimo.

$$\chi^2(\vec{p} + \Delta \vec{p}) = \chi^2(\vec{p}) + \sum_k \frac{\partial \chi^2}{\partial p_k} \Delta p_k + \frac{1}{2} \sum_k \sum_l \frac{\partial^2 \chi^2}{\partial p_k \partial p_l} \Delta p_k \Delta p_l + \dots$$

- Calculando la derivada respecto a  $\Delta p_k$  e igualando a cero resulta un sistema lineal de ecuaciones, donde las incógnitas son los desplazamientos de los parámetros y cuya estructura es equivalente al caso de ajustes lineales

$$\frac{\partial \chi^2(\vec{p} + \Delta \vec{p})}{\partial (\Delta p_k)} = \frac{\partial \chi^2}{\partial p_k} + \frac{1}{2} \left( 2 \sum_l \frac{\partial^2 \chi^2}{\partial p_k \partial p_l} \Delta p_l \right) + \dots = 0 \Rightarrow \sum_l \frac{\partial^2 \chi^2}{\partial p_k \partial p_l} \Delta p_l = - \frac{\partial \chi^2}{\partial p_k}$$

siendo ahora:

$$F_{kl} = \frac{1}{2} \frac{\partial^2 \chi^2}{\partial p_k \partial p_l} \approx \sum_i \frac{1}{\sigma_i^2} \frac{\partial f(x_i, \vec{p})}{\partial p_k} \frac{\partial f(x_i, \vec{p})}{\partial p_l} ; \quad v_k = -\frac{1}{2} \frac{\partial \chi^2}{\partial p_k} = \sum_i \left( \frac{y_i - f(x_i, \vec{p})}{\sigma_i} \right) \frac{1}{\sigma_i} \frac{\partial f(x_i, \vec{p})}{\partial p_k}$$

Una vez resuelto el sistema lineal, la matriz inversa de  $F$  es una aproximación a la matriz de covarianzas y con el mismo significado que en el caso de modelos lineales.

Básicamente hay cuatro pasos a seguir:

- Partir de valores aproximados de los parámetros.
- Utilizar un algoritmo (simplex) para situarnos tan cerca del mínimo como sea posible.
- Construir la matriz  $F$  y el vector columna  $\vec{v}$ .
- Resolver el problema lineal, determinando los parámetros óptimos y la matriz de covarianzas.

### 20.7.1.3 Bondad del ajuste

Cuando se estiman parámetros mediante el  $\chi^2$  es necesario determinar los errores o intervalos de confianza asociados a éstos. Asimismo, hay que concluir si el ajuste tiene sentido o no. Necesitamos, por tanto, un **control de calidad** del ajuste basado en una serie de criterios.

Recordemos que:  $E[\chi^2] = \nu = n - m$  ,  $\sigma(\chi^2) = \sqrt{2\nu}$  .

Criterios:

1. Si el ajuste da como resultado un valor de  $\chi^2$  que difiere de  $\nu$  hasta un total de 3 desviaciones típicas, diremos que el ajuste es aceptable.
2. Si  $\chi^2 \ll \nu$  , probablemente hemos asignado errores  $\sigma_i$  demasiado grandes. Conviene volver a analizar el experimento y los datos.
3. Si  $\chi^2 \gg \nu$  (más de 3 desviaciones), puede que:
  - a) Estemos trabajando con un conjunto incorrecto de medidas.
  - b) Hayamos asignado errores demasiado pequeños.
  - c) La función modelo utilizada es incorrecta.

### 20.7.1.4 Errores asociados a los parámetros del ajuste

Veamos cómo es la distribución de los valores obtenidos para  $p_1, \dots, p_m$  en los ajustes y cómo es su distribución conjunta.

#### **Caso 1: Errores con un solo parámetro**

- La función depende de un solo parámetro  $p$  , que en el mínimo de  $\chi^2$  vale  $p_0$  .

Desarrollando  $\chi^2$  en el mínimo se obtiene:  $\chi^2(p) = \chi^2_{\min} + \frac{1}{2} \frac{d^2 \chi^2}{dp^2} \Big|_{p_0} (p - p_0)^2 + \dots$  ,

siendo la derivada primera nula y la segunda positiva. La ecuación obtenida es una parábola

alrededor del mínimo, la derivada segunda especifica su forma y se pueden presentar las siguientes situaciones:

$$\left. \frac{d^2 \chi^2}{dp^2} \right|_{p_0} \begin{cases} \text{grande} & \left\{ \begin{array}{l} \bullet \text{ parábola muy cerrada} \\ \bullet \text{ buena estimación del parámetro} \end{array} \right. \\ \text{pequeña} & \left\{ \begin{array}{l} \bullet \text{ parábola muy abierta} \\ \bullet \text{ estimación pobre del parámetro} \end{array} \right. \end{cases}$$

por lo que la derivada segunda es una medida del error de  $p$ . En el caso que nos ocupa, dado que la función depende de un solo parámetro, la matriz de covarianzas es un único número  $C$ . Además, de acuerdo con un análisis estadístico, la distribución de los valores de  $p$  será gaussiana, centrada en  $p_0$  y con una desviación típica  $\sigma(p) = \sqrt{C}$ . Por tanto, podemos escribir:

$$\chi^2(p) = \chi^2_{\min} + \frac{1}{C}(p - p_0)^2$$

esta expresión permite evaluar el incremento que se produce en la variable  $\chi^2$  cuando el valor del parámetro varía en una determinada cantidad con respecto a  $p_0$ . En particular, para un incremento  $\sigma$  en el valor del parámetro, el  $\chi^2$  se incrementa en una unidad. Mientras que si el incremento es de  $2\sigma$ , el cambio en  $\chi^2$  es de 4 unidades. Basta, por tanto, ir variando el parámetro hasta observar el incremento de una unidad en  $\chi^2$  para tener una estimación del valor de  $\sigma$ , o lo que es lo mismo del error del parámetro.

$$\Delta p = (p - p_0) = \sigma \Rightarrow \chi^2 \rightarrow \chi^2 + 1 \qquad \Delta p = (p - p_0) = 2\sigma \Rightarrow \chi^2 \rightarrow \chi^2 + 4$$

### **Caso 2: Errores con muchos parámetros**

Cuando se trabaja con dos o más parámetros, el desarrollo de  $\chi^2$  cerca del mínimo  $\vec{p}^o$  corresponde a la expresión:

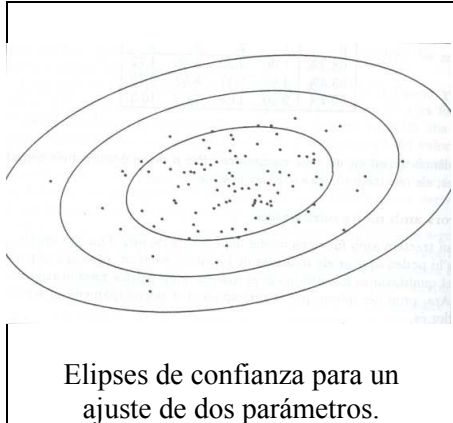
$$\chi^2(\vec{p}) = \chi^2_{\min} + F_{jk}(p_j - p_j^o)(p_k - p_k^o) + \dots$$

se trata de un paraboloide en un espacio de dimensión  $m + 1$ , que se estrechará en unas direcciones (parámetros bien determinados) y se ensanchará en otras (parámetros con mayor error). La estadística nos dice que los valores de los parámetros se distribuyen de acuerdo con una distribución gaussiana de muchas variables dada por:

$$f_g(\vec{p}) = \text{cte} \cdot e^{-\frac{1}{2}[\vec{p} - \vec{p}^o]^T F [\vec{p} - \vec{p}^o]}$$

y que ahora, en lugar de hablar de intervalos de confianza hemos de hablar de regiones de confianza. Por ejemplo, para el caso de dos parámetros las regiones son elipses dentro de las cuales el valor de  $\chi^2$  es menor que el valor en el mínimo más una cierta cantidad. La siguiente tabla proporciona valores de los incrementos de  $\chi^2$  para diferentes niveles de confianza  $P$  y número de parámetros.

<b>Incrementos de <math>\chi^2</math> según el Nivel de Confianza <math>P</math> y número de parámetros</b>				
<b><math>P</math></b>	<b><math>m = 1</math></b>	<b><math>m = 2</math></b>	<b><math>m = 3</math></b>	<b><math>m = 4</math></b>
68.3 %	1.00	2.30	3.53	4.72
95.4 %	4.00	6.17	8.02	9.70
99.7 %	9.00	11.8	14.2	16.3



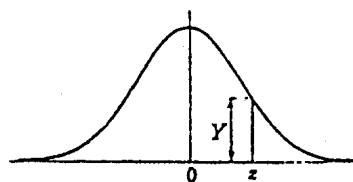
Si hacemos muchas medidas y ajustes, cada punto es resultado de un ajuste. La primera elipse contiene el 68.3% de los puntos, la segunda el 95.4%, etc.

La forma de la elipse depende de la matriz  $F$

$$F_{11}p_1^2 + 2F_{12}p_1p_2 + F_{22}p_2^2 = (\chi^2 - \chi_{\min}^2)$$

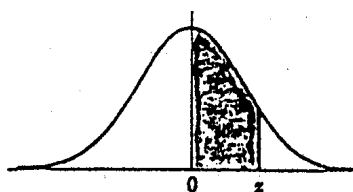
los elementos de la diagonal hacen de semiejes (errores) y el resto miden la rotación de la elipse (correlaciones).

ORDENADAS (Y)  
DE LA  
CURVA NORMAL  
TIPIFICADA  
EN z



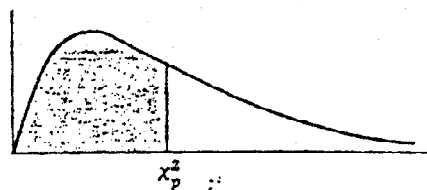
z	0	1	2	3	4	5	6	7	8	9
0.0	0.3989	0.3989	0.3989	0.3988	0.3986	0.3984	0.3982	0.3980	0.3977	0.3973
0.1	0.3970	0.3965	0.3961	0.3956	0.3951	0.3945	0.3939	0.3932	0.3925	0.3918
0.2	0.3910	0.3902	0.3894	0.3885	0.3876	0.3867	0.3857	0.3847	0.3836	0.3825
0.3	0.3814	0.3802	0.3790	0.3778	0.3765	0.3752	0.3739	0.3725	0.3712	0.3697
0.4	0.3683	0.3668	0.3653	0.3637	0.3621	0.3605	0.3589	0.3572	0.3555	0.3538
0.5	0.3521	0.3503	0.3485	0.3467	0.3448	0.3429	0.3410	0.3391	0.3372	0.3352
0.6	0.3332	0.3312	0.3292	0.3271	0.3251	0.3230	0.3209	0.3187	0.3166	0.3144
0.7	0.3123	0.3101	0.3079	0.3056	0.3034	0.3011	0.2989	0.2966	0.2943	0.2920
0.8	0.2897	0.2874	0.2850	0.2827	0.2803	0.2780	0.2756	0.2732	0.2709	0.2685
0.9	0.2661	0.2637	0.2613	0.2589	0.2565	0.2541	0.2516	0.2492	0.2468	0.2444
1.0	0.2420	0.2396	0.2371	0.2347	0.2323	0.2299	0.2275	0.2251	0.2227	0.2203
1.1	0.2179	0.2155	0.2131	0.2107	0.2083	0.2059	0.2036	0.2012	0.1989	0.1965
1.2	0.1942	0.1919	0.1895	0.1872	0.1849	0.1826	0.1804	0.1781	0.1758	0.1736
1.3	0.1714	0.1691	0.1669	0.1647	0.1626	0.1604	0.1582	0.1561	0.1539	0.1518
1.4	0.1497	0.1476	0.1456	0.1435	0.1415	0.1394	0.1374	0.1354	0.1334	0.1315
1.5	0.1295	0.1276	0.1257	0.1238	0.1219	0.1200	0.1182	0.1163	0.1145	0.1127
1.6	0.1109	0.1092	0.1074	0.1057	0.1040	0.1023	0.1006	0.0989	0.0973	0.0957
1.7	0.0940	0.0925	0.0909	0.0893	0.0878	0.0863	0.0848	0.0833	0.0818	0.0804
1.8	0.0790	0.0775	0.0761	0.0748	0.0734	0.0721	0.0707	0.0694	0.0681	0.0669
1.9	0.0656	0.0644	0.0632	0.0620	0.0608	0.0596	0.0584	0.0573	0.0562	0.0551
2.0	0.0540	0.0529	0.0519	0.0508	0.0498	0.0488	0.0478	0.0468	0.0459	0.0449
2.1	0.0440	0.0431	0.0422	0.0413	0.0404	0.0396	0.0387	0.0379	0.0371	0.0363
2.2	0.0355	0.0347	0.0339	0.0332	0.0325	0.0317	0.0310	0.0303	0.0297	0.0290
2.3	0.0283	0.0277	0.0270	0.0264	0.0258	0.0252	0.0246	0.0241	0.0235	0.0229
2.4	0.0224	0.0219	0.0213	0.0208	0.0203	0.0198	0.0194	0.0189	0.0184	0.0180
2.5	0.0175	0.0171	0.0167	0.0163	0.0158	0.0154	0.0151	0.0147	0.0143	0.0139
2.6	0.0136	0.0132	0.0129	0.0126	0.0122	0.0119	0.0116	0.0113	0.0110	0.0107
2.7	0.0104	0.0101	0.0099	0.0096	0.0093	0.0091	0.0088	0.0086	0.0084	0.0081
2.8	0.0079	0.0077	0.0075	0.0073	0.0071	0.0069	0.0067	0.0065	0.0063	0.0061
2.9	0.0060	0.0058	0.0056	0.0055	0.0053	0.0051	0.0050	0.0048	0.0047	0.0046
3.0	0.0044	0.0043	0.0042	0.0040	0.0039	0.0038	0.0037	0.0036	0.0035	0.0034
3.1	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026	0.0025	0.0025
3.2	0.0024	0.0023	0.0022	0.0022	0.0021	0.0020	0.0020	0.0019	0.0018	0.0018
3.3	0.0017	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014	0.0013	0.0013
3.4	0.0012	0.0012	0.0012	0.0011	0.0011	0.0010	0.0010	0.0010	0.0009	0.0009
3.5	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007	0.0007	0.0007	0.0006
3.6	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0004
3.7	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003	0.0003	0.0003	0.0003
3.8	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002	0.0002	0.0002	0.0002	0.0002
3.9	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0001	0.0001

AREAS  
BAJO LA  
CURVA NORMAL  
TIPIFICADA  
DE 0 a  $z$



$z$	0	1	2	3	4	5	6	7	8	9
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0754
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2258	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2518	0.2549
0.7	0.2580	0.2612	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2996	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998
3.5	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998
3.6	0.4998	0.4998	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.7	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.8	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999
3.9	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000

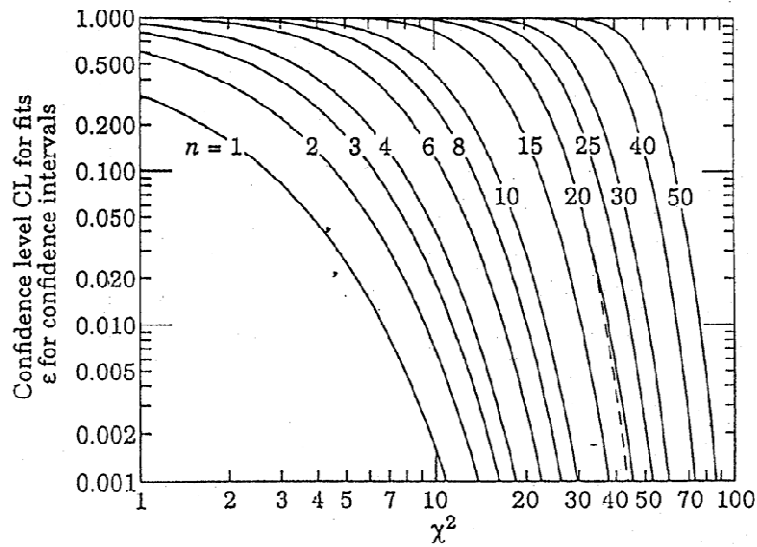
PERCENTILES ( $\chi_p^2$ )  
DE LA  
DISTRIBUCION CHI-CUADRADO  
CON  $\nu$  GRADOS DE LIBERTAD  
(AREA SOMBREADA =  $p$ )



$\nu$	$\chi_{0.995}^2$	$\chi_{0.99}^2$	$\chi_{0.975}^2$	$\chi_{0.95}^2$	$\chi_{0.90}^2$	$\chi_{0.75}^2$	$\chi_{0.50}^2$	$\chi_{0.25}^2$	$\chi_{0.10}^2$	$\chi_{0.05}^2$	$\chi_{0.025}^2$	$\chi_{0.01}^2$	$\chi_{0.005}^2$
1	7,88	6,63	5,02	3,84	2,71	1,32	0,455	0,102	0,0158	0,0039	0,0010	0,0002	0,0000
2	10,6	9,21	7,38	5,99	4,61	2,77	1,39	0,575	0,211	0,103	0,0506	0,0201	0,0100
3	12,8	11,3	9,35	7,81	6,25	4,11	2,37	1,21	0,584	0,352	0,216	0,115	0,072
4	14,9	13,3	11,1	9,49	7,78	5,39	3,36	1,92	1,06	0,711	0,484	0,297	0,207
5	16,7	15,1	12,8	11,1	9,24	6,63	4,35	2,67	1,61	1,15	0,831	0,554	0,412
6	18,5	16,8	14,4	12,6	10,6	7,84	5,35	3,45	2,20	1,64	1,24	0,872	0,676
7	20,3	18,5	16,0	14,1	12,0	9,04	6,35	4,25	2,83	2,17	1,69	1,24	0,989
8	22,0	20,1	17,5	15,5	13,4	10,2	7,34	5,07	3,49	2,73	2,18	1,65	1,34
9	23,6	21,7	19,0	16,9	14,7	11,4	8,34	5,90	4,17	3,33	2,70	2,09	1,73
10	25,2	23,2	20,5	18,3	16,0	12,5	9,34	6,74	4,87	3,94	3,25	2,56	2,16
11	26,8	24,7	21,9	19,7	17,3	13,7	10,3	7,58	5,58	4,57	3,82	3,05	2,60
12	28,3	26,2	23,3	21,0	18,5	14,8	11,3	8,44	6,30	5,23	4,40	3,57	3,07
13	29,8	27,7	24,7	22,4	19,8	16,0	12,3	9,30	7,04	5,89	5,01	4,11	3,57
14	31,3	29,1	26,1	23,7	21,1	17,1	13,3	10,2	7,79	6,57	5,63	4,66	4,07
15	32,8	30,6	27,5	25,0	22,3	18,2	14,3	11,0	8,55	7,26	6,26	5,23	4,60
16	34,3	32,0	28,8	26,3	23,5	19,4	15,3	11,9	9,31	7,96	6,91	5,81	5,14
17	35,7	33,4	30,2	27,6	24,8	20,5	16,3	12,8	10,1	8,67	7,56	6,41	5,70
18	37,2	34,8	31,5	28,9	26,0	21,6	17,3	13,7	10,9	9,39	8,23	7,01	6,26
19	38,6	36,2	32,9	30,1	27,2	22,7	18,3	14,6	11,7	10,1	8,91	7,63	6,84
20	40,0	37,6	34,2	31,4	28,4	23,8	19,3	15,5	12,4	10,9	9,59	8,26	7,43
21	41,4	38,9	35,5	32,7	29,6	24,9	20,3	16,3	13,2	11,6	10,3	8,90	8,03
22	42,8	40,3	36,8	33,9	30,8	26,0	21,3	17,2	14,0	12,3	11,0	9,54	8,64
23	44,2	41,6	38,1	35,2	32,0	27,1	22,3	18,1	14,8	13,1	11,7	10,2	9,26
24	45,6	43,0	39,4	36,4	33,2	28,2	23,3	19,0	15,7	13,8	12,4	10,9	9,89
25	46,9	44,3	40,6	37,7	34,4	29,3	24,3	19,9	16,5	14,6	13,1	11,5	10,5
26	48,3	45,6	41,9	38,9	35,6	30,4	25,3	20,8	17,3	15,4	13,8	12,2	11,2
27	49,6	47,0	43,2	40,1	36,7	31,5	26,3	21,7	18,1	16,2	14,6	12,9	11,8
28	51,0	48,3	44,5	41,3	37,9	32,6	27,3	22,7	18,9	16,9	15,3	13,6	12,5
29	52,3	49,6	45,7	42,6	39,1	33,7	28,3	23,6	19,8	17,7	16,0	14,3	13,1
30	53,7	50,9	47,0	43,8	40,3	34,8	29,3	24,5	20,6	18,5	16,8	15,0	13,8
40	66,8	63,7	59,3	55,8	51,8	45,6	39,3	33,7	29,1	26,5	24,4	22,2	20,7
50	79,5	76,2	71,4	67,5	63,2	56,3	49,3	42,9	37,7	34,8	32,4	29,7	28,0
60	92,0	88,4	83,3	79,1	74,4	67,0	59,3	52,3	46,5	43,2	40,5	37,5	35,5
70	104,2	100,4	95,0	90,5	85,5	77,6	69,3	61,7	55,3	51,7	48,8	45,4	43,3
80	166,3	112,3	106,6	101,9	96,6	88,1	79,3	71,1	64,3	60,4	57,2	53,5	51,2
90	128,3	124,1	118,1	113,1	107,6	98,6	89,3	80,6	73,3	69,1	65,6	61,8	59,2
100	140,2	135,8	129,6	124,3	118,5	109,1	99,3	90,1	82,4	77,9	74,2	70,1	67,3

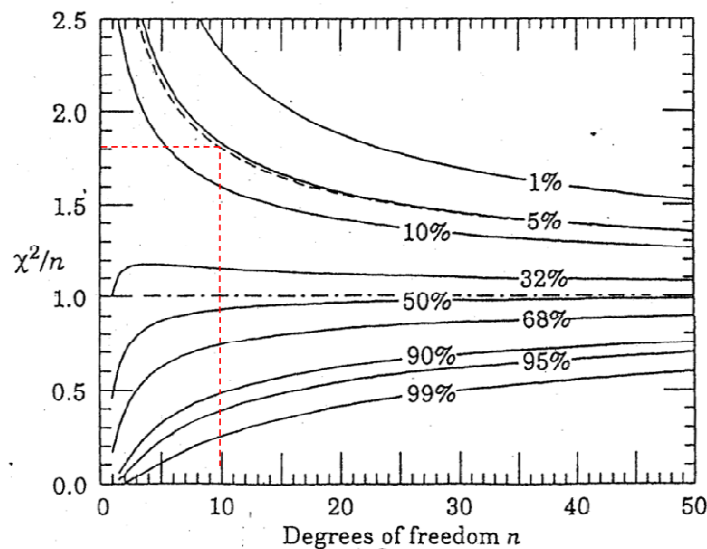


**Niveles de Confianza (CL) en función de  $\chi^2$  para diferentes grados de libertad  $n$**



**Figure 27.1:** The confidence level versus  $\chi^2$  for  $n$  degrees of freedom, as defined in Eq. (27.24). The curve for a given  $n$  gives the probability that a value at least as large as  $\chi^2$  will be obtained in an experiment; *e.g.*, for  $n = 10$ , a value  $\chi^2 \gtrsim 18$  will occur in 5% of a large number of experiments. For a fit, the CL is a measure of goodness-of-fit, in that a good fit to a correct model is expected to yield a low  $\chi^2$  (see Sec. 28.5.0). For a confidence interval,  $\alpha$  measures the probability that the interval *does not* cover the true value of the quantity being estimated (see Sec. 28.6). The dashed curve for  $n = 20$  is calculated using the approximation of Eq. (27.25).

**CL (%) en función del  $\chi^2$  reducido ( $\chi^2/n$ ) y el número de grados de libertad  $n$**



**Figure 27.3:** Confidence levels as a function of the “reduced  $\chi^2$ ”  $\equiv \chi^2/n$  and the number of degrees of freedom  $n$ . Curves are labeled by the probability that a measurement will give a value of  $\chi^2/n$  greater than that given on the  $y$  axis; *e.g.*, for  $n = 10$ , a value  $\chi^2/n \gtrsim 1.8$  can be expected 5% of the time.

## 20.8 Referencias

**Techniques for nuclear and particle physics experiments.** W. R. Leo. Ed. *Springer Verlag*. 1987.

**Mètodes Numèrics per a la Física.** R. Guardiola, E. Higón, J. Ros. Ed. *Universitat de València*, 1995.

**Radiation detection and measurement.** G.F. Knoll. Ed. *John Wiley and Sons*. 1979. *Segunda Edición* 1989.

**Statistics for nuclear and particle physicist.** L. Lyons. Ed. *Cambridge University Press*.

**Probability and statistics in particle physics.** A.G. Frodesen, O. Skjeggstad and H. Tofte. Ed. *Universitetsforlaget*, 1979.

**Data reduction and error analysis for the physical sciences.** Ph.D. Bevington. *Mc Graw Hill Book Co.* 1969, 1995.

**Análisis de errores.** C. Sánchez del Río. *Eudema Universidad*. 1989.