



Diploma de Postítulo en Geo-Minero-Metalurgia

Curso: Análisis estadístico de datos

Profesor: Xavier Emery

2008

Índice

Objetivos y alcances	2
Capítulo 1: Fundamentos de estadística matemática.....	3
1. Conceptos básicos de estadística.....	3
2. Conceptos básicos de probabilidades.....	6
3. Intervalos de confianza.....	14
Capítulo 2: Errores en datos.....	17
1. Precisión y exactitud	17
2. Propagación de errores	17
Capítulo 3: Estadística comparativa	20
1. Generalidades sobre los tests estadísticos	20
2. Comparación de dos proporciones	23
3. Comparación de dos medias	24
4. Otros tests usuales	28
Capítulo 4: Análisis de varianza	31
1. Análisis de varianza para diseños simples.....	31
2. Análisis de varianza para diseños anidados.....	33
Capítulo 5: Diseño de experimentos	36
1. Generalidades sobre el diseño de experimentos	36
2. Diseño de bloques aleatorios	38
3. Diseño de cuadrado latino	42
4. Diseño factorial.....	43
5. Diseño de pruebas industriales	49
Capítulo 6: Métodos de mínimos cuadrados	52
1. Modelamiento de datos	52
2. Principios básicos de los métodos de mínimos cuadrados	52
3. Ajuste de un modelo lineal de una variable.....	54
4. Ajuste de un modelo lineal de varias variables.....	59
5. Ajuste de un modelo polinomial.....	63
6. Ajuste de un modelo no lineal	65
Referencias	67
Anexo: Tablas estadísticas	68

Objetivos y alcances

Este curso busca entregar conocimientos teóricos y prácticos para estudiar, modelar e interpretar datos procedentes de muestreo o de experimentos o pruebas industriales. Se dirige a ingenieros de minas, metalurgistas y geólogos cuya labor requiere del diseño de experimentos, análisis de datos, prueba de hipótesis o modelamiento de relación causa-efecto entre variables.

El apunte se divide en los siguientes capítulos.

Parte 1: Análisis de datos

Lección 1: fundamentos de estadística matemática

Lección 2: errores en datos

Lección 3: estadística comparativa

Parte 2: Diseño de pruebas industriales

Lección 4: análisis de varianza

Lección 5: diseño de experimentos

Parte 3: Ajuste estadístico y modelamiento de datos

Lección 6: métodos de mínimos cuadrados

Capítulo 1: Fundamentos de estadística matemática

La estadística se puede definir como un conjunto de procedimientos, herramientas y técnicas utilizadas para recolectar, presentar y analizar datos, sobre los cuales basar decisiones en una situación de incertidumbre o frente a información incompleta, cuando no se puede conocer la realidad en forma exhaustiva. El modelamiento estadístico permite organizar nuestras elecciones y decisiones, para que éstas sean coherentes con lo que se conoce del fenómeno estudiado, aunque no permite legitimar las elecciones de manera absoluta (siempre existe la posibilidad de tomar decisiones erróneas).

Ejemplos de aplicación incluyen:

- control de calidad: estimar el promedio de vida y la dispersión de vida de un equipo;
- comparar las características de los equipos de distintos abastecedores;
- control de estándares de calidad en la toma, preparación y análisis de muestras;
- análisis de las características del mineral procesado (densidad, granulometría, ley, etc.);
- diseño de experimentos e interpretación de sus resultados.

1. Conceptos básicos de estadística

1.1. Población y muestra

A continuación, definiremos nociones básicas para el modelamiento estadístico.

La **población** es el conjunto, ya sea finito o infinito, de elementos o *individuos* que interesa considerar y podrían ser accesibles en el estudio. En general no se conoce la población entera (a menos de hacer un censo completo, en caso de que la población sea finita y que las pruebas a realizar no sean destructivas). El **muestreo** consiste en observar una fracción de la población, con la finalidad de hacer inferencias sobre dicha población. El error cometido al sacar conclusiones sobre cierta realidad (la población) a partir de la observación de sólo una parte de ella (la muestra), se conoce como el *error de muestreo*. Claramente, la manera con la cual se obtiene la muestra a partir de la población determina la calidad y la precisión de la información aportada por la muestra. Para que la inducción sea válida, la muestra debe ser **representativa** de la población.

Los métodos de muestreo *probabilísticos* son aquellos que se basan en el principio de **equiprobabilidad**, es decir, aquellos en los que todos los individuos de la población tienen la misma probabilidad de ser elegidos para formar parte de la muestra. Sólo estos métodos aseguran la representatividad de la muestra extraída y son, por lo tanto, los más recomendables.

Dentro de los métodos de muestreo probabilísticos encontramos:

- El **muestreo aleatorio simple**. Consiste en numerar cada individuo de la población y, a través de un mecanismo de sorteo (bolas dentro de una bolsa, tablas de números aleatorios, números aleatorios generados con una calculadora o con un computador), elegir tantos individuos como sea necesario para completar el tamaño de muestra requerido.
- El **muestreo aleatorio estratificado**. Consiste en definir categorías diferentes entre sí (*estratos*) que sean homogéneas con respecto a alguna característica. Los estratos funcionan en forma independiente, pudiendo aplicarse dentro de ellos un muestreo aleatorio simple para elegir los individuos que formarán parte de la muestra. Este tipo de muestreo busca asegurarse de que todos los estratos estén adecuadamente representados en la muestra.
- El **muestreo sistemático**. Sea k el resultado de dividir el tamaño de la población por el tamaño de la muestra deseada. Se numera todos los elementos de la población y se elige un número (i) al azar entre 1 y k . Los individuos que integran la muestra son los que ocupan los lugares $i, i+k, i+2k, \dots, i+(n-1)k$, es decir se toman los individuos de k en k . Este tipo de muestreo es riesgoso cuando existen periodicidades en la población: al elegir a los individuos de la muestra con una periodicidad constante (k), se puede introducir una homogeneidad que no existe en la población.

Se llama *parámetros* a las características de la población que se desea investigar y que suelen ser desconocidas a priori. Se habla de **variables** cuando las características de la población son numéricas, es decir, se pueden medir (ya sea en una escala continua o una escala discreta), y de **atributos** cuando estas características no son susceptibles de medirse numéricamente, sino que sólo presentan modalidades o categorías. A menudo, es posible codificar un atributo en una variable numérica discreta, ya sea categórica o binaria (presencia / ausencia de una característica).

1.2. Distribución de frecuencia

La **distribución de frecuencia** fracciona los datos en grupos o clases e indica el número de observaciones en cada clase (Tabla 1), o el número de observaciones en cada clase dividido por el número total de observaciones. Un *histograma* es un gráfico de barras que representa una distribución de frecuencia: las clases se miden en el eje de abscisa, mientras que el número de observaciones o las frecuencias se miden en el eje de ordenada (Figura 1).

Clase	Número de observaciones	Clase	Número de observaciones
0	0	1.6	77
0.1	0	1.7	64
0.2	15	1.8	45
0.3	75	1.9	42
0.4	132	2	48
0.5	178	2.1	34
0.6	152	2.2	19
0.7	187	2.3	14
0.8	192	2.4	13
0.9	185	2.5	9
1	177	2.6	10
1.1	174	2.7	10
1.2	144	2.8	3
1.3	132	2.9	2
1.4	119	3	4
1.5	95	y mayor...	25

Tabla 1. Tabla de frecuencia para datos de ley de cobre

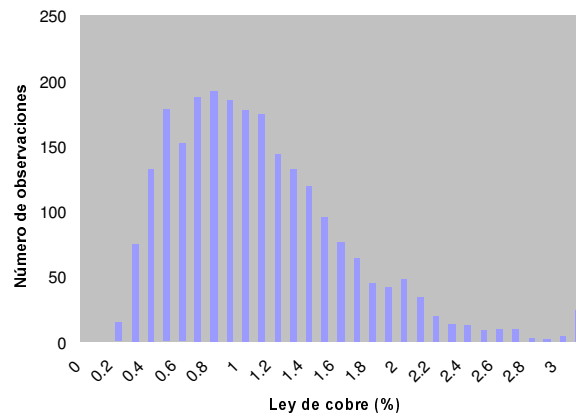


Figura 1. Histograma asociado a la tabla 1 de frecuencia.

La *distribución de frecuencia acumulada* muestra, para cada clase, el número total de observaciones en todas las clases inferiores así como en la clase en cuestión, dividido eventualmente por el número total de observaciones (Tabla 2). La representación gráfica de dicha distribución se hace mediante un *histograma acumulado* (Figura 2).

Clase de tamaño de partícula	Malla superior (cm)	Malla inferior (cm)	Proporción en clase	Proporción acumulada
L0	----	5.000	0.0204	0.0204
L1	5.000	3.800	0.0597	0.0801
L2	3.800	3.200	0.0597	0.1398
L3	3.200	2.500	0.0759	0.2157
L4	2.500	1.900	0.1383	0.3540
L5	1.900	1.300	0.1622	0.5162
L6	1.300	1.000	0.0813	0.5975
L7	1.000	0.600	0.0962	0.6937
L8	0.600	0.055	0.2412	0.9349
L9	0.055	0.000	0.0651	1.0000

Tabla 2. Tabla de frecuencia acumulada para datos de granulometría

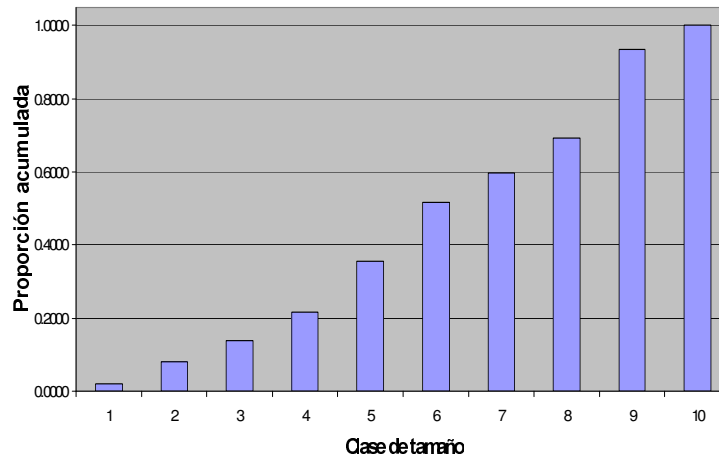


Figura 2. Histograma acumulado asociado a la tabla 2 de frecuencia acumulada.

2. Conceptos básicos de probabilidades

2.1. *Espacio muestral*

Una experiencia aleatoria se modela a partir de un espacio muestral definido por tres elementos:

- el conjunto de los resultados posibles o *universo*.
- el conjunto de proposiciones (*eventos*) que se puede enunciar sobre los elementos del universo.
- una probabilidad (es decir, una medida positiva de peso total 1) definida sobre los eventos.

Así, se sustituye la experiencia aleatoria real por la colección de todos sus resultados posibles y de las probabilidades de realización de los eventos asociados. Por ejemplo, para el lanzamiento de un dado no cargado, el universo es $\Omega = \{1,2,3,4,5,6\}$, los eventos son el conjunto de las partes de Ω (subconjuntos de Ω), mientras que la probabilidad atribuye a cada elemento de Ω un peso igual a $1/6$.

2.2. *Variables aleatorias y distribuciones de probabilidad*

Una **variable aleatoria** X es una magnitud que, a todo elemento del universo, asocia un valor numérico o *realización* en \mathbb{R} (variable aleatoria continua) o \mathbb{N} (variable discreta entera). Una variable aleatoria X se caracteriza por una distribución de probabilidad, la cual se representa por medio de

- Una **función de distribución**, definida por:

$$\forall x \in \mathbb{R}, P(x) = \text{Prob}(X < x).$$

P es una función no decreciente que toma valores en el intervalo $[0,1]$. Se tiene: $P(-\infty) = 0$ y $P(+\infty) = 1$.

- Una **densidad de probabilidad** (si la variable es continua): se trata de una función p positiva tal que

$$\forall x \in \mathbb{R}, P(x) = \int_{-\infty}^x p(t) dt.$$

- Una **masa de probabilidad** (si la variable es discreta, por ejemplo, entera), definida por

$$\forall n \in \mathbb{N}, p(n) = \text{Prob}(X = n).$$

Al sortear numerosos valores independientes de X , la distribución de frecuencia de los valores sorteados (realizaciones) tiende a la distribución de probabilidad, acorde a la *ley de los grandes números*.

2.3. Momentos de una variable aleatoria

Se suele considerar parámetros sintéticos (*momentos*) para describir la distribución de probabilidad. Entre ellos, los más importantes son:

- la **esperanza** o **valor esperado**, que representa el valor medio en torno al cual los valores de X se distribuyen según su distribución de probabilidad. Para una variable continua, se define por

$$\mu = E(X) = \int_{\mathbb{R}} x p(x) dx.$$

En cambio, para una variable discreta (entera), la esperanza es:

$$\mu = E(X) = \sum_{n \in \mathbb{N}} n p(n).$$

- la **varianza**, que mide la dispersión en torno al valor esperado:

$$\sigma^2 = \text{var}(X) = E\{(X - \mu)^2\} = E(X^2) - \mu^2 \geq 0.$$

La raíz cuadrada de la varianza es la **desviación estándar**, denotada σ , y expresada en la misma unidad que la variable X .

2.4. Estimadores de los momentos

Dado un conjunto de n observaciones independientes de X , o sea $\{X_1, \dots, X_n\}$, se puede estimar el valor esperado y la varianza por

- la **media experimental**, definida por $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$
- la **varianza experimental**, definida por $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$.

Si se restituye a X_1, \dots, X_n su carácter aleatorio, se ve que la media experimental y la varianza experimental son variables aleatorias. Se tiene los siguientes resultados:

- $E(\bar{X}) = \mu$: la media experimental estima la esperanza sin sesgo (error sistemático).
- $\text{var}(\bar{X}) = \frac{\sigma^2}{n}$: la media tiene menos dispersión que cada observación individual.
- $\bar{X} \xrightarrow{n \rightarrow \infty} \mu$: la media converge a la esperanza cuando el tamaño de la muestra n se vuelve infinitamente grande.
- $E(S^2) = \sigma^2$: la varianza experimental estima la varianza sin sesgo.
- $S^2 \xrightarrow{n \rightarrow \infty} \sigma^2$: la varianza experimental converge a la varianza cuando el tamaño de la muestra n se vuelve infinitamente grande.

2.5. Ejemplos de distribuciones de probabilidad

A continuación, introduciremos algunas distribuciones de probabilidad comúnmente utilizadas en el análisis de datos.

Distribución uniforme

La densidad de probabilidad es constante en un intervalo $[a, b]$:

$$\forall x \in \mathbb{R}, p(x) = \begin{cases} \frac{1}{b-a} & \text{si } x \in [a, b] \\ 0 & \text{en caso contrario} \end{cases}$$

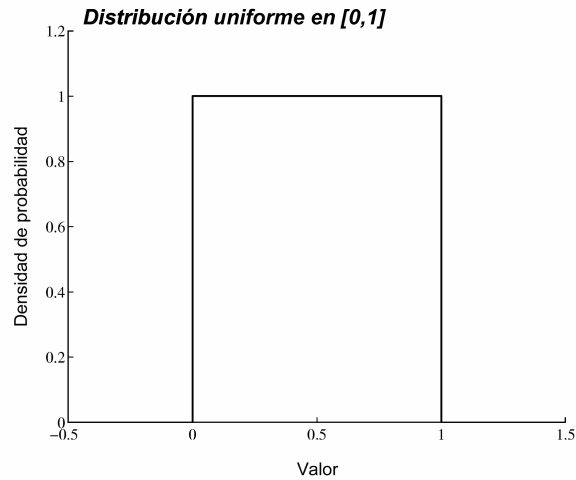


Figura 3. Densidad de probabilidad uniforme.

Distribución Gaussiana o normal

La densidad de probabilidad es la campana de Gauss:

$$\forall x \in \mathbb{R}, p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}.$$

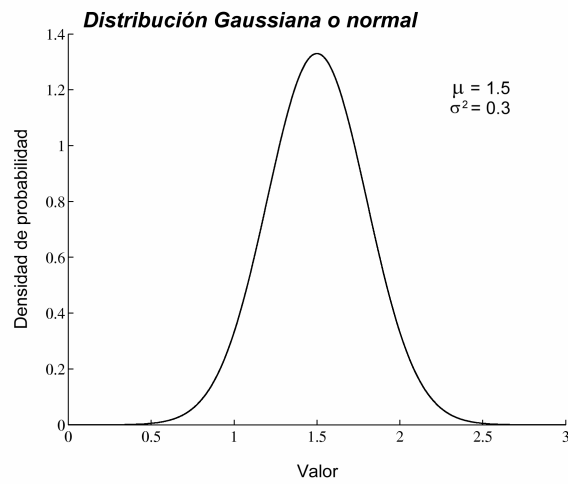


Figura 4. Densidad de probabilidad Gaussiana (normal).

La distribución normal **estándar** corresponde a aquella de media 0 y varianza 1; se denota usualmente como $N(0,1)$. La función de distribución normal estándar no tiene expresión analítica simple, aunque se puede aproximar (con un error menor a 10^{-5}) de la siguiente forma:

$$\begin{aligned}\forall x \in \mathbb{R}, P(x) &= \int_{-\infty}^x p(t) dt \\ &\approx 1 - p(x)(0.4361836t - 0.1201676t^2 + 0.9372980t^3)\end{aligned}$$

$$\text{con } t = \frac{1}{1 + 0.33267x}.$$

Una variable aleatoria normal de media μ y varianza σ^2 se obtiene de una variable aleatoria normal estándar al plantear:

$$N(\mu, \sigma^2) = \mu + \sigma N(0, 1).$$

Asimismo, la suma de n variables aleatorias normales independientes de media μ y varianza σ^2 es una variable aleatoria normal de media $n \times \mu$ y varianza $n \times \sigma^2$.

Distribución lognormal

X tiene distribución lognormal cuando su logaritmo sigue una distribución normal. La densidad de probabilidad es:

$$\forall x > 0, p(x) = \frac{1}{x \sigma_{\ln(X)} \sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left(\frac{\ln(x) - \mu_{\ln(X)}}{\sigma_{\ln(X)}} \right)^2 \right\}$$

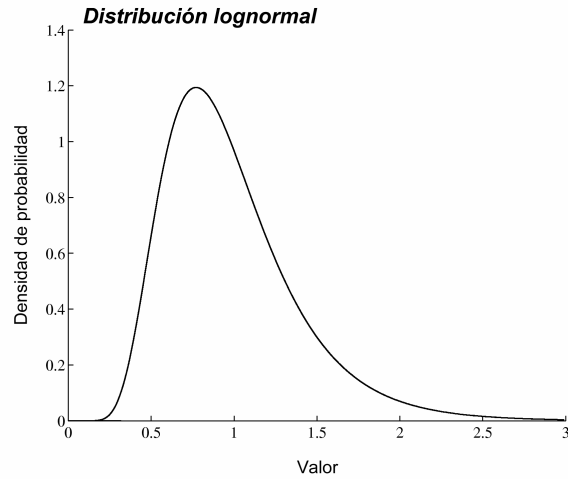


Figura 5. Densidad de probabilidad lognormal.

Distribución gamma y del chi cuadrado

La densidad de probabilidad de la distribución gamma estándar depende de un parámetro positivo θ (*parámetro de forma*), igual a la media y a la varianza:

$$\forall x > 0, p(x) = \frac{1}{\Gamma(\theta)} e^{-x} x^{\theta-1}.$$

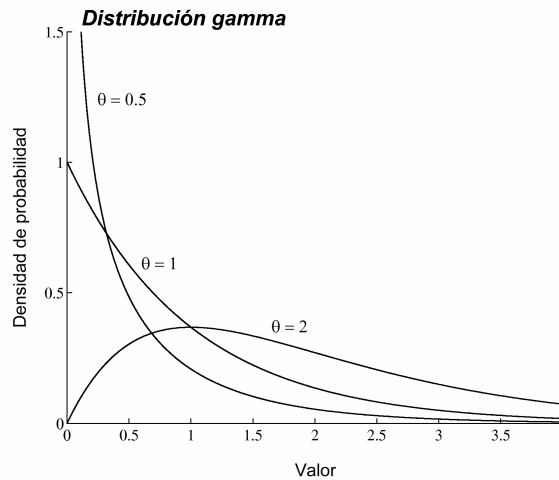


Figura 6. Densidad de probabilidad gamma.

El caso $\theta = 1$ corresponde a la **distribución exponencial**.

Si se consideran n variables normales estándar independientes (X_1, \dots, X_n) , entonces la variable U definida por

$$U = \frac{1}{2} \sum_{i=1}^n X_i^2$$

tiene una distribución gamma estándar de parámetro $\theta = n/2$.

La distribución de $2U$ se conoce como distribución del **chi cuadrado** con n grados de libertad y se denota como χ_n^2 . Su esperanza es igual a n y su varianza a $2n$.

Distribución de Weibull

Una variable X sigue una distribución de Weibull estándar de parámetro θ (positivo) si X^θ tiene una distribución exponencial:

$$\forall x > 0, p(x) = \theta x^{\theta-1} \exp(-x^\theta)$$

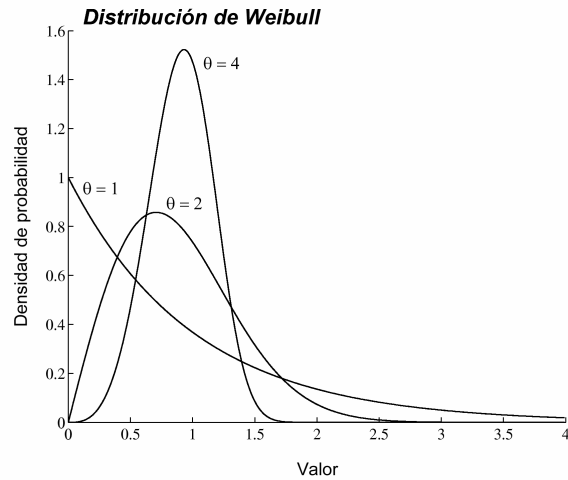


Figura 7. Densidad de probabilidad de Weibull.

Distribución de Bernoulli

Esta distribución es discreta y sólo tiene dos valores: 0 y 1. Si p_1 es la probabilidad de obtener 1 (probabilidad de éxito), entonces la esperanza de la variable de Bernoulli es p_1 y su varianza es $p_1 (1 - p_1)$.

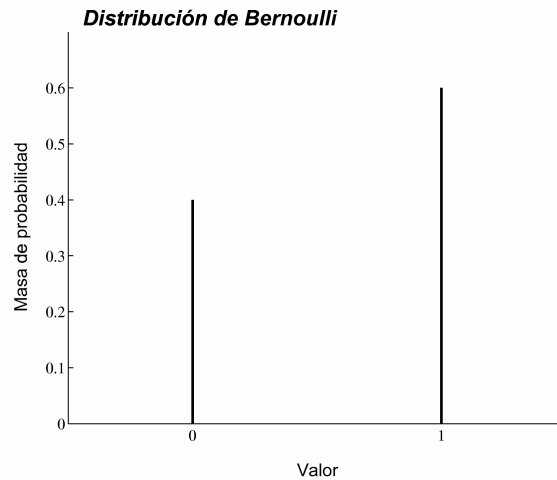


Figura 8. Masa de probabilidad de Bernoulli.

Distribución binomial

Una variable binomial se obtiene al sumar M variables de Bernoulli independientes, de misma probabilidad de éxito p_1 . La masa de probabilidad es:

$$\forall n \in \{0, \dots, M\}, p(n) = \mathbf{C}_M^n p_1^n (1 - p_1)^{M-n}$$

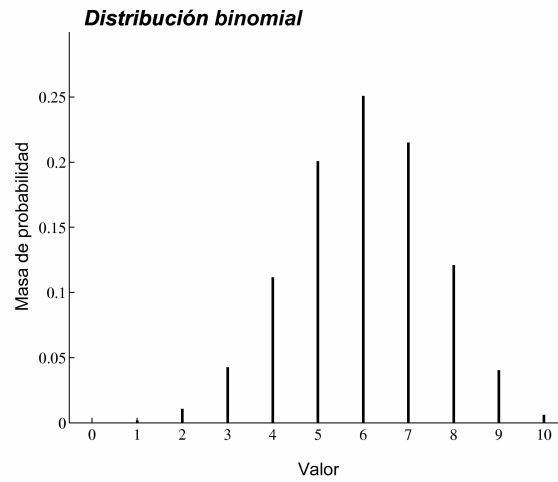


Figura 9. Masa de probabilidad binomial.

Distribución binomial negativa

La distribución binomial negativa depende de dos parámetros $\theta > 0$ y $\eta \in [0,1[$:

$$\forall n \in \mathbb{N}, p(n) = (1-\eta)^\theta \frac{\Gamma(\theta+n)}{\Gamma(\theta)} \frac{\eta^n}{n!}$$

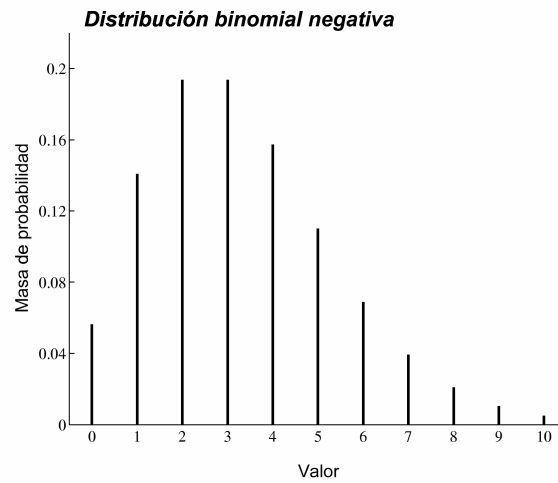


Figura 10. Masa de probabilidad binomial negativa.

Distribución de Poisson

La masa de probabilidad depende de un parámetro positivo θ igual tanto a la media como a la varianza:

$$\forall n \in \mathbb{N}, p(n) = \exp(-\theta) \frac{\theta^n}{n!}$$

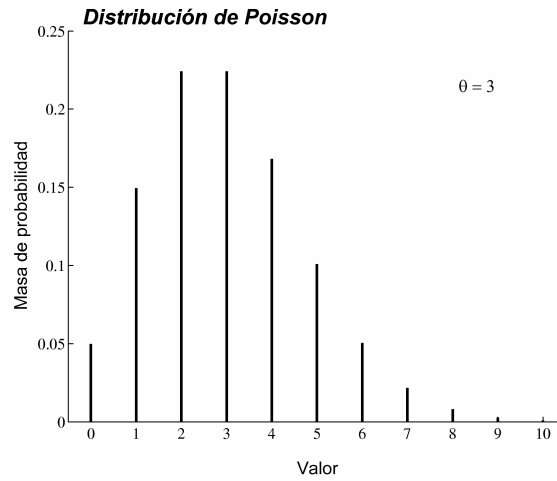


Figura 11. Masa de probabilidad de Poisson.

2.6. Teorema del límite central

Si una variable aleatoria X tiene una esperanza μ y una varianza σ^2 (ambas finitas), entonces

$$\frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \rightarrow N(0,1) \text{ si } n \rightarrow \infty$$

donde \bar{X} es la media experimental calculada en una muestra de X de tamaño n .

En otros términos, independientemente de la distribución inicial de X la distribución de la media experimental de una muestra de gran tamaño es normal. Usualmente, se considera que la convergencia se alcanza si $n > 50$.

3. Intervalos de confianza

Una aplicación directa de la teoría de probabilidades es la definición de intervalos de confianza para los parámetros de una población, dada la información de una muestra de tamaño n .

3.1. Intervalos de confianza para una proporción

Se considera una variable aleatoria de Bernoulli X , igual a 1 si se cumple una cierta propiedad A , 0 en caso contrario. Se tiene

$$\begin{cases} E(X) = \text{Prob}(X=1) = p_1 \\ \text{var}(X) = p_1(1-p_1) \end{cases}$$

Dada una muestra de tamaño n , se puede estimar la proporción (probabilidad) p_1 con la frecuencia empírica

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

En virtud del teorema del límite central, se tiene:

$$\frac{\bar{X} - p_1}{\sqrt{p_1(1-p_1)/n}} \approx N(0,1).$$

Utilizando las tablas de la distribución normal estándar:

$$\text{Prob}\left\{-z_{\alpha/2} < \frac{\bar{X} - p_1}{\sqrt{p_1(1-p_1)/n}} < z_{\alpha/2}\right\} = 1 - \alpha.$$

donde z_α es el valor tal que $P(z_\alpha) = 1 - \alpha$ (donde P es la función de distribución normal estándar). En particular, $z_{0.025} = 1.96$, valor clásicamente utilizado para determinar intervalos con un 95% de confianza. Invertiendo la relación anterior para despejar p_1 , se obtiene:

$$\text{Prob}\{p_{\min} < p_1 < p_{\max}\} = 1 - \alpha$$

$$\text{con } \begin{cases} p_{\min} = \frac{n}{n + z_{\alpha/2}^2} \left(\bar{X} + \frac{z_{\alpha/2}^2}{2n} - z_{\alpha/2} \sqrt{\frac{\bar{X}(1-\bar{X})}{n} + \frac{z_{\alpha/2}^2}{4n^2}} \right) \\ p_{\max} = \frac{n}{n + z_{\alpha/2}^2} \left(\bar{X} + \frac{z_{\alpha/2}^2}{2n} + z_{\alpha/2} \sqrt{\frac{\bar{X}(1-\bar{X})}{n} + \frac{z_{\alpha/2}^2}{4n^2}} \right) \end{cases}$$

Generalmente, se hace la siguiente aproximación:

$$p_{\min} = \bar{X} - z_{\alpha/2} \sqrt{\frac{\bar{X}(1-\bar{X})}{n}} \quad p_{\max} = \bar{X} + z_{\alpha/2} \sqrt{\frac{\bar{X}(1-\bar{X})}{n}}.$$

3.2. Intervalos de confianza para una esperanza

Supongamos que se conoce el valor de la varianza σ^2 de una población, pero no se tiene certeza sobre el valor de la esperanza μ . En virtud del teorema del límite central, se tiene:

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \approx N(0,1).$$

Utilizando las tablas de la distribución normal estándar, se obtiene un intervalo de confianza $1 - \alpha$:

$$\text{Prob}\left\{\bar{X} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right\} = 1 - \alpha.$$

En particular, μ tiene 95% de probabilidad de encontrarse entre $\bar{X} - 1.96\sigma/\sqrt{n}$ y $\bar{X} + 1.96\sigma/\sqrt{n}$.

Si se desconoce el valor de la varianza σ^2 , se buscará reemplazarlo por la varianza experimental S^2 . Para ello, se deberá utilizar la llamada **distribución de Student** en lugar de la distribución normal. Sea $N(0,1)$ una variable normal estándar (de esperanza 0 y varianza 1) y χ_{n-1}^2 una variable independiente del chi cuadrado con $n - 1$ grados de libertad. Se define la variable de Student con $n - 1$ grados de libertad (denotada T_{n-1}) como

$$T_{n-1} = N(0,1) \sqrt{\frac{n-1}{\chi_{n-1}^2}}.$$

Este resultado se puede aplicar al considerar

$$N(0,1) = \frac{\bar{X}-\mu}{\sigma} \sqrt{n} \quad \text{y} \quad \chi_{n-1}^2 = \frac{(n-1)S^2}{\sigma^2},$$

lo que da:

$$\frac{\bar{X}-\mu}{S/\sqrt{n}} = T_{n-1}.$$

Este resultado es independiente del valor de σ , por lo cual se puede usar aun cuando se desconoce σ . Con respecto al caso donde se conoce σ , basta con reemplazar σ por S y la distribución normal estándar por la distribución de Student de $n - 1$ grados de libertad:

$$\text{Prob}\left\{\bar{X} - t_{n-1,\alpha/2} \frac{S}{\sqrt{n}} < \mu < \bar{X} + t_{n-1,\alpha/2} \frac{S}{\sqrt{n}}\right\} = 1 - \alpha$$

donde $t_{n-1,\alpha}$ es el valor tal que $P(t_{n-1,\alpha}) = 1 - \alpha$ (donde P es la función de distribución de Student con $n - 1$ grados de libertad).

Capítulo 2: Errores en datos

1. Precisión y exactitud

La **precisión** mide la dispersión de una medición y puede expresarse bajo la forma de una desviación estándar o de una varianza. Una baja precisión implica incertidumbre y reduce la confianza que uno tiene en una medición.

La **exactitud** mide la desviación de la medición con respecto al resultado correcto. Mediciones inexactas implican la existencia de sesgos (errores sistemáticos), debidos a errores instrumentales, muestreos no representativos, equivocaciones, etc.

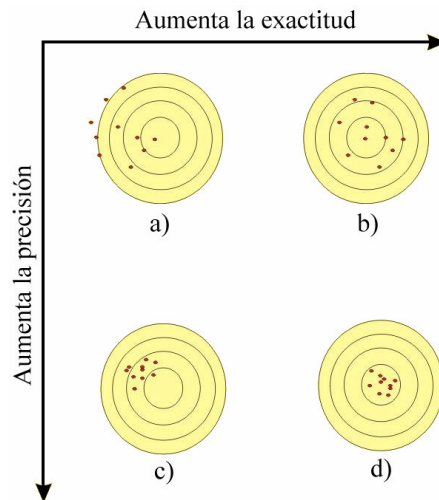


Figura 1. Ilustración de los conceptos de precisión y exactitud.

Toda medición posee incertidumbre, es decir, un *error*. Una estimación del error se puede llevar a cabo al replicar la medición (suponiendo que se mide la misma magnitud bajo las mismas condiciones) y calcular la desviación estándar o la varianza de los valores medidos.

2. Propagación de errores

En esta sección, interesa saber cómo los errores se propagan a través de los cálculos que uno es susceptible realizar.

2.1. Propagación a través de una suma o una diferencia

Si una variable z es la suma o la diferencia de dos variables (x,y) independientes cuyos errores son pequeños, entonces el error máximo en z es:

$$\delta z = \delta x + \delta y$$

En términos de varianzas de las mediciones, se tiene:

$$\sigma_z^2 = \sigma_x^2 + \sigma_y^2$$

Más generalmente, en el caso donde z es una combinación lineal de varias variables independientes:

$$z = \sum_{i=1}^n a_i x_i ,$$

se tendrá:

$$\sigma_z^2 = \sum_{i=1}^n a_i^2 \sigma_{x_i}^2 .$$

2.2. Propagación a través de una función cualquiera

La propagación de errores a través de una función puede ser determinada usando las derivadas parciales de esta función. Si $z = f(x,y)$, entonces la fórmula de Taylor permite escribir

$$\delta z \approx \frac{\partial f}{\partial x} \delta x + \frac{\partial f}{\partial y} \delta y ,$$

siempre que los errores $(\delta x, \delta y)$ sean pequeños. En términos de varianzas, se tiene:

$$\sigma_z^2 \approx \left(\frac{\partial f}{\partial x} \right)^2 \sigma_x^2 + \left(\frac{\partial f}{\partial y} \right)^2 \sigma_y^2 .$$

A modo de ejemplo, consideremos una función que es el producto de dos variables:

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

La fórmula anterior equivale a:

$$\frac{\sigma_{f(x,y)}^2}{f^2(x,y)} = \frac{\sigma_x^2}{x^2} + \frac{\sigma_y^2}{y^2} ,$$

es decir, la varianza *relativa* de $f(x,y)$ es igual a la sumatoria de las varianzas relativas de x y de y . Esta identidad queda válida si se considera el cuociente de las dos variables en lugar de su producto.

Más generalmente, para una función definida como el producto o cuociente de varias variables independientes, se tendrá

$$\frac{\sigma_{f(x,y,z,\dots)}^2}{f^2(x,y,z,\dots)} = \frac{\sigma_x^2}{x^2} + \frac{\sigma_y^2}{y^2} + \frac{\sigma_z^2}{z^2} + \dots$$

Como ejemplo de aplicación, si se tiene una estimación de los errores relativos en el tonelaje y en la ley media, se podrá estimar el error relativo en la cantidad total de fino.

Capítulo 3: Estadística comparativa

En este capítulo, revisaremos la teoría de los tests (o pruebas) de significancia para juzgar la relevancia de diferencias observadas en distintas muestras. El test se basa en la formulación de una hipótesis sobre parámetros de las muestras o sobre su distribución, y una prueba estadística de su validez.

1. Generalidades sobre los tests estadísticos

Sólo se testea una hipótesis con respecto a otras posibilidades. Es necesario precisar cuál es la hipótesis alternativa (H_1) con la cual se compara la hipótesis testada (llamada **hipótesis nula** y denotada como H_0). En importante destacar que la hipótesis nula es la que se privilegia, es decir, se supondrá válida salvo si existe una fuerte evidencia para probar lo contrario (en cuyo caso, se rechazará la hipótesis nula en favor de la hipótesis alternativa). Por ende, la elección de la hipótesis nula es de suma importancia, dado que los tests estadísticos siempre tienden a favorecer el status quo. Se distingue dos tipos de errores:

- **error de primera especie** (tipo I): rechazar la hipótesis nula cuando en realidad esta hipótesis es válida (ocurre con una probabilidad α);
- **error de segunda especie** (tipo II): aceptar la hipótesis nula cuando en realidad esta hipótesis es incorrecta (ocurre con una probabilidad β).

Los tests estadísticos permiten controlar los errores de tipo I. Cuando el umbral de significancia está excedido, se justifica rechazar la hipótesis nula. En el caso contrario, no se ha probado que la hipótesis nula sea la correcta; sólo se acepta por falta de prueba de que sea incorrecta.

Decisión	Realidad de la población	
	H_0 es correcta	H_0 es falsa
Se rechaza H_0	Error tipo I (probabilidad α)	Decisión correcta (probabilidad $1 - \beta$)
Se acepta H_0	Decisión correcta (probabilidad $1 - \alpha$)	Error tipo II (probabilidad β)

Tabla 1. Errores de primera y segunda especie.

La cantidad $1 - \alpha$ se llama la *eficiencia* del test, mientras que la cantidad $1 - \beta$ se llama la *potencia* del test.

Dado una muestra de tamaño n fijo, se busca definir una partición del espacio \mathbb{R}^n de las posibles realizaciones de la muestra, para poder optar por una u otra hipótesis, según los valores observados en la muestra. La partición debe ser tal que la *región crítica* o *región de rechazo* (para la cual se rechaza la hipótesis nula H_0) contenga la mayor parte de las realizaciones posibles de H_1 (hipótesis alternativa). En otras palabras, entre todas las particiones posibles, se busca aquella que maximiza la potencia $1 - \beta$ del test.

La construcción de un test se basa en las siguientes etapas:

- 1) Elegir las hipótesis H_0 y H_1 . La hipótesis que se privilegia es H_0 , mientras que H_1 es la hipótesis alternativa.
- 2) Determinar la variable de decisión.
- 3) Determinar la forma de la región crítica.
- 4) Calcular esta región crítica, según el valor escogido del riesgo de primera especie α .
- 5) Calcular la potencia $1 - \beta$ del test.
- 6) Hacer un experimento y medir el valor de la variable de decisión.
- 7) Concluir, aceptando o rechazando la hipótesis H_0 .

Ejemplo: control de calidad

Se compra miles de piezas con cierta característica. El control de la calidad se hace sobre una muestra, por ejemplo, 100 piezas sorteadas al azar. Debido a los aleas de fabricación, no se puede esperar que todas las piezas cumplen la característica deseada a cabalidad. Por lo tanto, se fija una proporción máxima de piezas defectuosas en toda la población, por ejemplo $p_0 = 10\%$. Dado que la muestra presentó una frecuencia f de piezas defectuosas, ¿es esta frecuencia compatible con la proporción p_0 ? o al contrario ¿es más acorde a una mayor proporción, $p_1 = 15\%$?

Una primera manera de responder a esta pregunta es fijar una frecuencia crítica, por ejemplo $p = 12\%$, y decidir rechazar la hipótesis nula ($H_0: p_0 = 10\%$) si la frecuencia observada supera 12% , para aceptar que la hipótesis alternativa ($H_1: p_1 = 15\%$) es más verosímil.

El rechazar equivocadamente la hipótesis H_0 (error de primera especie) es un riesgo para el fabricante, mientras que el aceptar equivocadamente la hipótesis H_0 es un riesgo para el cliente (error de segunda especie). Si el cliente desconfía del fabricante, elegiría testear la hipótesis H_1 ($p_1 = 15\%$). De este modo, descartaría esta hipótesis y aceptaría la hipótesis H_0 ($p_0 = 10\%$) sólo en base a una muestra muy probatoria.

Dado que en una población de N piezas, se sorteó n individuos sin reposición, el número de piezas defectuosas $n f$ sigue una distribución *hipergeométrica*, que depende de tres parámetros: el tamaño de la población (N), el tamaño de la muestra (n) y la verdadera proporción de piezas defectuosas (p). Si n/N es menor que 10%, se puede asimilar la distribución hipergeométrica con una *binomial*, con lo cual nos liberamos del parámetro N . Además, si n es suficientemente grande, la binomial se puede simplificar a una distribución normal (teorema del límite central):

$$f \approx p + \sqrt{\frac{p(1-p)}{n}} N(0,1).$$

- **Error de primera especie**

Bajo la hipótesis H_0 ($p = 10\% = 0.1$), se tiene:

$$\alpha = \text{Prob}\{f > 12\% | p = 0.1\} = \text{Prob}\{N(0,1) > 0.666\} = 0.253.$$

Una de cada cuatro veces, se rechazará una fabricación correcta.

- **Error de segunda especie**

Bajo la hipótesis H_1 ($p = 15\% = 0.15$), se tiene:

$$\beta = \text{Prob}\{f < 12\% | p = 0.15\} = \text{Prob}\{N(0,1) < -0.8402\} = 0.2005.$$

Una de cada cinco veces, se aceptará una fabricación defectuosa.

Al variar el criterio de decisión (valor de la frecuencia crítica), se puede disminuir α pero aumenta β : los dos riesgos son antagónicos.

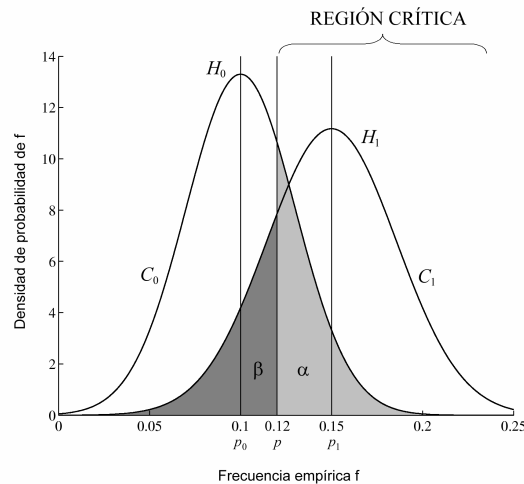


Figura 1. Antagonismos de los riesgos de primera y segunda especie.

Para cambiar los riesgos α y β , se puede cambiar el criterio de decisión (valor de la frecuencia crítica p) o aumentar el tamaño de la muestra (n):

$$\begin{cases} \alpha = \text{Prob} \left\{ N(0,1) > \frac{p - p_0}{\sqrt{p_0(1-p_0)/n}} \right\} \\ \beta = \text{Prob} \left\{ N(0,1) < \frac{p - p_1}{\sqrt{p_1(1-p_1)/n}} \right\} \end{cases}$$

Por ejemplo, si se quiere riesgos más pequeños ($\alpha = 0.05$, $\beta = 0.10$), se puede tomar los siguientes valores: $p = 0.126$ y $n = 362$.

2. Comparación de dos proporciones

Interesa comparar las proporciones de individuos que cumplen cierta característica en dos poblaciones. Sean p_1 y p_2 las proporciones reales (es decir, las probabilidades de ocurrencia) de la característica en las poblaciones, y f_1 y f_2 las frecuencias empíricas observadas en las dos muestras correspondientes.

La hipótesis nula y la hipótesis alternativa son:

$$H_0: p_1 = p_2 = p$$

$$H_1: p_1 \neq p_2$$

Se supondrá además que los tamaños de las muestras, n_1 y n_2 , son grandes, de modo que se podrá aplicar el teorema del límite central. En particular, se puede considerar que las frecuencias observadas tienen distribuciones normales:

$$f_1 \approx p_1 + \sqrt{\frac{p_1(1-p_1)}{n_1}} N(0,1)$$

$$f_2 \approx p_2 + \sqrt{\frac{p_2(1-p_2)}{n_2}} N(0,1)$$

Bajo la hipótesis H_0 , se tiene entonces:

$$f_1 - f_2 \approx \sqrt{p(1-p)} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} N(0,1).$$

Se rechaza H_0 si, para el riesgo α asumido, $|f_1 - f_2|$ supera un valor máximo dado por las tablas de la distribución normal.

3. Comparación de dos medias

3.1. Hipótesis

Se hace las siguientes suposiciones sobre los datos muestreados:

- su distribución es normal (Gaussiana);
- cada dato es una muestra independiente de la población a la cual pertenece;
- las características de las poblaciones muestreadas son constantes.

Ahora, los tests que utilizaremos a continuación (tests de Student) son **robustos**, por lo que se puede tolerar desviaciones con respecto a estos supuestos, en particular, que la distribución de los datos no sea exactamente normal. La justificación se debe a que se trabaja con medias experimentales, las cuales siguen una distribución normal aún si los datos no tienen tal distribución (teorema del límite central).

3.2. Poblaciones de misma varianza

Se busca ahora testear la significancia de la diferencia $\bar{X}_1 - \bar{X}_2$ entre las medias de dos muestras de tamaños n_1 y n_2 y desviaciones estándares s_1 y s_2 , respectivamente.

La hipótesis nula y la hipótesis alternativa son:

H_0 : las muestras provienen de poblaciones de misma media y misma varianza

H_1 : las poblaciones tienen medias distinta, aunque la misma varianza.

Se define la siguiente variable de decisión:

$$T = \frac{\bar{X}_1 - \bar{X}_2}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad \text{con} \quad s = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}.$$

Bajo la hipótesis nula, T tiene una distribución de Student con $n_1 + n_2 - 2$ grados de libertad. Luego, con las tablas de esta distribución, se decide si el valor observado de $|T|$ es significativamente alto, con un determinado riesgo de primera especie α . Cuando las muestras tienen el mismo tamaño ($n_1 = n_2 = n$), las fórmulas anteriores se simplifican:

$$T = \frac{\bar{X}_1 - \bar{X}_2}{s \sqrt{\frac{2}{n}}} \quad \text{con} \quad s = \sqrt{\frac{s_1^2 + s_2^2}{2}}.$$

3.3. Tests unilaterales o bilaterales?

Según la hipótesis alternativa elegida, se usará un test **unilateral** o **bilateral**.

- Hipótesis alternativa (H_1): las dos poblaciones tienen medias distintas ($\mu_1 \neq \mu_2$). En este caso, se usará un test bilateral (el valor de T puede ser significativo, cualquiera sea su signo) con un valor crítico $t_{n_1+n_2-2, \alpha/2}$.
- Hipótesis alternativa (H_1): $\mu_1 > \mu_2$. En este caso, se usará un test unilateral (sólo los valores positivos de T pueden ser significativos) con un valor crítico $t_{n_1+n_2-2, \alpha}$.
- Hipótesis alternativa (H_1): $\mu_1 < \mu_2$. En este caso también, se usará un test unilateral con un valor crítico $-t_{n_1+n_2-2, \alpha}$.

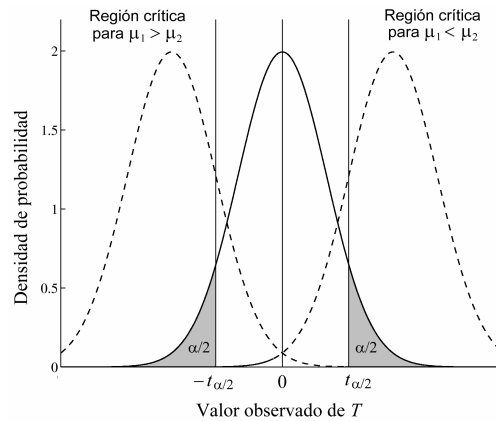


Figura 2. Tests unilaterales y bilaterales, según la hipótesis alternativa elegida.

3.4. Poblaciones de varianzas distintas

Cuando las varianzas de las dos poblaciones son distintas, el test debe modificarse de la siguiente forma. Bajo la hipótesis nula ($\mu_1 = \mu_2$), la variable

$$T = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

tiene una distribución aproximadamente de Student, cuyo número de grados de libertad es igual a

$$v = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}}.$$

3.5. Selección del tamaño de muestra

Al planificar un experimento comparativo, es importante decidir el tamaño de las muestras que se usarán en el test. Sean:

- n el tamaño de muestras (idéntico para ambas muestras)
- s la desviación estándar esperada
- δ la diferencia mínima que se quiere detectar
- z_α el valor normal (unilateral o bilateral, según la hipótesis alternativa) para un nivel de significancia $1 - \alpha$.

Para controlar el error de primera especie, se debe tomar

$$n \geq 2 \left(\frac{s z_\alpha}{\delta} \right)^2$$

El haber considerado el valor normal (z_α) para un nivel de significancia $1 - \alpha$ es una aproximación, porque se desconoce el número de grados de libertad de la distribución de Student (precisamente porque no se conoce el tamaño de muestra n). Para refinar la solución, una vez calculado el valor de n , se puede reemplazar z_α por el valor de Student con $2n - 2$ grados de libertad para el mismo nivel $1 - \alpha$, luego recalcular el valor de n . Se puede iterar este procedimiento hasta que no haya más cambios en el valor encontrado para n . En la práctica, utilizar el valor normal en lugar del valor de Student no tiene mucho impacto cuando n es mayor que 10.

Si se desea controlar los errores de primera (tipo I) y de segunda especie (tipo II), se debe tomar el siguiente valor de n :

$$n \geq 2 \left(\frac{s(z_\alpha + z_\beta)}{\delta} \right)^2$$

donde z_α es el valor para un nivel de significancia $1 - \alpha$ (error tipo I) y z_β es el valor para un nivel de significancia $1 - \beta$ (error tipo II) (el valor z_β siempre es unilateral).

Nuevamente, el haber considerado la distribución normal para definir los valores z_α y z_β constituye una aproximación. Como mencionado más arriba, se puede refinar la solución encontrada, utilizando un procedimiento iterativo. Otra opción consiste en:

1) calcular n con la fórmula anterior: $n_{aprox} = 2 \left(\frac{s(z_\alpha + z_\beta)}{\delta} \right)^2$

2) determinar el número de grados de libertad $f = 2(n_{aprox} - 1)$

3) calcular el valor corregido de n como $n = n_{aprox} \frac{f+3}{f+1}$.

Con el tamaño de muestra n así determinado, si el valor de T supera el valor crítico, entonces se rechaza la hipótesis nula con una probabilidad de error (tipo I) igual a α . Al contrario, si el valor crítico de T no está excedido, se acepta la hipótesis nula con una probabilidad β de equivocarse (error tipo II).

Ahora, como el tamaño de muestra suele ser grande para controlar ambos errores, es común tomar un valor de β mayor que α , por ejemplo $\alpha = 0.05$ y $\beta = 0.20$. En la prueba de nuevos procedimientos, esta elección privilegia el *status quo* (mayor probabilidad β de aceptar la hipótesis nula cuando en realidad existe una diferencia entre el nuevo procedimiento y el antiguo).

3.6. Muestras pareadas

Se busca comparar dos métodos, dos tratamientos o dos aparatos, que se aplican a cantidades no constantes, por ejemplo leyes en datos duplicados en un protocolo de muestreo. En este caso, es apropiado testear la significancia de la *diferencia* promedio entre pares de valores.

Se define la variable de diferencia $D = X_1 - X_2$ y se plantea

$$T = \frac{\bar{D}}{s/\sqrt{n}},$$

donde s es la desviación estándar de D . Bajo la hipótesis nula (D tiene una esperanza nula), T tiene una distribución de Student con $n - 1$ grados de libertad. Por lo tanto, se rechaza la hipótesis nula si el valor observado de T supera el valor crítico para el nivel de significancia deseado.

Este test de comparación de muestras pareadas conduce a mejores resultados que la comparación de dos muestras, cuando las diferencias entre pares de datos son importantes y cuando las dos series de valores están correlacionadas. Al contrario, si las diferencias entre pares son pequeñas, el test puede ser menos potente que la comparación de las dos muestras. También es peligroso utilizar este test cuando no existe un acoplamiento natural de las dos muestras, y que el usuario define los pares arbitrariamente.

Para la selección de tamaño de muestra, se tiene los mismos resultados que con el test de Student de comparación de dos medias, salvo que se debe omitir el factor 2 en las fórmulas que proporcionan el tamaño de muestra. Para controlar el error de tipo I, se debe tomar

$$n \geq \left(\frac{s z_{\alpha}}{\delta} \right)^2$$

mientras que para controlar los errores de tipo I y II, se debe tomar $n \geq \left(\frac{s(z_{\alpha} + z_{\beta})}{\delta} \right)^2$.

3.7. Comparación de una media con un valor predeterminado

Se busca determinar si la media μ de una población (de la cual se tiene una muestra de tamaño n , media \bar{X} y desviación estándar s) es distinta a un valor predeterminado μ_0 . Para ello, se plantea:

$$T = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}.$$

Bajo la hipótesis nula ($\mu = \mu_0$), T tiene una distribución de Student con $n - 1$ grados de libertad. Se rechaza la hipótesis nula si el valor observado de $|T|$ supera el valor crítico para el nivel de significancia deseado.

4. Otros tests usuales

4.1. Comparación de dos varianzas

Se busca comparar las varianzas S_1^2 y S_2^2 de dos muestras normales independientes de tamaños n_1 y n_2 . Se ordena las muestras de manera que $S_1^2 \geq S_2^2$. Se define

$$F = \frac{S_1^2}{S_2^2}$$

Bajo la hipótesis nula (H_0 : las muestras provienen de poblaciones de distribuciones normales con la misma varianza), F tiene una distribución de Fisher con $n_1 - 1$ y $n_2 - 1$ grados de libertad. Luego se compara el valor de F con el valor crítico para el nivel de significancia deseado (test bilateral o unilateral dependiendo de la hipótesis alternativa). Es importante recalcar que este test es válido solamente cuando las muestras tienen distribuciones normales y no es robusto frente a una violación de esta hipótesis.

4.2. Ajuste de distribuciones

Supongamos que se tiene un conjunto de datos clasificados en k clases o categorías exclusivas. El **test del chi cuadrado** permite determinar si los números de ocurrencias observados en cada clase (O_i , $i = 1 \dots k$) son significativamente distintos de los números esperados (E_i , $i = 1 \dots k$) bajo un determinado modelo de distribución. Si el modelo de distribución postulado es correcto, la variable

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

sigue una distribución del chi cuadrado con $k - 1$ grados de libertad. Consecuentemente, se rechazará el modelo si el valor de χ^2 supera el valor crítico para el riesgo asumido.

El test del chi cuadrado es *libre*, es decir, independiente del modelo de distribución propuesto. Idealmente, se requiere al menos 5 individuos por clase; en caso contrario se agrupan varias clases para cumplir esta condición. Si, además de la distribución, se estima ℓ parámetros (media, varianza...), entonces el chi cuadrado tiene $k - 1 - \ell$ grados de libertad en lugar de $k - 1$.

Se puede extender el test para ajustar una distribución de probabilidad bivariable. Se considera un conjunto de datos clasificados según dos variables categóricas (A y B) y resumidas en una tabla de contingencia de r columnas y s filas (Tabla 2).

	A_1	...	A_i	...	A_r	total
B_1	n_{11}		n_{i1}		n_{r1}	$n_{\bullet 1}$
...						
B_j	n_{1j}		n_{ij}		n_{rj}	$n_{\bullet j}$
...						
B_s	n_{1s}		n_{is}		n_{rs}	$n_{\bullet s}$
total	$n_{1\bullet}$		$n_{i\bullet}$		$n_{r\bullet}$	n

Tabla 2. Tabla de contingencia para dos variables categóricas.

Se denota como $(O_{ij}, i = 1 \dots r, j = 1 \dots s)$ los números de ocurrencias observados en cada clase y $(E_{ij}, i = 1 \dots r, j = 1 \dots s)$ los números esperados bajo un determinado modelo de distribución. Si el modelo es correcto, la variable

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^s (O_{ij} - E_{ij})^2 / E_{ij}$$

tiene una distribución del chi cuadrado con $(r - 1)(s - 1)$ grados de libertad.

Existen otros enfoques para probar el ajuste de distribuciones:

- Visualización del histograma de los datos. Para un histograma simétrico, se podrá considerar un modelo de distribución normal o de Student, mientras que para un histograma asimétrico se podrá buscar un modelo lognormal, gamma, chi cuadrado, exponencial o de Weibull.
- Ajuste gráfico utilizando *gráficos cuantiles contra cuantiles* (visualizan los cuantiles de la distribución empírica de datos contra los cuantiles de una distribución teórica), *gráficos de probabilidad normal* o *lognormal* (visualizan el histograma acumulado de los datos, con ejes de escalas tales que el gráfico dibuja una recta en caso de tener una distribución normal o lognormal) (Figura 3).
- Test libres (de Kolmogorov o de Cramer von Mises) que comparan la distribución acumulada empírica con una distribución acumulada teórica.

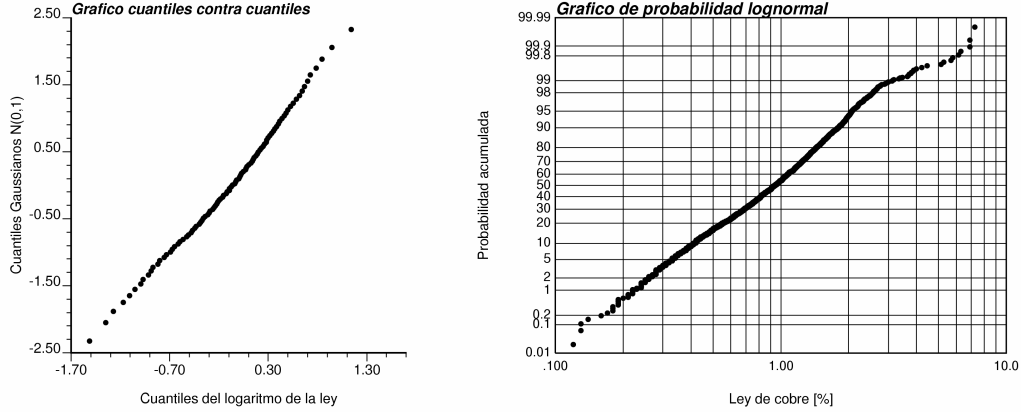


Figura 3. Gráfico cuantiles contra cuantiles y gráfico de probabilidad lognormal para datos de ley de cobre de distribución aproximadamente lognormal.

4.3. Tests de outliers

Un **outlier** es un dato atípico o inusual. Cabe preguntarse si se puede o no eliminar este dato atípico. La respuesta más simple y conservadora es la negativa, a menos que uno encuentra una explicación física del outlier (error instrumental, error de transcripción del dato, muestra contaminada, etc.). Ahora, es posible diseñar un test estadístico para saber si este dato proviene de la misma población que los demás, luego para corroborar la decisión de remover o no el outlier. En todo caso, el conocimiento extra-estadístico y el juicio del usuario es primordial.

A continuación, se presenta el *test de Grubbs* (o test del residuo de máxima norma), el cual se basa en la hipótesis de que los datos provienen de una distribución normal. Dada una muestra de tamaño n , media experimental \bar{X} y varianza experimental s^2 , se considera la siguiente variable:

$$G = \frac{\max(|X - \bar{X}|)}{s}$$

y se compara su valor con un valor crítico que depende del tamaño de la muestra y del riesgo asumido. La hipótesis H_0 de no tener outlier se rechaza si

$$G > \frac{n-1}{\sqrt{n}} \sqrt{\frac{t_{n-2, \alpha/2n}^2}{n-2 + t_{n-2, \alpha/2n}^2}}$$

donde $t_{n, \alpha}$ representa el valor de la distribución de Student con n grados de libertad que corresponde a una probabilidad acumulada $1 - \alpha$.

Capítulo 4: Análisis de varianza

El análisis de varianza (*ANOVA*, por su acrónimo inglés) permite determinar si las medias de varias muestras son significativamente distintas una de otra. La idea básica es dividir el total de la varianza de los datos entre varios componentes (varianza inter-muestra o explicada y varianza intra-muestra o residual) y comparar estos componentes utilizando un test de Fisher.

Se supone que las muestras tienen una distribución normal, de misma varianza, pero posiblemente de diferentes medias. Si no se verifica la hipótesis de normalidad, se debe buscar otros métodos, como el *test de Kruskal-Wallis*.

1. Análisis de varianza para diseños simples

Supongamos que, en un experimento donde se comparan k tratamientos (realizando una muestra por tratamiento), existen dos fuentes de variaciones: el tratamiento mismo y el error de medición. La variación debida al tratamiento (*variación explicada* o inter-muestra) se puede medir por:

$$SS_1 = \sum_{i=1}^k n_i (\bar{X}_i - \bar{X})^2$$

con \bar{X} : media global de todos los datos

\bar{X}_i : media de los datos asociados al tratamiento n° i

n_i : número de datos asociados al tratamiento n° i .

La variación debida al error (variación *residual* o intra-muestra) se mide por

$$SS_0 = \sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2$$

donde X_{ij} es el j -ésimo dato asociado al tratamiento n° i .

La propiedad clave que utiliza el análisis de varianza es la **aditividad de las sumas de cuadrados**:

$$SS = \sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X})^2 = SS_0 + SS_1.$$

En lugar de trabajar con sumas, se prefiere considerar las medias de cuadrados (las cuales tienen sentido de varianza). Para ello, se divide cada suma de cuadrados por el número de grados de libertad correspondiente:

$$MS_1 = \frac{SS_1}{k-1} \quad MS_0 = \frac{SS_0}{n-k} \quad MS = \frac{SS_0 + SS_1}{n-1}$$

donde $n = n_1 + \dots + n_k$ es el número total de datos.

El test de Fisher permite entonces decidir si la media de cuadrados asociada a los diferentes tratamientos es significativamente mayor que aquella asociada a los errores. Bajo la hipótesis nula (las medias de todos los tratamientos son iguales: $\mu_1 = \dots = \mu_k$), se tiene que:

$$F = \frac{MS_1}{MS_0} = \frac{SS_1/(k-1)}{SS_0/(n-k)}$$

es una variable aleatoria de Fisher con $k-1$ y $n-k$ grados de libertad. Por lo tanto, si el valor observado de F es mayor que el valor crítico para el riesgo asumido (valor del test unilateral), se rechazará la hipótesis nula.

Los resultados de análisis de varianza suelen ser presentados en una tabla, como la tabla mostrada a continuación.

Fuente de variación	Suma de cuadrados	Grados de libertad	Media de cuadrados	F
Tipo de mineral (inter-muestra)				
Error (intra-muestra)				
Total				

Tabla 1. Tabla de análisis de varianza

Ejemplo: se considera los siguientes datos ($k = 6$)

tratamiento	1	2	3	4	5	6
número	12	13	12	13	13	12
promedio	7.25	6.62	6.33	4.08	4.23	8.17

Tabla 2. Números y promedios de datos por tratamiento

En el total de los 6 tratamientos, se ha observado $n=75, \bar{X}=6.16$ y $s^2=10.65$.

Se llena la tabla de análisis de varianza:

fuelle de variación	Suma de cuadrados	grados de libertad	Media de cuadrados
explicada	170	5	170/5 = 34
residual	618	69	618/69 = 8.96
total	788	74	

Tabla 3. Tabla de análisis de varianza para los datos de la Tabla 2

Luego, $F = 34/8.96 = 3.79$. Se rechazará la hipótesis nula, dado que $F(5,69) = 2.35$ para $\alpha = 5\%$. En conclusión, existen diferencias significativas entre los tratamientos.

2. Análisis de varianza para diseños anidados

2.1. Diseños anidados

Se busca diseñar un esquema de análisis para tener la mayor precisión al menor costo en determinados procedimientos o ensayos de laboratorio. Para este análisis, la información necesaria se obtiene al considerar un diseño anidado, en el cual las fuentes de variaciones son independientes. Por ejemplo:

- considerar varios lotes, muestrear varias veces cada lote y analizar varias veces cada muestra (esquema anidado: *lotes > muestras > análisis*)
- considerar distintas fechas, operadores e instrumentos. En cada fecha y para cada operador, se realiza una medición con cada instrumento (esquema anidado: *fechas > operadores > instrumentos*).

Se desea atribuir una varianza a cada fuente (ej: fecha, operador, instrumento). La varianza de una medición se descompone como:

$$V_1 = \sigma_{\text{fecha}}^2 + \sigma_{\text{operador}}^2 + \sigma_{\text{instrumento}}^2 .$$

Para reducir el error, se puede repetir la medición y promediar los resultados. A modo de ejemplo, si cada operador realiza dos análisis por instrumento, se reducirá el error instrumental:

$$V_2 = \sigma_{\text{fecha}}^2 + \sigma_{\text{operador}}^2 + \sigma_{\text{instrumento}}^2 / 2 .$$

En cambio, al promediar las mediciones obtenidas por dos operadores que usan instrumentos distintos en fechas distintas, se obtendría una reducción de la varianza en todas las componentes:

$$V_3 = (\sigma_{\text{fecha}}^2 + \sigma_{\text{operador}}^2 + \sigma_{\text{instrumento}}^2) / 2.$$

Si el objetivo de los análisis es comparar dos instrumentos distintos, es mejor pedir que cada operador haga mediciones con los dos instrumentos en la misma fecha. En este caso, la varianza de la *diferencia* entre las dos mediciones efectuadas por cada operador sería

$$2\sigma_{\text{instrumento}}^2$$

y se cancelarían los errores correspondientes a fechas y operadores. Al contrario, si dos operadores distintos hicieran mediciones en fechas distintas, la varianza de la diferencia entre las mediciones sería mayor y aumentaría el riesgo de equivocar la conclusión:

$$2(\sigma_{\text{fecha}}^2 + \sigma_{\text{operador}}^2 + \sigma_{\text{instrumento}}^2).$$

2.2. Tabla de análisis de varianza

Consideremos un diseño anidado con a clases asociadas a una fuente de variación A . En cada una de estas clases, se tiene b sub-clases asociadas a una fuente de variación B , y así sucesivamente (c sub-clases de una fuente C y d sub-clases de una fuente D). Finalmente, cada sub-clase de D tiene n mediciones (réplicas). El número total de datos es $N = abcdn$ (nota: se puede extender el método presentado a números variables de ítems en cada clase). Basándose en la aditividad de las sumas de cuadrados, se obtiene la siguiente tabla de análisis de varianza.

Fuente	Suma de cuadrados	Grados de libertad	Media de cuadrados
A	$SS_A = \sum_{i=1}^a bcdn (\bar{X}_i - \bar{X})^2$	$a - 1$	$MS_A = \frac{SS_A}{a-1}$
B	$SS_B = \sum_{i=1}^a \sum_{j=1}^b cdn (\bar{X}_{ij} - \bar{X}_i)^2$	$a(b - 1)$	$MS_B = \frac{SS_B}{a(b-1)}$
C	$SS_C = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^c dn (\bar{X}_{ijk} - \bar{X}_{ij})^2$	$ab(c - 1)$	$MS_C = \frac{SS_C}{ab(c-1)}$
D	$SS_D = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^c \sum_{l=1}^d n (\bar{X}_{ijkl} - \bar{X}_{ijk})^2$	$abc(d - 1)$	$MS_D = \frac{SS_D}{abc(d-1)}$
Error	$SS_E = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^c \sum_{l=1}^d \sum_{m=1}^n (X_{ijklm} - \bar{X}_{ijkl})^2$	$abcd(n - 1)$	$MS_E = \frac{SS_E}{abcd(n-1)}$
Total	$SS = SS_A + SS_B + SS_C + SS_D + SS_E$	$N - 1$	$MS = \frac{SS}{N-1}$

Tabla 4. Tabla de análisis de varianza para un diseño anidado.

La significancia de cada media de cuadrados se testea al formar una variable de Fisher F definida como la razón entre esta media de cuadrados y la media siguiente en la tabla.

Ejemplo (ensayos de laboratorio).

Se considera 2 lotes, en cada uno de los cuales se elige 3 muestras y se realizan 3 análisis por muestra. Los resultados se indican en la siguiente tabla.

Muestra	Lote 1	Lote 2
1	8.0, 7.4, 7.8	6.2, 7.4, 6.9
2	7.7, 7.3, 5.1	5.7, 5.8, 5.2
3	5.8, 5.6, 5.7	5.4, 6.1, 4.6
Media	6.711	5.922

Tabla 5. Datos de ensayos de laboratorio

La tabla de análisis de varianza se presenta a continuación.

Fuente de variación	Suma de cuadrados	Grados de libertad	Media de cuadrados	F
Lote	2.8006	1	2.8006	1.120
Muestra	9.9978	4	2.4994	4.848
Análisis	6.1867	12	0.5156	
Total	18.9850	17		

Tabla 6. Tabla de análisis de varianza para los datos de la Tabla 5

La media de cuadrados de las muestras es significativamente mayor que la media de cuadrados de los análisis ($F = 4.848$ es significativo). Luego, para optimizar el diseño sin aumentar el presupuesto total, sería preferible tomar más muestras y hacer menos análisis.

Capítulo 5: Diseño de experimentos

1. Generalidades sobre el diseño de experimentos

1.1. Definiciones y pasos preliminares

Un experimento se puede definir como una prueba o una serie de pruebas donde se cambia a propósito el valor de variables de entrada de un proceso o sistema, de manera de observar e identificar los cambios en las variables de salida.

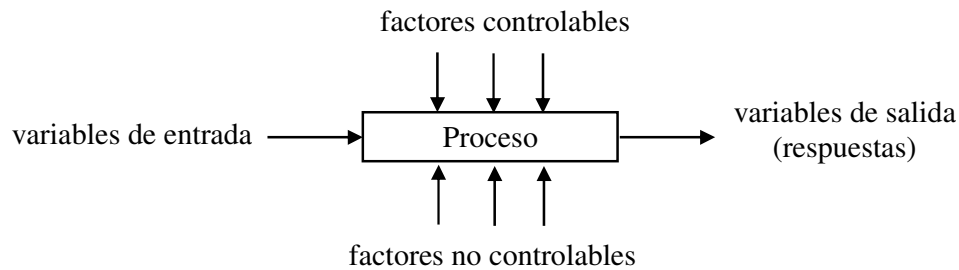


Figura 1. Esquema general de un experimento

Se define a la **respuesta** o **variable dependiente** como la variable que se investiga y que, en general, se busca optimizar. Un **factor** o **variable independiente** es una variable que el experimentador puede hacer variar para medir su efecto en la respuesta. Puede ser *cualitativo* o *cuantitativo*. El valor de un factor en una prueba particular se conoce como su *nivel*. Un **tratamiento** es una combinación de niveles de factores, y una **prueba** una operación del experimento con un tratamiento.

Un experimento se articula en torno a los siguientes pasos:

- 1) Definir los objetivos específicos del experimento. Por ejemplo:
 - Determinar qué factor tiene mayor influencia sobre la respuesta.
 - Determinar para qué valores de los factores la variabilidad de la respuesta es pequeña.
 - Determinar los valores de los factores controlables que minimizan los efectos de los factores no controlables.
- 2) Formular los objetivos en términos experimentales. ¿Qué experimentos se requiere hacer para lograr los objetivos?

- 3) ¿Qué variables de respuesta y factores se requiere seleccionar? ¿Qué se conoce del error experimental (o sea, de las variaciones aleatorias de las respuestas)? Conducir pruebas a escala piloto si es necesario.
- 4) ¿Sobre qué rangos se debe alterar los factores? ¿Interesa conocer la interacción entre los factores? ¿Existe una o varias variable(s) no deseada(s) o no controlada(s) cuyos efectos deban ser eliminados?
- 5) Acorde a lo anterior, al tiempo y presupuesto disponible y al efecto posible sobre los procesos en operación, seleccionar un diseño experimental apropiado.

1.2. Conceptos claves en el diseño de experimento

Aleatorización

La asignación del material experimental y el orden en que se desarrollan las pruebas se eligen al azar. Esto permite que las observaciones (o los errores experimentales) puedan ser consideradas como variables aleatorias independientes, y mitigar los efectos que puedan tener factores externos no controlables.

Replicación

Se trata de una repetición *independiente* de un mismo tratamiento (combinación de factores). Permite tener una idea del error experimental y determinar si las diferencias observadas en los datos son estadísticamente significativas. También permite tener mejores estimaciones de los parámetros (por ejemplo, tomando el valor promedio observado en las réplicas).

Rebloqueo

Permite mejorar la precisión con la cual se realiza la comparación entre los factores de interés, al reducir la variabilidad provocada por factores que no son de interés.

1.3. Guía general para diseñar experimentos

La primera etapa consiste en **reconocer y plantear el problema** a investigar. Para ello, se recomienda trabajar en equipo multidisciplinario. Se debe plantear las *preguntas* que ha de responder el experimento y los *objetivos* del mismo, por ejemplo:

- diseñar un nuevo sistema y caracterizarlo;
- optimizar un sistema existente;
- confirmar un comportamiento observado en el pasado;
- descubrir lo que ocurre bajo nuevas condiciones de operación;
- determinar la robustez de la respuesta a cambios en los factores.

En lugar de diseñar un experimento complejo para responder a múltiples preguntas y objetivos, una estrategia común es realizar una serie de experimentos pequeños, cada uno asociado a un objetivo específico. Se habla entonces de experimentos iterativos o secuenciales.

La segunda etapa en el diseño consiste en **seleccionar una variable de respuesta**. A menudo, se trata de la media o de la desviación estándar (o ambas) de la característica medida. Si el error de medición es muy importante, se puede tomar varias réplicas y utilizar la media de las mediciones replicadas como variable de respuesta.

Luego, se debe **elegir los factores**, sus niveles y rangos. En esta etapa, se puede distinguir *factores potenciales de diseño* (factores que el experimentador puede y quiere variar durante el experimento, o factores cuyos niveles se dejan constantes) y *factores de molestia* que no son de interés (estos últimos pueden ser controlables o no). Se puede buscar factores relacionados con la medición, el material, los operadores, el entorno, los métodos y las máquinas. Para decidir los rangos y niveles de los factores, se requiere cierto conocimiento del proceso (tanto experiencia práctica como conocimiento teórico).

Posteriormente, se **selecciona un diseño experimental**. Hay que definir el tamaño de la muestra (número de réplicas), el orden de las pruebas experimentales, y determinar si considerar un rebloqueo u otras restricciones. Es aconsejable elegir el diseño lo más sencillo posible.

En la **realización del experimento**, se debe monitorear el proceso para asegurarse que todo se haga como planeado. Se puede hacer unos ensayos o pruebas piloto antes de realizar el experimento, para verificar el sistema de medición, tener una idea del error experimental y practicar la técnica de experimento.

La etapa siguiente corresponde al **análisis estadístico de los datos**, el cual permite tomar decisiones y conclusiones objetivas. Se podrá recurrir a interpretaciones gráficas, pruebas (tests) de hipótesis o modelos experimentales (por ejemplo, ajustes de curvas de regresión).

Finalmente, se emiten las **conclusiones y recomendaciones**. Es aconsejable utilizar su conocimiento no estadístico del problema y tener presente que una significancia *estadística* no necesariamente implica que exista una significancia *práctica* (aunque sea estadísticamente probada, una diferencia en una respuesta puede ser irrelevante del punto de vista práctico). No olvidar documentar el experimento y los datos obtenidos.

2. Diseño de bloques aleatorios

El efecto de *un* factor no deseado (por ejemplo: el operador o factor humano; el tiempo; el cambio de maquinaria) puede ser eliminado al dividir las pruebas en distintos grupos o **bloques**, cada uno de los cuales presenta un nivel del factor no deseado. Si el orden de los tratamientos en cada bloque se elige al azar, entonces el experimento se conoce como **experimento de bloques aleatorios**. El análisis de tal experimento permite separar los efectos de cambios de tratamiento y de cambios de bloque.

Por ejemplo, supongamos que un experimento se lleva a cabo durante seis días consecutivos e interesa medir las respuestas a un factor con tres posibles niveles (A, B, C). Podemos definir cada día como un bloque distinto y considerar el siguiente orden de tratamientos:

Día 1: CAB
Día 2: BAC
Día 3: BCA
Día 4: ABC
Día 5: CBA
Día 6: ACB

El diseño de bloque aleatorio se puede resumir en una tabla de datos, como aquella mostrada a continuación. El número de tratamiento (en las filas) se denota f , mientras que el número de bloques (en las columnas) se denota como c . El número total de datos es $n = fc$.

	Bloque 1	Bloque 2	...	Bloque c	Media
Tratamiento 1	X_{11}	X_{12}	...	X_{1c}	$\bar{X}_{1\bullet}$
Tratamiento 2	X_{21}	X_{22}	...	X_{2c}	$\bar{X}_{2\bullet}$
...					
Tratamiento f	X_{f1}	X_{f2}	...	X_{fc}	$\bar{X}_{f\bullet}$
Media	$\bar{X}_{\bullet 1}$	$\bar{X}_{\bullet 2}$		$\bar{X}_{\bullet c}$	\bar{X}

Tabla 1. Tabla de datos para un diseño de bloques aleatorios

La herramienta clave para determinar si existe algún efecto de los bloques o de los tratamientos es el **análisis de varianza**. En este contexto, se define la siguiente tabla.

Fuente de variación	Suma de cuadrados	Grados de libertad	Media de cuadrados
Entre filas	$SS_f = \sum_{i=1}^f c(\bar{X}_{i\bullet} - \bar{X})^2$	$f - 1$	$MS_f = \frac{SS_f}{f-1}$
Entre columnas	$SS_c = \sum_{j=1}^c f(\bar{X}_{\bullet j} - \bar{X})^2$	$c - 1$	$MS_c = \frac{SS_c}{c-1}$
Error	$SS_e = \sum_{i=1}^f \sum_{j=1}^c (X_{ij} - \bar{X}_{i\bullet})(X_{ij} - \bar{X}_{\bullet j})$	$(f - 1)(c - 1)$	$MS_e = \frac{SS_e}{(f-1)(c-1)}$
Total	$SS = SS_f + SS_c + SS_e$	$n - 1$	

Tabla 2. Tabla de análisis de varianza para un diseño de bloques aleatorios

Una vez construida la tabla de análisis de varianza, se puede comparar MS_f y MS_e , así como MS_c y MS_e , mediante tests de Fisher, para determinar si las diferencias son significativas.

Ejemplo 1: Diseño de bloques aleatorios sin réplicas

Se considera los datos de la Tabla 3. Se desea saber si las diferencias entre bloques y/o entre tratamientos son significativas.

	Bloque 1	Bloque 2	Bloque 3	Bloque 4	Bloque 5
Tratamiento 1	6325	7020	5896	6619	6123
Tratamiento 2	7186	7361	6132	7106	6250
Tratamiento 3	6643	6511	5928	6532	5938

Tabla 3. Datos obtenidos de un diseño de bloques aleatorios

Se obtiene la siguiente tabla de análisis de varianza.

Fuente de variación	Suma de cuadrados	Grados de libertad	Media de cuadrados
Entre tratamientos	704 117	2	352 058
Entre bloques	2 245 035	4	561 259
Error	314 351	8	39 294
Total	3 263 503	14	

Tabla 4. Tabla de análisis de varianza para los datos de la Tabla 3

La media de cuadrados entre tratamientos aparece significativamente más alta que la media de cuadrados de errores (el cociente entre ambas medias es de 8.96, mientras que el valor crítico de la distribución de Fisher de 2 y 8 grados de libertad para un nivel de confianza de 99% es igual a 8.65). Asimismo, la media de cuadrados entre bloques también es significativamente más alta que la media de cuadrados de errores (cociente igual a 14.28, en comparación con un valor crítico igual a 7.01 para una distribución de Fisher de 4 y 8 grados de libertad para un nivel de confianza de 99%). En conclusión, las variaciones debidas tanto a los bloques como a los tratamientos son significativas. Ambas variaciones son del mismo orden de magnitud (el cociente entre las medias de cuadrados de tratamientos y de bloques, 1.59, no es significativo para una distribución de Fisher de 4 y 2 grados de libertad).

Ejemplo 2: Diseño con tratamientos replicados

Si cada tratamiento dentro de cada bloque está replicado, se puede obtener una mejor estimación del error experimental y mejorar la comparación entre tratamientos, así como entre tratamientos y bloques. El análisis de varianza se lleva a cabo de manera similar, pero es posible extraer una fuente adicional de variación debida a la interacción entre tratamientos y bloques.

Sea f el número de tratamientos, c el número de bloques, r el número de réplicas y $n = fcr$ el número total de datos. Para $i = 1 \dots f$, $j = 1 \dots c$ y $k = 1 \dots r$, denotaremos como X_{ijk} el dato correspondiente al tratamiento $n^\circ i$, bloque $n^\circ j$ y réplica $n^\circ k$. La tabla de análisis de varianza se indica a continuación.

Fuente de variación	Suma de cuadrados	Grados de libertad	Media de cuadrados
Entre tratamientos	$SS_f = r \sum_{i=1}^f c (\bar{X}_{i..} - \bar{X})^2$	$f - 1$	$MS_f = \frac{SS_f}{f-1}$
Entre bloques	$SS_c = r \sum_{j=1}^c f (\bar{X}_{.j.} - \bar{X})^2$	$c - 1$	$MS_c = \frac{SS_c}{c-1}$
Interacción bloques / tratamientos	$SS_{fc} = r \sum_{i=1}^f \sum_{j=1}^c (\bar{X}_{ij.} - \bar{X}_{i..} - \bar{X}_{.j.} + \bar{X})^2$	$(f-1)(c-1)$	$MS_{fc} = \frac{SS_{fc}}{(f-1)(c-1)}$
Error	$SS_e = \sum_{i=1}^f \sum_{j=1}^c \sum_{k=1}^r (X_{ijk} - \bar{X}_{ij.})^2$	$fc(r-1)$	$MS_e = \frac{SS_e}{fc(r-1)}$
Total	$SS = SS_f + SS_c + SS_{fc} + SS_e$	$n - 1$	

Tabla 5. Tabla de análisis de varianza para un diseño de bloques aleatorios con réplicas

Por ejemplo, consideremos la prueba de dos tratamientos (A y B) en tres bloques, con dos replicas para cada prueba.

	Bloques / replicas					
tratamiento	1/1	1/2	2/1	2/2	3/1	3/2
A	7.3	8.3	7.1	6.3	4.4	4.3
B	8.0	8.5	9.6	7.0	7.0	6.9

Tabla 6. Datos obtenidos de un diseño de bloques aleatorios con réplicas

Fuente de variación	Suma de cuadrados	Grados de libertad	Media de cuadrados
Entre bloques	12.452	2	$MS_1 = 6.226$
Entre tratamientos	7.208	1	$MS_2 = 7.208$
Interacción bloques / tratamientos	2.314	2	$MS_3 = 1.157$
Error	4.335	6	$MS_4 = 0.723$
Total	26.309	11	

Tabla 7. Tabla de análisis de varianza para los datos de la Tabla 6

Se puede comentar lo siguiente:

- $MS_3 / MS_4 = 1.60$, lo cual no es significativo para una variable de Fisher de 2 y 6 grados de libertad. Por ende, se asume que no existe interacción entre tratamientos y bloques, y se combina ambos valores MS_3 y MS_4 para tener una mejor estimación de la varianza del error experimental:

$$MS_{\text{combinado}} = \frac{2.315 + 4.335}{6 + 2} = 0.831.$$

El error experimental en la medición se puede cuantificar por la desviación estándar

$$\sqrt{MS_{\text{combinado}}} = 0.912.$$

- $MS_1 / MS_{\text{combinado}} = 7.49$, lo cual es significativo para una variable de Fisher de 2 y 8 grados de libertad. El cambio de bloque tiene un efecto real sobre la medición.
- De igual manera, $MS_2 / MS_{\text{combinado}} = 8.67$ es significativo para una variable de Fisher de 1 y 8 grados de libertad. El cambio de tratamiento tiene un efecto sobre la medición (de la tabla 6, se ve que el tratamiento B proporciona sistemáticamente valores mayores, en promedio de 1.55).

3. Diseño de cuadrado latino

Supongamos ahora que existen *dos* factores no deseados (por ejemplo, el operador y el día en el que se hace el experimento) que no tienen interacción. El llamado *diseño de cuadrado latino* consiste en probar cada tratamiento una vez con cada operador y una vez por día. Debe existir igual número (p) de tratamientos, operadores y días.

Día	Operador				
	1	2	3	4	5
1	A	B	C	D	E
2	B	C	D	E	A
3	C	D	E	A	B
4	D	E	A	B	C
5	E	A	B	C	D

Tabla 8. Ejemplo de diseño de cuadrado latino

Aquí, el análisis de varianza consiste en dividir la suma total de cuadrados en cuatro componentes (filas, columnas, tratamientos y errores):

$$SS = SS_f + SS_c + SS_t + SS_e$$

con los respectivos grados de libertad

$$p^2 - 1 = (p-1) + (p-1) + (p-1) + (p-2)(p-1).$$

Una vez construida la tabla de análisis de varianza se compara las diferentes medias de cuadrados (MS_t / MS_e), (MS_f / MS_e) y (MS_c / MS_e) para determinar si las diferencias son significativas.

4. Diseño factorial

4.1. Estrategias posibles

Cuando existen varios factores en un mismo experimento, se puede considerar las siguientes estrategias:

- **Mejor apuesta:** se hace una serie de pruebas, cambiando el valor de un solo factor entre una prueba y la siguiente y esperando observar una mejora en la respuesta. Este procedimiento no garantiza llegar a la combinación óptima de los factores.
- **Un factor a la vez:** se parte de una combinación de niveles de factores y se realiza varias pruebas, en cada una de las cuales se modifica el valor de *un solo* factor con respecto a la combinación de partida. Este procedimiento no permite detectar interacciones entre factores (un factor puede no producir el mismo efecto en la respuesta según el valor de otros factores) y entrega generalmente soluciones no óptimas.
- **Todos los factores a la vez:** se conduce un experimento llamado *factorial*, donde los niveles de los factores se modifican conjuntamente. Este tipo de experimento es el que hace el uso más eficiente de los datos.

4.2. Experimento factorial

Se dice que un experimento es **factorial** cuando se elige un conjunto de niveles para los factores, y una o varias pruebas se realizan con cada posible combinación de cada factor. Si se tiene n factores con el mismo número de niveles (x) y cada combinación se repite r veces, entonces el número total de pruebas necesarias es $r x^n$. Típicamente, se toma $x = 2$ (un nivel bajo y un nivel alto para cada factor) o $x = 3$, y el orden de las pruebas se elige al azar.

Las ventajas de realizar experimentos factoriales son:

- Determinar interacciones entre factores.
- Si no existen tales interacciones, se tiene la máxima eficiencia para estimar los efectos propios de los factores.
- Las conclusiones son válidas en un amplio rango de situaciones.

Por ejemplo, consideremos dos factores (A y B), cada uno con dos niveles (– y +). Al graficar la respuesta en función de los niveles del factor A, se puede determinar si existe o no interacción entre A y B (Figura 2). En el caso general, el análisis de varianza permite separar y cuantificar los efectos de los factores y los efectos de interacción entre factores.

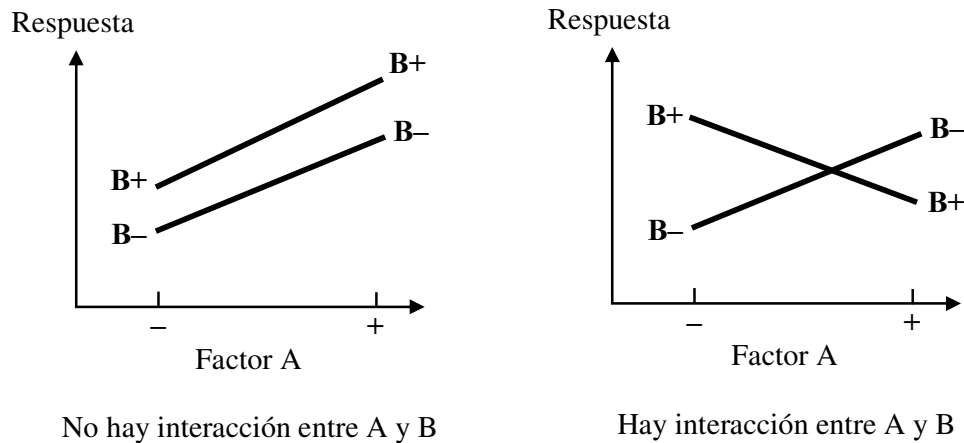


Figura 2. Determinación gráfica de la posible interacción entre los factores A y B

4.3. Experimento de dos factores

El modelo matemático para un experimento de dos factores (A,B) es:

$$Y_{ijk} = \bar{Y} + A_i + B_j + D_{ij} + \epsilon_{ijk}$$

con

- Y_{ijk} : respuesta de la k -ésima réplica con A en el nivel i y B en el nivel j .

- A_i : valor promedio de todas las pruebas con A en el nivel i
- B_j : valor promedio de todas las pruebas con B en el nivel j
- D_{ij} : componente de interacción
- ε_{ijk} : componente de error.

Se supone $i \in \{1 \dots a\}$, $j \in \{1 \dots b\}$ y $k \in \{1 \dots r\}$. La tabla de análisis de varianza tiene la siguiente forma:

Fuente de variación	Suma de cuadrados	Grados de libertad	Media de cuadrados
Efecto de A	$SS_A = br \sum_{i=1}^a (\bar{Y}_i - \bar{Y})^2$	$a - 1$	$MS_A = \frac{SS_A}{a-1}$
Efecto de B	$SS_B = ar \sum_{j=1}^b (\bar{Y}_j - \bar{Y})^2$	$b - 1$	$MS_B = \frac{SS_B}{b-1}$
Interacción AB	$SS_{AB} = r \sum_{i=1}^a \sum_{j=1}^b (\bar{Y}_{ij} - \bar{Y}_i - \bar{Y}_j + \bar{Y})^2$	$(a-1)(b-1)$	$MS_{AB} = \frac{SS_{AB}}{(a-1)(b-1)}$
Error	$SS_E = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^r (Y_{ijk} - \bar{Y}_{ij})^2$	$ab(r-1)$	$MS_E = \frac{SS_E}{ab(r-1)}$
Total	$SS = \sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^r (Y_{ijk} - \bar{Y})^2$	$abr - 1$	

Tabla 9. Tabla de análisis de varianza para un diseño de dos factores

Ejemplo 1: experimento con réplicas

Se considera un experimento con $r = 2$ réplicas, $n = 2$ factores (A y B), $x = 2$ niveles (0 = bajo y 1 = alto). Los resultados de las mediciones se indican a continuación.

	A ₀	A ₁
B ₀	23.2; 22.5	26.1; 27.3
B ₁	22.8; 21.1	23.9; 24.2

Tabla 10. Tabla de datos para un diseño de dos factores

Se busca determinar si A y/o B tienen efectos significativos, y si la interacción entre ambos factores es significativa. Se obtiene la tabla de análisis de varianza señalada a continuación.

Fuente de variación	Suma de cuadrados	Grados de libertad	Media de cuadrados
Efecto de A	17.70	1	17.70
Efecto de B	6.30	1	6.30
Interacción AB	1.53	1	1.53
Error	2.46	4	0.62
Total	27.99	7	

Tabla 11. Tabla de análisis de varianza para los datos de la Tabla 10

El valor crítico para una variable de Fisher de 1 y 4 grados de libertad y un nivel de confianza de 95% es de 7.71. Por lo tanto, los efectos de A y de B son significativos (las medias de cuadrados divididas por la media de cuadrados de error son iguales a 28.78 y 10.24, respectivamente). De la tabla 10, se puede calcular el efecto promedio de A:

$$\frac{26.1+27.3+23.9+24.2}{4} - \frac{23.2+22.5+22.8+21.1}{4} = 2.98$$

y el efecto promedio de B:

$$\frac{22.8+21.1+23.9+24.2}{4} - \frac{23.2+22.5+26.1+27.3}{4} = -1.78.$$

En cambio, la interacción entre A y B no tiene un efecto significativo, puesto que el cociente de la media de cuadrados (1.53) por la media de cuadrados de errores (0.62) es inferior al valor crítico 7.71.

Ejemplo 2: experimento sin réplicas

Si no hay réplica ($r = 1$), no se puede calcular el error residual SS_E . En este caso, se puede evaluar dicho error:

- haciendo r' pruebas con valores intermedios de los factores (en cuyo caso, el número de grados de libertad para el valor estimado de SSE será $r' - 1$);
- suponiendo que las componentes de interacción no son significativas, por lo cual se identifica estas componentes con el error residual;
- utilizando datos históricos.

A modo de ilustración, retomemos los datos de la Tabla 10, considerando solamente la primera réplica de cada experimento (en total, se tiene 4 datos con valores 23.2, 26.1, 22.8 y 23.9). Se llega a la siguiente tabla de análisis de varianza.

Fuente de variación	Suma de cuadrados	Grados de libertad	Media de cuadrados
Efecto de A	4.00	1	4.00
Efecto de B	1.69	1	1.69
Interacción AB	0.81	1	0.81
Error	?		
Total	6.50	3	

Tabla 12. Tabla de análisis de varianza para el experimento sin réplica

Se requiere ahora estimar la media de cuadrados (varianza) del error experimental. Por ejemplo, se puede realizar tres pruebas con niveles intermedios de los factores A y B, obteniendo los siguientes datos: 24.1, 25.0 y 24.5. Se tendrá entonces una varianza de 0.203, con $3 - 1 = 2$ grados de libertad. En este ejemplo, el efecto de A es juzgado significativo ($4.00/0.203 = 19.7$, valor significativo para una distribución de Fisher de 1 y 2 grados de libertad), mientras que el efecto de B y la interacción entre A y B no son significativos. Se podría combinar las medias de cuadrados de B y de AB para tener una mejor estimación de la varianza del error experimental.

4.4. Experimento de tres factores

Para un experimento de 3 factores (A,B,C), se tendría la siguiente tabla.

Fuente de variación	Suma de cuadrados	Grados de libertad	Media de cuadrados
A	SS_A	$a - 1$	MS_A
B	SS_B	$b - 1$	MS_B
C	SS_C	$c - 1$	MS_C
AB	SS_{AP}	$(a - 1)(b - 1)$	MS_{AP}
AC	SS_{AC}	$(a - 1)(c - 1)$	MS_{AC}
BC	SS_{BC}	$(b - 1)(c - 1)$	MS_{BC}
ABC	SS_{ABC}	$(a - 1)(b - 1)(c - 1)$	MS_{ABC}
Error	SS_E	$abc(r - 1)$	MS_E
Total	SS	$abc r - 1$	MS

Tabla 13. Tabla de análisis de varianza para un diseño de tres factores

Por ejemplo, se puede volver a tomar los datos de la Tabla 10 y, esta vez, considerar las réplicas como un tercer factor (C).

Fuente de variación	Suma de cuadrados	Grados de libertad	Media de cuadrados
A	17.70	1	17.70
B	6.30	1	6.30
C	0.10	1	0.10
AB	1.53	1	1.53
AC	1.90	1	1.90
BC	0.45	1	0.45
ABC	0.001	1	0.001
Error	?		
Total	27.98	7	

Tabla 14. Tabla de análisis de varianza para los datos de la Tabla 10

Dado que no hay réplicas, para estimar la varianza del error, se podría nuevamente hacer pruebas con valores intermedios para los factores. Una alternativa es considerar el valor de la interacción de mayor orden (en este caso, ABC) como primera estimación de la varianza (media de cuadrados) del error experimental. Luego, se puede mejorar esta estimación al combinar progresivamente las otras fuentes de variación que no son declaradas significativas.

En este ejemplo, tomando un nivel de confianza de 95%:

- Empezamos con una varianza de error estimada en 0.001 con 1 grado de libertad (fuente ABC).
- La variación debida al factor C (0.10) no es significativa para una distribución de Fisher con 1 y 1 grados de libertad. Se puede entonces incluir la variación debida a este factor y a sus interacciones (AC, BC) para estimar mejor la varianza de error:

$$(0.10 + 1.90 + 0.45 + 0.001) / 4 = 0.6128.$$

- Las variaciones debidas a A y a B son significativas para una distribución de Fisher con 1 y 4 grados de libertad (las medias de cuadrados correspondientes divididas por la varianza de error dan 28.90 y 10.29, respectivamente, y son mayores que el valor crítico 7.71). En cambio, la variación debida a la interacción AB no es significativa.

4.5. Experimentos fraccionales

Cuando el número de factores aumenta, el número de pruebas a realizar para tener un experimento factorial completo se vuelve demasiado grande para el experimentador. Por ejemplo, considerando 6 factores en dos niveles, se necesitaría 64 pruebas (sin hacer réplicas). En este caso, se tendría 63 grados de libertad, de los cuales 6 están asociados a

los efectos principales de los factores (A, B, C, D, E, F) y 15 a las interacciones de segundo orden entre factores (AB, AC, AD..., EF), mientras que 42 grados de libertad están asociados a interacciones de tercer orden o superior (ABC...).

Si el experimentador asume que las interacciones de orden alto son despreciables, la información sobre los efectos principales y sobre las interacciones de bajo orden puede ser obtenida con solamente una fracción del experimento factorial completo. En este caso, se habla de *diseños fraccionales*.

Un uso común de este tipo de diseño consiste en experimentos en que se consideran muchos factores y el objetivo es determinar qué factores tienen el mayor efecto. Esto se puede utilizar al principio de un proyecto, cuando el efecto de muchos de los factores inicialmente considerados es susceptible de ser nulo o muy poco. Los factores que serán identificados como importantes podrán luego ser investigados más detalladamente en otros experimentos.

A modo de ejemplo, considerando tres factores (A,B,C) y dos niveles, existen dos diseños de medio, los cuales pueden ser representados geoméricamente como en la Figura 3. En este tipo de diseño, no se puede diferenciar el efecto principal del factor A del efecto de la interacción entre B y C (similarmente, no se puede diferenciar los efectos de B y AC, así como los efectos de C y AB).

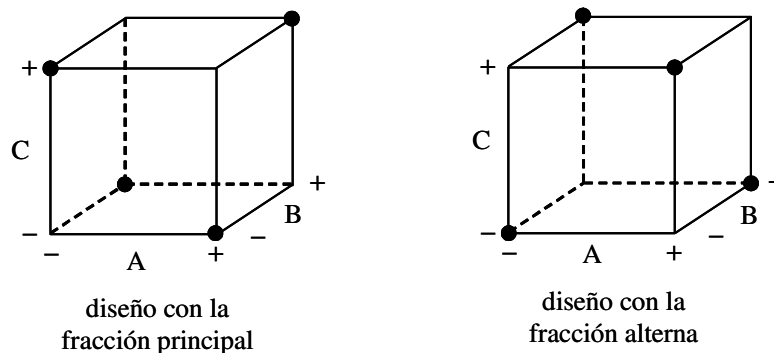


Figura 3. Representación geométrica de diseños de fracción de medio (los puntos corresponden a las pruebas seleccionadas)

5. Diseño de pruebas industriales

Interesa saber si un cambio en un proceso o en las condiciones de operación trae un beneficio. A menudo, la prueba a escala real es la única manera de determinar con confianza si tal beneficio existe (mejoras a nivel de laboratorio o escala piloto pueden desvanecerse a escala industrial). Ahora, las pruebas industriales son caras y pueden reducir la eficiencia de los procesos, por lo que se buscará llegar a una conclusión firme en el menor tiempo posible.

Se debe tener presente que varios factores son susceptibles alterar los resultados, en particular el factor tiempo (las condiciones operativas suelen cambiar en el tiempo). Muchas réplicas de la prueba son necesarias para mitigar los efectos de estos factores, lo

que puede ser obtenido al utilizar periodos de tiempo fijos para las réplicas (típicamente, periodos de un día).

Existen tres grandes enfoques para conducir una prueba industrial:

- 1) Comparando respuestas en pares de datos (en particular, cuando se quiere comparar dos tipos de condiciones experimentales), en cuyo caso se puede utilizar el test de Student para muestras pareadas.
- 2) Utilizando un diseño de bloques aleatorios, en donde el efecto de una variable no deseada (por ejemplo, el tiempo) se elimina al considerar bloques.
- 3) Sin un diseño formal, por ejemplo, comparando las respuestas en forma grafica para llegar a una conclusión. Este enfoque requiere generalmente tener más datos que en los dos enfoques anteriores, ya que no hay un control formal del error experimental. Se debe tomar más datos cuando las correlaciones entre las variables de entrada y de respuesta son bajas. Asimismo, se debe suponer que estas correlaciones no cambian significativamente durante el periodo de prueba.

5.1. Test de Student para muestras pareadas

La ventaja de este test es que permite eliminar las desviaciones debidas a errores, al considerar la *diferencia* entre dos respuestas (típicamente, la respuesta bajo condiciones normales de operación y la respuesta bajo nuevas condiciones). Para evitar sesgos, se recomienda alternar / variar los tratamientos, por ejemplo:

Día	Condición	Condición	Diferencia
1	normal	nueva	normal – nueva
2	nueva	normal	normal – nueva
3	nueva	normal	normal – nueva
4	normal	nueva	normal – nueva

A veces, al cambiar las condiciones de operación de la prueba, es necesario incluir un periodo de “estabilización” antes de poder obtener una respuesta confiable, por lo que se ignora la respuesta en el periodo de estabilización. Se puede entonces considerar grupos de tres o cuatro periodos, por ejemplo:

Día	Secuencia de condiciones
1	AABB
2	BAA
3	ABB
4	AABB
5	BAA
6	BBAA

En negrita, los pares utilizados para el test de Student.

5.2. Diseño de bloques aleatorios

Para experimentos con dos condiciones y sin réplicas, este enfoque proporciona los mismos resultados que el test de Student de muestras pareadas. Sin embargo, el diseño de bloques aleatorios es apropiado cuando se busca comparar más de dos condiciones. El factor tiempo es comúnmente la variable agrupada en bloques.

Es indispensable que cada tratamiento sea probado al menos una vez en cada bloque y que el orden de los tratamientos sea elegido al azar. Si estas dos condiciones no se cumplen, el análisis de varianza dará resultados inválidos.

Capítulo 6: Métodos de mínimos cuadrados

1. Modelamiento de datos

Un *modelo matemático* permite describir y resumir un conjunto de datos medidos y sus relaciones. Por ejemplo, puede tratarse de una clase de funciones (ej.: polinomios) cuyos coeficientes están ajustados a los datos. Modelos más complejos son obtenidos al considerar leyes físicas que rigen el comportamiento de los datos. A veces, uno combina los dos enfoques anteriores (*modelos mecánicos*), de tal modo que la estructura del modelo proviene de mecanismos o leyes físicas que se piensa son importantes, mientras que los detalles de estos mecanismos son descritos por parámetros ajustados de manera de reproducir los datos medidos. Estos modelos son generalmente más confiables que un simple ajuste de los datos para extrapolar los resultados a nuevas situaciones o nuevas condiciones operatorias.

Se puede clasificar los modelos matemáticos entre

- *modelos determinísticos*: el mismo input (entrada) genera el mismo output (salida);
- *modelos probabilísticos*: el output intenta reflejar la naturaleza aleatoria del proceso. Un modelo determinístico se puede convertir en un modelo probabilístico al agregar un error aleatorio al output. Ahora, para recurrir a un modelo probabilístico, es necesario asegurarse de la “aleatoriedad” de las mediciones. Visualizar el valor de la medición en función de su orden (fecha en que se toma la medición) puede revelar la presencia de problemas, como por ejemplo, tendencias sistemáticas (valor de la medición que disminuye o que aumenta con el tiempo).

Asimismo, los modelos (incluso determinísticos) recurren a la estadística, dado que se basan en datos medidos y muy raramente reproducen exactamente estos datos debido a errores de medición. Por lo tanto, se necesitan métodos para evaluar si el modelo es apropiado, es decir, testearlo acorde a estándares estadísticos.

La construcción de modelos es un proceso creativo que depende de la experiencia del modelador, sus preferencias y sus intuiciones, así como del objetivo para el cual se construye el modelo.

2. Principios básicos de los métodos de mínimos cuadrados

Se dispone de N **datos medidos** $\{y_i, i = 1 \dots N\}$ (vistos como variables aleatorias, puesto que existen errores de medición) con desviaciones estándares $\{\sigma_i, i = 1 \dots N\}$. Estas desviaciones miden la precisión de los datos y pueden estimarse si se dispone de réplicas en las mediciones. Se desea ajustar estos datos de modo de encontrar el “mejor” estimador $\{\hat{y}_i, i = 1 \dots N\}$ de los valores reales.

Supongamos también que se tiene **L restricciones** (por ejemplo, ecuaciones del modelo predictivo que queremos construir):

$$\forall j \in \{1, \dots, L\}, f_j(\hat{y}_1, \dots, \hat{y}_N; a_1, \dots, a_M) = 0$$

con **M parámetros** $\{a_k, k = 1 \dots M\}$ que se busca determinar junto con $\{\hat{y}_i, i = 1 \dots N\}$.

Una manera de ajustar los datos es minimizar la siguiente cantidad:

$$F = \sum_{i=1}^N \varepsilon_i^2 \quad \text{con} \quad \varepsilon_i = \frac{y_i - \hat{y}_i}{\sigma_i}$$

bajo las L restricciones anteriores.

La motivación de este procedimiento es la siguiente: si uno supone que los errores $\{\varepsilon_i, i = 1 \dots N\}$ son variables aleatorias normales, se puede establecer que minimizar F equivale a maximizar la *verosimilitud* de los parámetros $\{a_1, \dots, a_M\}$ y $\{\hat{y}_1, \dots, \hat{y}_N\}$, o sea, encontrar los parámetros que maximizan la probabilidad de haber encontrado los valores medidos $\{y_1, \dots, y_N\}$.

Ahora, cada uno de los N términos en la suma de cuadrados que define F aporta un grado de libertad, mientras que cada parámetro libre (ya sea \hat{y}_i o a_k) resta un grado de libertad. Asimismo, cada ecuación de restricción agrega un grado de libertad, puesto que elimina a un parámetro libre. El número total de grados de libertad es:

$$\underbrace{N}_{\text{datos}} - \underbrace{(M+N)}_{\text{parámetros}} + \underbrace{L}_{\text{restricciones}} = L - M.$$

Por lo tanto, cuando está en su valor mínimo (y suponiendo que la distribución de los errores es normal), F es una variable del chi cuadrado con $L - M$ grados de libertad. Se puede utilizar este resultado para medir la calidad del ajuste, dado que las tablas de la distribución del chi cuadrado indican la probabilidad de superar un determinado valor para F . A modo de ejemplo, si la probabilidad Q de superar el valor obtenido de F es muy pequeña, esto significa que el ajuste no es satisfactorio:

- el modelo es “malo” y puede ser rechazado;
- los valores de desviación estándar $\{\sigma_i, i = 1 \dots N\}$ están subestimados;
- los errores $\{\varepsilon_i, i = 1 \dots N\}$ no tienen una distribución normal. Usualmente, el tener errores no normales aumenta la ocurrencia de valores extremos (outliers) que hacen bajar la probabilidad Q , por lo que se puede aceptar valores bajos de Q (0.001).

Para verificar que los errores de medición $\{\varepsilon_i, i = 1 \dots N\}$ (residuos de la regresión) tienen una distribución normal, se puede aplicar un gráfico de probabilidad normal con estos residuos. Si se invalida la hipótesis de normalidad, el ajuste de mínimos cuadrados sigue proporcionando un modelo para determinar la respuesta esperada (valor de y).

Pero ya no se puede definir intervalos de confianza sobre los parámetros del ajuste o hacer pruebas (tests) de hipótesis.

Ejemplo

Se busca ajustar un modelo con una sola restricción:

$$\forall i \in \{1, \dots, N\}, \hat{y}_i = a,$$

bajo el supuesto que las desviaciones estándares $\{\sigma_i, i = 1 \dots N\}$ son todas iguales a σ . Se busca minimizar

$$\chi^2 = \sum_{i=1}^N \left(\frac{y_i - a}{\sigma} \right)^2.$$

El mínimo se obtiene al anular la derivada parcial de χ^2 con respecto al parámetro a :

$$\frac{\partial \chi^2}{\partial a} = -2 \sum_{i=1}^N \frac{y_i - a}{\sigma^2} = 0$$

o sea:

$$a = \frac{1}{N} \sum_{i=1}^N y_i = \bar{y}.$$

Es decir, el mejor ajuste de los datos $\{y_i, i = 1 \dots N\}$ por una constante corresponde a la media aritmética de estos datos (de haber considerado desviaciones σ_i variables, se hubiese encontrado una media ponderada de los datos). La calidad del ajuste se puede evaluar al calcular la suma de residuos cuadráticos

$$\chi^2 = \sum_{i=1}^N \left(\frac{y_i - \bar{y}}{\sigma} \right)^2$$

y comparar su valor con el valor crítico para una variable del chi cuadrado con $N - 1$ grados de libertad.

3. Ajuste de un modelo lineal de una variable

3.1. Ajuste de una recta que pasa por el origen

Buscamos ahora un modelo de la siguiente forma:

$$\forall i \in \{1, \dots, N\}, \hat{y}_i = b x_i$$

donde:

- se dispone de mediciones $\{y_i, i = 1 \dots N\}$ con desviaciones $\{\sigma_i, i = 1 \dots N\}$
- cada medición y_i está asociada a un valor x_i (supuesto sin error de medición).

En este ejemplo, existen tantas restricciones como datos ($L = N$) y un solo parámetro ($M = 1$), por lo que el número de grados de libertad es $N - 1$. La minimización de los residuos cuadráticos conduce a:

$$b = \frac{\sum_{i=1}^N \frac{x_i y_i}{\sigma_i^2}}{\sum_{i=1}^N \frac{x_i^2}{\sigma_i^2}}.$$

Como casos particulares, se tiene:

- $\sigma_i = \sigma$ constante: $b = \sum_{i=1}^N x_i y_i / \sum_{i=1}^N x_i^2$
- σ_i^2 proporcional a x_i : se obtiene una pendiente b igual al cociente \bar{y}/\bar{x}
- σ_i proporcional a x_i : la pendiente b es igual al promedio de las pendientes que se obtendrían al considerar cada dato individualmente:

$$b = \frac{1}{N} \sum_{i=1}^N \frac{y_i}{x_i}.$$

3.2. Modelo lineal de una variable

3.2.1. Desviaciones variables y conocidas

Consideremos la expresión general de un modelo lineal de una variable:

$$\forall i \in \{1, \dots, N\}, \hat{y}_i = a + b x_i$$

donde:

- se dispone de mediciones $\{y_i, i = 1 \dots N\}$ con desviaciones $\{\sigma_i, i = 1 \dots N\}$
- cada medición y_i está asociada a un valor x_i (supuesto sin error de medición).

Existen el mismo número de restricciones que de datos ($L = N$) y dos parámetros ($M = 2$), por lo que el número de grados de libertad es $N - 2$. Se busca minimizar

$$\chi^2 = \sum_{i=1}^N \left(\frac{y_i - a - b x_i}{\sigma_i} \right)^2.$$

El mínimo se obtiene al anular las derivadas parciales de χ^2 con respecto a ambos parámetros:

$$\begin{cases} \frac{\partial \chi^2}{\partial a} = -2 \sum_{i=1}^N \frac{y_i - a - bx_i}{\sigma_i^2} = 0 \\ \frac{\partial \chi^2}{\partial b} = -2 \sum_{i=1}^N \frac{x_i (y_i - a - bx_i)}{\sigma_i^2} = 0 \end{cases}$$

La solución es:

$$\begin{cases} a = \frac{S_{xx}S_y - S_xS_{xy}}{SS_{xx} - S_x^2} \\ b = \frac{SS_{xy} - S_xS_y}{SS_{xx} - S_x^2} \end{cases} \quad \text{con} \quad \begin{cases} S = \sum_{i=1}^N \frac{1}{\sigma_i^2} \\ S_x = \sum_{i=1}^N \frac{x_i}{\sigma_i^2} \\ S_{xx} = \sum_{i=1}^N \frac{x_i^2}{\sigma_i^2} \\ S_y = \sum_{i=1}^N \frac{y_i}{\sigma_i^2} \\ S_{xy} = \sum_{i=1}^N \frac{x_i y_i}{\sigma_i^2} \end{cases}$$

Alternativamente, se puede escribir

$$\begin{cases} b = \frac{\sum_{i=1}^N \frac{1}{\sigma_i^2} (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^N \frac{1}{\sigma_i^2} (x_i - \bar{x})^2} \\ a = \bar{y} - b\bar{x} \end{cases}$$

con

$$\begin{cases} \bar{x} = \frac{S_x}{S} = \sum_{i=1}^N \frac{x_i}{\sigma_i^2} / \sum_{i=1}^N \frac{1}{\sigma_i^2} \quad (\text{media de } x) \\ \bar{y} = \frac{S_y}{S} = \sum_{i=1}^N \frac{y_i}{\sigma_i^2} / \sum_{i=1}^N \frac{1}{\sigma_i^2} \quad (\text{media de } y) \end{cases}$$

Suponiendo que los y_i son independientes y utilizando la fórmula de propagación de errores y las expresiones de a y b en función de los y_i , se obtiene:

$$\begin{aligned} \text{var}\{a\} &= \frac{S_{xx}}{SS_{xx} - S_x^2} & \text{var}\{b\} &= \frac{S}{SS_{xx} - S_x^2} \\ \text{cov}\{a, b\} &= -\frac{S_x}{SS_{xx} - S_x^2} & \text{corr}\{a, b\} &= -\frac{S_x}{\sqrt{SS_{xx}}} \end{aligned}$$

Estos resultados permiten determinar intervalos de confianza para los parámetros “reales” (α y β) de la recta de regresión en torno a los valores obtenidos (a y b), o para testear hipótesis sobre sus valores (por ejemplo, determinar si la pendiente real β puede

ser igual a 0). Para ello, se supondrá que a y b tienen distribuciones normales, de medias α y β y varianzas dadas por las fórmulas anteriores. Por ejemplo, se se consideran las siguientes hipótesis nula y alternativa (*test de significancia de la regresión*):

$$\begin{aligned} H_0: \beta &= 0 \\ H_1: \beta &\neq 0 \end{aligned}$$

se examinará si el valor de $(b-0)/\sqrt{\text{var}(b)}$ es compatible con una normal $N(0,1)$.

3.2.2. Desviaciones conocidas sólo con un factor de proporcionalidad

Se supone que las desviaciones $\{\sigma_i, i = 1 \dots N\}$ sólo se conocen con un factor de proporcionalidad:

$$\forall i \in \{1, \dots, N\}, \sigma_i = \sigma r_i \text{ con } \sigma \text{ desconocido.}$$

Por ejemplo, si y_i es la media de n_i mediciones y que todas las mediciones tienen la misma precisión, se puede plantear

$$r_i = 1/\sqrt{n_i}.$$

La cantidad a minimizar se puede escribir como:

$$F = \sum_{i=1}^N \left(\frac{y_i - a - b x_i}{r_i} \right)^2.$$

por lo que se obtiene la solución (a y b) de la misma forma que anteriormente, salvo que se debe **reemplazar σ_i por r_i** . Pero el valor de F no permite evaluar la calidad del ajuste (F no tiene una distribución del chi cuadrado puesto que no es la suma de variables $N(0,1)$ al cuadrado; solamente F/σ^2 tiene una distribución del chi cuadrado, pero se ignora el valor de σ).

Con el reemplazo de σ_i por r_i , las fórmulas previas permiten calcular las varianzas y covarianzas *relativas* de a y b :

$$\begin{aligned} \frac{\text{var}\{a\}}{\sigma^2} &= \frac{S_{xx}}{SS_{xx} - S_x^2} & \frac{\text{var}\{b\}}{\sigma^2} &= \frac{S}{SS_{xx} - S_x^2} \\ \frac{\text{cov}\{a,b\}}{\sigma^2} &= -\frac{S_x}{SS_{xx} - S_x^2} & \text{corr}\{a,b\} &= -\frac{S_x}{\sqrt{SS_{xx}}} \end{aligned}$$

Si N es grande, la distribución de F/σ^2 (chi cuadrado) es aproximadamente normal de media igual al número de grados de libertad ($N - 2$). Luego, se puede estimar el valor desconocido de σ al plantear

$$\hat{\sigma} = \sqrt{\frac{F}{N-2}}.$$

Nuevamente, se puede determinar intervalos de confianza para los valores reales de la recta de regresión en torno a los valores obtenidos (a, b) , y testear hipótesis sobre estos valores reales. Para ello, basta con reemplazar σ por $\hat{\sigma}$, y utilizar la distribución de Student de $N - 2$ grados de libertad en lugar de la distribución normal.

Por ejemplo, para testear la significancia de la regresión ($H_0: \beta = 0$), se examinará si el valor de $\frac{b-0}{\sqrt{\text{var}(b)}}$ es compatible con una variable de Student de $N - 2$ grados de libertad. Un test equivalente consiste en calcular el coeficiente de correlación empírico entre x e y :

$$R = \frac{SS_{xy} - S_x S_y}{\sqrt{(SS_{xx} - S_x^2)(SS_{yy} - S_y^2)}} = \frac{\sum_{i=1}^N \frac{1}{r_i^2} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\left(\sum_{i=1}^N \frac{1}{r_i^2} (x_i - \bar{x})^2 \right) \left(\sum_{i=1}^N \frac{1}{r_i^2} (y_i - \bar{y})^2 \right)}}$$

Bajo la hipótesis nula ($\beta = 0$), la variable

$$T_{N-2} = R \sqrt{\frac{N-2}{1-R^2}}$$

sigue una distribución de Student con $N - 2$ grados de libertad. Por consiguiente, se rechazará la hipótesis nula si $|T_{N-2}|$ es demasiado grande.

3.2.3. Desviaciones constantes

Cuando las desviaciones $\{\sigma_i, i = 1 \dots N\}$ son iguales, el método lleva el nombre de *mínimos cuadrados ordinarios*, es decir, no ponderados. Se tiene:

$$\forall i \in \{1, \dots, N\}, \sigma_i = \sigma \text{ desconocido}$$

o, equivalentemente, $r_i = 1$. Para validar la hipótesis que las desviaciones son iguales, se puede graficar los residuos de regresión $y_i - \hat{y}_i$ en función de los valores estimados \hat{y}_i : se espera que la dispersión de los residuos sea aproximadamente constante.

El modelo de regresión permite prever los valores de y para un valor dado de x , al plantear

$$\hat{y} = a + bx$$

En este caso, se tiene:

$$\text{var}\{y\} = \sigma^2 \left\{ 1 + \frac{1}{S} + \frac{S(x - \bar{x})^2}{SS_{xx} - S_x^2} \right\}.$$

Esta varianza depende de x y es mínima cuando $x = \bar{x}$. Por lo tanto, el intervalo de confianza sobre y (determinado al suponer que y tiene una distribución normal de media \hat{y} y varianza dada por la fórmula anterior) es más amplio cuando x se aleja de \bar{x} .

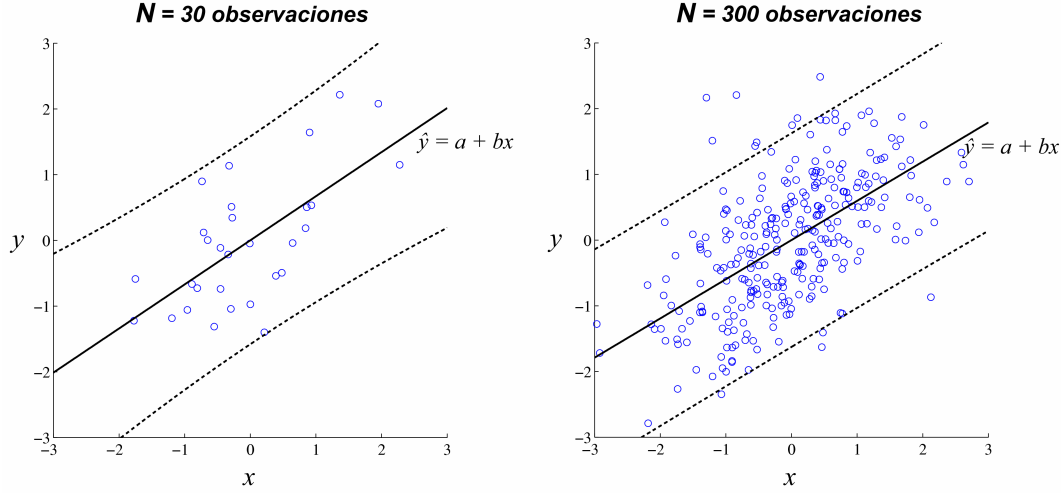


Figura 1. Regiones con 95% de confianza sobre los valores de y

4. Ajuste de un modelo lineal de varias variables

Consideremos ahora un modelo lineal de M variables:

$$\forall i \in \{1, \dots, N\}, \hat{y}_i = a_1 x_{i1} + \dots + a_M x_{iM} = \sum_{j=1}^M a_j x_{ij}$$

donde

- el índice i se refiere al número del dato (entre 1 y N)
- el índice j se refiere al número de la variable (entre 1 y M).

Si se toma $x_{i1} = 1$ para todo i , se obtiene la regresión multilineal usual (la regresión lineal examinada en la sección anterior corresponde al caso cuando M es igual a 2). La linealidad se refiere a los coeficientes a_j y no se hace supuesto sobre los valores x_{ij} (podrían ser funciones no lineales de una misma variable, por ejemplo $x_{ij} = x_i^j$). En este ejemplo, existen tantas restricciones como datos ($L = N$) y M parámetros, por lo que el número de grados de libertad es $N - M$.

4.1. Desviaciones variables y conocidas

Se supone que cada medición y_i tiene una desviación estándar σ_i conocida. Se busca minimizar

$$\chi^2 = \sum_{i=1}^N \left(\frac{y_i - \sum_{j=1}^M a_j x_{ij}}{\sigma_i} \right)^2.$$

El mínimo se encuentra al plantear

$$\forall j \in \{1, \dots, M\}, \frac{\partial \chi^2}{\partial a_j} = 0,$$

o sea:

$$\forall j \in \{1, \dots, M\}, \sum_{i=1}^N \frac{2}{\sigma_i^2} (y_i - \sum_{k=1}^M a_k x_{ik}) x_{ij} = 0.$$

Para expresar la solución, es conveniente usar notaciones vectoriales y matriciales. Sea

- \mathbf{X} una matriz de tamaño $N \times M$ cuyo término genérico es $\frac{x_{ij}}{\sigma_i}$
- \mathbf{A} un vector de tamaño $M \times 1$ cuyo término genérico es a_j
- \mathbf{Y} un vector de tamaño $N \times 1$ cuyo término genérico es $\frac{y_i}{\sigma_i}$.

El sistema de ecuaciones anteriores se escribe:

$$\mathbf{X}' \mathbf{X} \mathbf{A} = \mathbf{X}' \mathbf{Y}$$

llamado sistema de **ecuaciones normales** del problema de mínimos cuadrados. Por lo anterior, la solución está dada por

$$\mathbf{A} = (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \mathbf{Y}$$

Existen varios métodos de resolución de las ecuaciones normales:

- pivote de Gauss
- inversión matricial de $\mathbf{X}' \mathbf{X}$
- descomposición de Choleski (LU) de la matriz $\mathbf{X}' \mathbf{X}$

Se demuestra que la matriz $\mathbf{C} = (\mathbf{X}' \mathbf{X})^{-1}$ es la matriz de varianza – covarianza de los parámetros $\{a_j, j = 1 \dots M\}$. Se puede utilizar este resultado para establecer intervalos o regiones de confianza para estos parámetros, o para testear hipótesis sobre sus valores reales (en particular, para testear si es factible considerar $a_j = 0$).

Asimismo, el valor mínimo de χ^2 puede ser utilizado para calcular la probabilidad de que una variable del chi cuadrado sea mayor que este valor, lo que permite evaluar la calidad del modelo.

4.2. Desviaciones conocidas sólo con un factor de proporcionalidad

Supongamos que se conocen las desviaciones estándares solamente con un factor de proporcionalidad:

$$\forall i \in \{1, \dots, N\}, \sigma_i = \sigma r_i \text{ con } \sigma \text{ desconocido.}$$

Por ejemplo, si los y_i tienen la misma desviación desconocida, se planteará $r_i = 1$. En este caso, se resuelve las ecuaciones normales para minimizar

$$F = \sum_{i=1}^N \left(\frac{y_i - \sum_{j=1}^M a_j x_{ij}}{r_i} \right)^2.$$

(El formalismo es igual que en la sección anterior, salvo que se reemplaza σ_i por r_i). Se puede estimar el factor de proporcionalidad por:

$$\hat{\sigma} = \sqrt{\frac{F}{N-M}}$$

y obtener una estimación de la matriz de varianza – covarianza de los parámetros $\{a_j, j = 1 \dots M\}$ al plantear:

$$\hat{\mathbf{C}} = \hat{\sigma}^2 (\mathbf{X}' \mathbf{X})^{-1}$$

donde \mathbf{X} es una matriz de tamaño $N \times M$ cuyo término genérico es $\frac{x_{ij}}{r_i}$.

4.3. Significancia del modelo

Se busca evaluar la utilidad de un modelo lineal de varias variables

$$\forall i \in \{1, \dots, N\}, y_i = \sum_{j=1}^M a_j x_{ij} + \varepsilon_i$$

bajo el supuesto que los residuos ε_i son normales, independientes, de esperanza 0 y de misma varianza. Las hipótesis nula (H_0) y alternativa (H_1) son:

$$H_0: a_1 = a_2 = \dots = a_M = 0$$

$$H_1: \text{al menos un } a_i \text{ es distinto de cero}$$

La idea es realizar un **análisis de varianza**, dividiendo la suma total de residuos cuadráticos (S) en una suma debida al modelo (regresión) (S_R) y una suma debida al residuo o error (S_E). Más precisamente, se define:

$$S = \sum_{i=1}^N (y_i - \bar{y})^2 \quad S_R = \sum_{i=1}^N (\hat{y}_i - \bar{y})^2 \quad S_E = \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

tal que $S = S_R + S_E$. El procedimiento consiste en definir una variable que sigue (bajo la hipótesis nula) una distribución de Fisher con $M - 1$ y $N - M$ grados de libertad:

$$F = \frac{S_R / (M - 1)}{S_E / (N - M)}$$

y rechazar la hipótesis nula si el valor observado de F es mayor que el valor crítico para el riesgo asumido.

Se define el **coeficiente de determinación múltiple** como

$$R^2 = \frac{S_R}{S} = 1 - \frac{S_E}{S}.$$

Este coeficiente, comprendido entre 0 y 1, mide cuánto se explica la variable y al utilizar el modelo de regresión con las variables x . Cuando $M = 2$ (regresión lineal de una variable), el coeficiente de determinación múltiple no es otra cosa que el cuadrado del coeficiente de correlación entre x e y . Cuando se aumenta el número M de variables x , el coeficiente de determinación múltiple aumenta y se acerca a 1. Por lo tanto, sólo permite comparar modelos del mismo nivel, es decir, con el mismo número de variables explicativas.

Se introduce también el *coeficiente de determinación múltiple ajustado*:

$$R'^2 = 1 - \frac{S_E / (N - M)}{S / (N - 1)}$$

que introduce un castigo por el número de parámetros a estimar (nivel de la regresión). Este coeficiente ajustado no siempre aumenta al incluir variables explicativas; de hecho, si variables innecesarias están consideradas, es muy probable que baje.

4.4. Test de un modelo lineal sujeto a condiciones

Supongamos que se ha ajustado un **modelo inicial**, con M variables explicativas:

$$\hat{y} = a_1 x_1 + a_2 x_2 + \dots + a_M x_M$$

Se considera k condiciones lineales sobre los coeficientes, por ejemplo:

- $a_1 = 0$ (una condición)
- $a_1 = 0$ y $a_2 = 1$ (dos condiciones)
- $a_1 + a_2 = 0$ (una condición).

El modelo inicial puede re-escribirse en un **modelo transformado**, el cual toma en cuenta las condiciones lineales:

- $\hat{y} = a_2 x_2 + \dots + a_M x_M$
- $\hat{y} = x_2 + a_3 x_3 + \dots + a_M x_M$
- $\hat{y} = a_1 (x_1 - x_2) + \dots + a_M x_M$

El número de coeficientes libres (M) disminuye del número k de condiciones lineales. Con el modelo inicial, los coeficientes cumplen las condiciones lineales sólo de manera aproximada. Se quiere testear si el modelo transformado es aceptable:

H_0 : el modelo transformado es aceptable
 H_1 : el modelo inicial es aceptable

Se supone que los residuos (errores) de la regresión son independientes y tienen una distribución normal. Sea N el número de observaciones disponibles y M el número de variables explicativas. Sean $S_E(H_0)$ y $S_E(H_1)$ las sumas de errores cuadráticos obtenidos con el modelo transformado y con el modelo inicial, respectivamente. Se demuestra que, bajo la hipótesis nula H_0 , la cantidad

$$F = \frac{\frac{S_E(H_0) - S_E(H_1)}{k}}{\frac{S_E(H_1)}{N - M}}$$

sigue una distribución de Fisher con k y $N - M$ grados de libertad. Por consiguiente, se rechazará la hipótesis H_0 si la cantidad encontrada es demasiado grande para el nivel de confianza deseado.

Este test permite determinar si es preferible usar un modelo completo (expandido) o un modelo simplificado (reducido). Como otros casos particulares, señalemos:

- el test de nulidad de los coeficientes de la regresión ($a_1 = a_2 = \dots = a_M = 0$) visto anteriormente;
- el test de nulidad de la pendiente de una regresión lineal de una sola variable.

5. Ajuste de un modelo polinomial

Consideremos un modelo polinomial de la siguiente forma:

$$\forall i \in \{1, \dots, N\}, \hat{y}_i = a_0 + a_1 x_i + \dots + a_M x_i^M = \sum_{j=0}^M a_j x_i^j$$

donde

- el índice i se refiere al número del dato (entre 1 y N)
- el índice j se refiere al grado del monomio (entre 0 y M).

En este ejemplo, el número de grados de libertad es $N - M - 1$. Los parámetros a_0, a_1, \dots, a_M son determinados al resolver el sistema de ecuaciones normales (se trata de un caso particular de la regresión multilineal).

5.1. Grado del polinomio

La elección del grado M puede hacerse en base a la nube de correlación entre x e y , considerando también el número de datos disponibles (un grado muy alto puede llevar a una curva de regresión que fluctúa demasiado entre los puntos experimentales, por lo que el modelo está sobre-ajustado) (Figura 2).

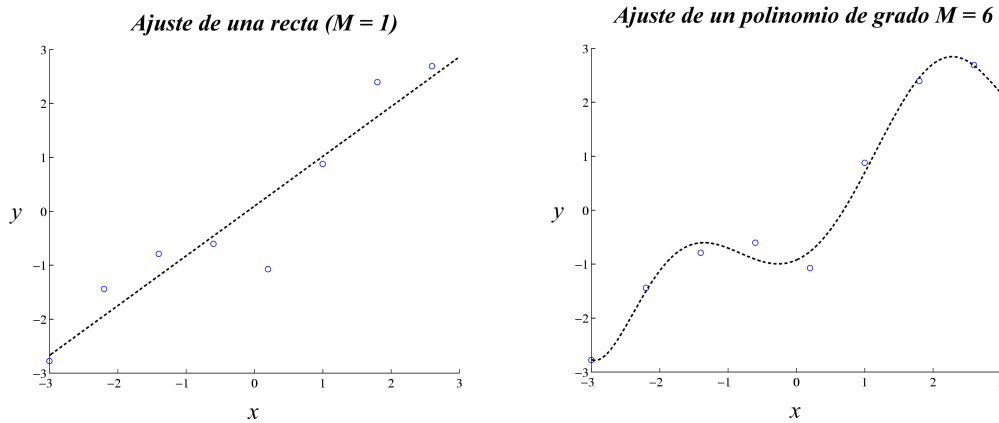


Figura 2. Modelo lineal y modelo polinomial sobre-ajustado

5.2. Significancia del ajuste polinomial

Supongamos que se ha ajustado un modelo polinomial con un determinado grado M . Se quiere saber si un modelo de grado inferior hubiese sido suficiente. Se suele proceder de forma iterativa:

- Se testea la hipótesis de que el término de mayor grado no es necesario: $a_M = 0$.
- Si se rechaza esta hipótesis, entonces el modelo de grado M es necesario. Si no, se continúa el testeo, buscando si un modelo aún más simple sería suficiente. El paso siguiente es testear la hipótesis $a_M = a_{M-1} = 0$.
- Si es preciso, continuar hasta testear $a_M = a_{M-1} = \dots = a_1 = 0$. Si se acepta esta última hipótesis, esto significa que la variable x no sirve para modelar la variable y .

Los resultados de estos tests pueden presentarse en una tabla de análisis de varianza.

Fuente de variación	Grados de libertad	Suma de cuadrados	Media de cuadrados	Razón
a_M	1	$S_{M-1} - S_E$	MS_{M-1}	MS_{M-1} / MS_E
$a_M \dots a_{M-1}$	2	$S_{M-2} - S_E$	MS_{M-2}	MS_{M-2} / MS_E
...
$a_M \dots a_2$	$M-1$	$S_1 - S_E$	MS_1	MS_1 / MS_E
$a_M \dots a_2, a_1$	M	$S_0 - S_E$	MS_0	MS_0 / MS_E
Error	$N - M - 1$	S_E	MS_E	
Total	$N - 1$	S_0		

Tabla 1. Tabla de análisis de varianza (S_p denota la suma de errores cuadráticos al ajustar una regresión polinomial de grado p , y se plantea $S_E = S_M$)

6. Ajuste de un modelo no lineal

En esta última sección, abordamos brevemente el problema donde los parámetros $\{\hat{y}_i, i = 1 \dots N\}$ son funciones no lineales de los $\{a_j, j = 1 \dots M\}$:

$$\forall i \in \{1, \dots, N\}, \hat{y}_i = f(a_1, \dots, a_M; x_{i1}, \dots, x_{iM}) = f(\mathbf{A}; \mathbf{X}_i)$$

Para minimizar la suma de cuadrados

$$\chi^2 = \sum_{i=1}^N \left(\frac{y_i - f(\mathbf{A}; \mathbf{X}_i)}{\sigma_i} \right)^2$$

se requiere anular las derivadas parciales

$$\forall j \in \{1, \dots, M\}, \frac{\partial \chi^2}{\partial a_j} = - \sum_{i=1}^N \left(\frac{y_i - f(\mathbf{A}; \mathbf{X}_i)}{\sigma_i^2} \right) \frac{\partial f(\mathbf{A}; \mathbf{X}_i)}{\partial a_j} = 0.$$

Para resolver las ecuaciones anteriores (ecuaciones normales), se suele recurrir a **métodos iterativos**, como por ejemplo los métodos de Levenberg-Marquardt y Gauss-Newton. A continuación, se explica el principio de dichos métodos. Denotemos como \mathbf{A}^k el estimador de \mathbf{A} en la k -ésima iteración, y \mathbf{A}^{min} el vector solución de las ecuaciones normales. Se tiene la siguiente aproximación (linealización de f):

$$f(\mathbf{A}^{min}; \mathbf{X}_i) \approx f(\mathbf{A}^k; \mathbf{X}_i) + \sum_{j=1}^M \frac{\partial f(\mathbf{A}^k; \mathbf{X}_i)}{\partial a_j} \delta a_j$$

con $\delta a_j = a_j^{min} - a_j^k$.

La suma de cuadrados se escribe entonces bajo la forma

$$\chi^2 = \sum_{i=1}^N \left(\frac{y_i - f(\mathbf{A}^k; \mathbf{X}_i) - \sum_{j=1}^M \frac{\partial f(\mathbf{A}^k; \mathbf{X}_i)}{\partial a_j} \delta a_j}{\sigma_i} \right)^2.$$

Se trata de un problema de mínimos cuadrados, con un modelo *lineal con respecto a las incógnitas* $\{\delta a_j, j = 1 \dots M\}$. La solución a este problema proporciona los valores de estas incógnitas, los cuales a su vez permiten actualizar los valores de los parámetros $\{a_j, j = 1 \dots M\}$:

$$\forall j \in \{1, \dots, M\}, a_j^{k+1} = a_j^k + \delta a_j.$$

Después de numerosas iteraciones, se obtiene la solución para los parámetros $\{a_j, j = 1 \dots M\}$, o sea $\{a_j^{min}, j = 1 \dots M\}$. Se obtiene también una aproximación de la matriz de varianza – covarianza de estos parámetros:

$$\mathbf{C} = (\mathbf{X}' \mathbf{X})^{-1}$$

donde \mathbf{X} es una matriz de tamaño $N \times M$ cuyo término genérico es $\frac{1}{\sigma_i} \frac{\partial f(\mathbf{A}^{min}; \mathbf{X}_i)}{\partial a_j}$.

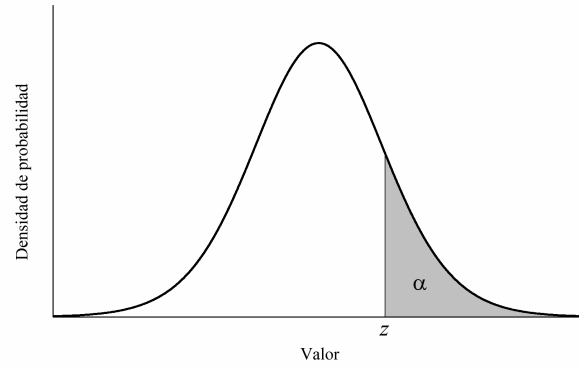
Referencias

- [1] Box, G.E.P., Hunter, W.G., Hunter, J.S., 1978. *Statistics for Experimenters*. John Wiley and Sons, New York, 653 p.
- [2] Dean, A., Voss, D., 1999. *Design and Analysis of Experiments*. Springer, New York, 740 p.
- [3] Deming, W.E., 1964. *Statistical Adjustment of Data*. Dover Publications, New York, 261 p.
- [4] Green, J.R., Margerison, D., 1978. *Statistical Treatment of Experimental Data*. Elsevier Scientific Publishing Company, Amsterdam, 382 p.
- [5] Johnson, R.A, Bhattacharyya, G.K., 1996. *Statistics: Principles and Methods*. Wiley, New York
- [6] Lapin, L.L, 1990. *Probability and Statistics for Modern Engineering*. PWS-Kent, Boston.
- [7] Mason, R.L.,Gunst, R.F., Hess, J.L., 1989. *Statistical Design and Analysis of Experiments*. John Wiley and Sons, New York.
- [8] Montgomery, D.C., 2005. *Design and Analysis of Experiments*, 6th edition. John Wiley and Sons, New York, 643 p.
- [9] Montgomery, D.C., Runger, G.C., 1999. *Applied Statistics and Probability for Engineers*. John Wiley and Sons, New York.
- [10] Snedecor, G.W., Cochran, W.G., 1968. *Statistical Methods*. Iowa State University Press, 6th edition.

Anexo: Tablas estadísticas

TABLA 1 (NORMAL)

**Distribución normal
estándar $N(0,1)$**

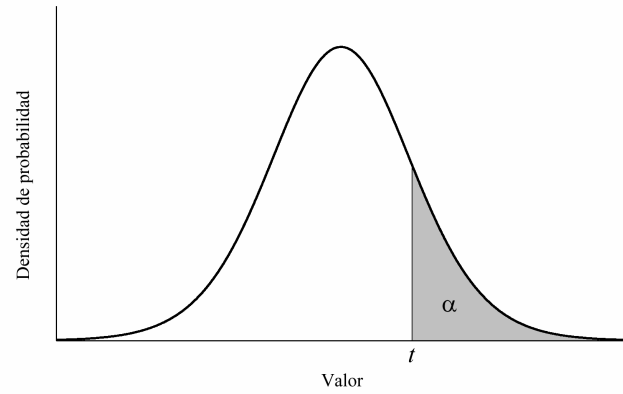


Valores con probabilidad α de ser superados

α	0.25	0.2	0.15	0.1	0.05	0.025	0.01	0.005	0.0005
z	0.675	0.84	1.03	1.28	1.64	1.96	2.32	2.57	3.27

TABLA 2 (STUDENT)

**Distribución de Student T_n
con n grados de libertad**

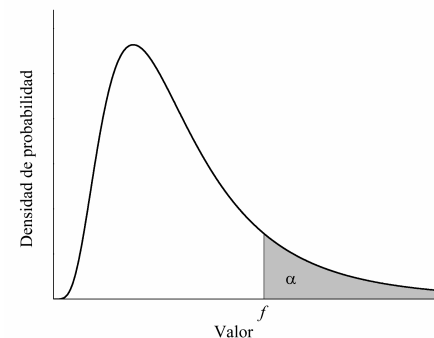


Valores con probabilidad α de ser superados

α n	0.25	0.2	0.15	0.1	0.05	0.025	0.01	0.005	0.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	636.6
2	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	31.599
3	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	12.924
4	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.781
10	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.587
11	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.437
12	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	4.318
13	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	4.140
15	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	4.073
16	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	4.015
17	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.965
18	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.922
19	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.883
20	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.850
21	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.819
22	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.792
23	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.768
24	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.745
25	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.725
26	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.707
27	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.690
28	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.674
29	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.659
30	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.646
40	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.551
60	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.460
120	0.677	0.845	1.041	1.289	1.658	1.980	2.358	2.617	3.373
∞	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.291

TABLA 3A (FISHER)

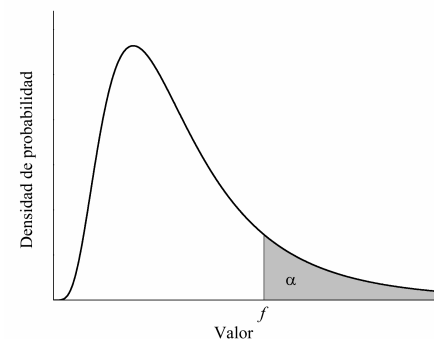
**Distribución de Fisher $F(n_1, n_2)$ con n_1 y n_2 grados de libertad
Valores con probabilidad $\alpha = 0.05$ de ser superados**



$n_2 \backslash n_1$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	16	20	24	30	40	50	75	100	200	500	∞
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	242.98	243.91	244.69	245.36	246.46	248.01	249.05	250.10	251.14	251.77	252.62	253.04	253.68	254.06	254.31
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.40	19.41	19.42	19.42	19.43	19.45	19.45	19.46	19.47	19.48	19.48	19.49	19.49	19.49	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.76	8.74	8.73	8.71	8.69	8.66	8.64	8.62	8.59	8.58	8.56	8.55	8.54	8.53	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.94	5.91	5.89	5.87	5.84	5.80	5.77	5.75	5.72	5.70	5.68	5.66	5.65	5.64	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.70	4.68	4.66	4.64	4.60	4.56	4.53	4.50	4.46	4.44	4.42	4.41	4.39	4.37	4.36
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00	3.98	3.96	3.92	3.87	3.84	3.81	3.77	3.75	3.73	3.71	3.69	3.68	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.60	3.57	3.55	3.53	3.49	3.44	3.41	3.38	3.34	3.32	3.29	3.27	3.25	3.24	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.31	3.28	3.26	3.24	3.20	3.15	3.12	3.08	3.04	3.02	2.99	2.97	2.95	2.94	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.10	3.07	3.05	3.03	2.99	2.94	2.90	2.86	2.83	2.80	2.77	2.76	2.73	2.72	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.94	2.91	2.89	2.86	2.83	2.77	2.74	2.70	2.66	2.64	2.60	2.59	2.56	2.55	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.82	2.79	2.76	2.74	2.70	2.65	2.61	2.57	2.53	2.51	2.47	2.46	2.43	2.42	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.72	2.69	2.66	2.64	2.60	2.54	2.51	2.47	2.43	2.40	2.37	2.35	2.32	2.31	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.63	2.60	2.58	2.55	2.51	2.46	2.42	2.38	2.34	2.31	2.28	2.26	2.23	2.22	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.57	2.53	2.51	2.48	2.44	2.39	2.35	2.31	2.27	2.24	2.21	2.19	2.16	2.14	2.13
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.46	2.42	2.40	2.37	2.33	2.28	2.24	2.19	2.15	2.12	2.09	2.07	2.04	2.02	2.01
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.31	2.28	2.25	2.22	2.18	2.12	2.08	2.04	1.99	1.97	1.93	1.91	1.88	1.86	1.84
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.22	2.18	2.15	2.13	2.09	2.03	1.98	1.94	1.89	1.86	1.82	1.80	1.77	1.75	1.73
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.13	2.09	2.06	2.04	1.99	1.93	1.89	1.84	1.79	1.76	1.72	1.70	1.66	1.64	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.04	2.00	1.97	1.95	1.90	1.84	1.79	1.74	1.69	1.66	1.61	1.59	1.55	1.53	1.51
50	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07	2.03	1.99	1.95	1.92	1.89	1.85	1.78	1.74	1.69	1.63	1.60	1.55	1.52	1.48	1.46	1.44
75	3.97	3.12	2.73	2.49	2.34	2.22	2.13	2.06	2.01	1.96	1.92	1.88	1.85	1.83	1.78	1.71	1.66	1.61	1.55	1.52	1.47	1.44	1.39	1.36	1.34
100	3.94	3.09	2.70	2.46	2.31	2.19	2.10	2.03	1.97	1.93	1.89	1.85	1.82	1.79	1.75	1.68	1.63	1.57	1.52	1.48	1.42	1.39	1.34	1.31	1.28
200	3.89	3.04	2.65	2.42	2.26	2.14	2.06	1.98	1.93	1.88	1.84	1.80	1.77	1.74	1.69	1.62	1.57	1.52	1.46	1.41	1.35	1.32	1.26	1.22	1.19
500	3.86	3.01	2.62	2.39	2.23	2.12	2.03	1.96	1.90	1.85	1.81	1.77	1.74	1.71	1.66	1.59	1.54	1.48	1.42	1.38	1.31	1.28	1.21	1.16	1.11
∞	3.75	2.97	2.59	2.36	2.21	2.09	2.01	1.94	1.88	1.83	1.79	1.75	1.72	1.69	1.64	1.57	1.52	1.46	1.39	1.35	1.28	1.24	1.17	1.11	1.00

TABLA 3B (FISHER)

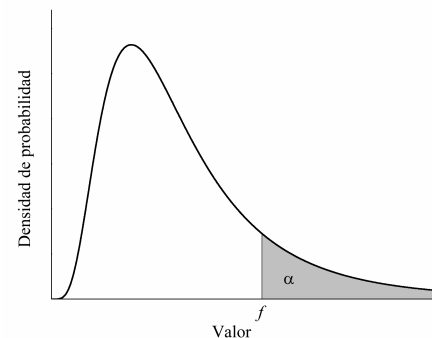
Distribución de Fisher $F(n_1, n_2)$ con n_1 y n_2 grados de libertad
Valores con probabilidad $\alpha = 0.025$ de ser superados



$n_2 \backslash n_1$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	16	20	24	30	40	50	75	100	200	500	∞
1	647.79	799.50	864.16	899.58	921.85	937.11	948.22	956.66	963.28	968.63	973.03	976.71	979.84	982.53	986.92	993.10	997.25	1001.4	1005.6	1008.1	1011.5	1013.2	1015.7	1017.2	1018.3
2	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40	39.41	39.41	39.42	39.43	39.44	39.45	39.46	39.46	39.47	39.48	39.48	39.49	39.49	39.50	39.50
3	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42	14.37	14.34	14.30	14.28	14.23	14.17	14.12	14.08	14.04	14.01	13.97	13.96	13.93	13.91	13.90
4	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.79	8.75	8.71	8.68	8.63	8.56	8.51	8.46	8.41	8.38	8.34	8.32	8.29	8.27	8.26
5	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.57	6.52	6.49	6.46	6.40	6.33	6.28	6.23	6.18	6.14	6.10	6.08	6.05	6.03	6.02
6	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.41	5.37	5.33	5.30	5.24	5.17	5.12	5.07	5.01	4.98	4.94	4.92	4.88	4.86	4.85
7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.71	4.67	4.63	4.60	4.54	4.47	4.41	4.36	4.31	4.28	4.23	4.21	4.18	4.16	4.14
8	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.24	4.20	4.16	4.13	4.08	4.00	3.95	3.89	3.84	3.81	3.76	3.74	3.70	3.68	3.67
9	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.91	3.87	3.83	3.80	3.74	3.67	3.61	3.56	3.51	3.47	3.43	3.40	3.37	3.35	3.33
10	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.66	3.62	3.58	3.55	3.50	3.42	3.37	3.31	3.26	3.22	3.18	3.15	3.12	3.09	3.08
11	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.47	3.43	3.39	3.36	3.30	3.23	3.17	3.12	3.06	3.03	2.98	2.96	2.92	2.90	2.88
12	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.32	3.28	3.24	3.21	3.15	3.07	3.02	2.96	2.91	2.87	2.82	2.80	2.76	2.74	2.72
13	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31	3.25	3.20	3.15	3.12	3.08	3.03	2.95	2.89	2.84	2.78	2.74	2.70	2.67	2.63	2.61	2.60
14	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21	3.15	3.09	3.05	3.01	2.98	2.92	2.84	2.79	2.73	2.67	2.64	2.59	2.56	2.53	2.50	2.49
16	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05	2.99	2.93	2.89	2.85	2.82	2.76	2.68	2.63	2.57	2.51	2.47	2.42	2.40	2.36	2.33	2.32
20	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.72	2.68	2.64	2.60	2.55	2.46	2.41	2.35	2.29	2.25	2.20	2.17	2.13	2.10	2.09
24	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70	2.64	2.59	2.54	2.50	2.47	2.41	2.33	2.27	2.21	2.15	2.11	2.05	2.02	1.98	1.95	1.94
30	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.46	2.41	2.37	2.34	2.28	2.20	2.14	2.07	2.01	1.97	1.91	1.88	1.84	1.81	1.79
40	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45	2.39	2.33	2.29	2.25	2.21	2.15	2.07	2.01	1.94	1.88	1.83	1.77	1.74	1.69	1.66	1.64
50	5.34	3.97	3.39	3.05	2.83	2.67	2.55	2.46	2.38	2.32	2.26	2.22	2.18	2.14	2.08	1.99	1.93	1.87	1.80	1.75	1.69	1.66	1.60	1.57	1.55
75	5.23	3.88	3.30	2.96	2.74	2.58	2.46	2.37	2.29	2.22	2.17	2.12	2.08	2.05	1.99	1.90	1.83	1.76	1.69	1.65	1.58	1.54	1.48	1.44	1.42
100	5.18	3.83	3.25	2.92	2.70	2.54	2.42	2.32	2.24	2.18	2.12	2.08	2.04	2.00	1.94	1.85	1.78	1.71	1.64	1.59	1.52	1.48	1.42	1.38	1.35
200	5.10	3.76	3.18	2.85	2.63	2.47	2.35	2.26	2.18	2.11	2.06	2.01	1.97	1.93	1.87	1.78	1.71	1.64	1.56	1.51	1.44	1.39	1.32	1.27	1.23
500	5.05	3.72	3.14	2.81	2.59	2.43	2.31	2.22	2.14	2.07	2.02	1.97	1.93	1.89	1.83	1.74	1.67	1.60	1.52	1.46	1.38	1.34	1.25	1.19	1.14
∞	4.93	3.67	3.11	2.78	2.56	2.41	2.29	2.19	2.11	2.05	1.99	1.94	1.90	1.87	1.80	1.71	1.64	1.57	1.48	1.43	1.34	1.30	1.21	1.13	1.00

TABLA 3C (FISHER)

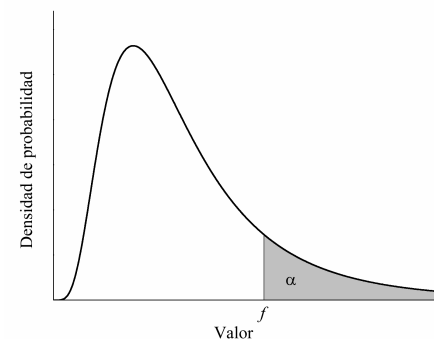
Distribución de Fisher $F(n_1, n_2)$ con n_1 y n_2 grados de libertad
Valores con probabilidad $\alpha = 0.01$ de ser superados



$n_2 \backslash n_1$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	16	20	24	30	40	50	75	100	200	500	∞
1	4052.2	4999.5	5403.4	5624.6	5763.7	5859.0	5928.4	5981.1	6022.5	6055.9	6083.3	6106.3	6125.9	6142.7	6170.1	6208.7	6234.6	6260.7	6286.8	6302.5	6323.6	6334.1	6350.0	6359.5	6365.9
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.41	99.42	99.42	99.43	99.44	99.45	99.46	99.47	99.47	99.48	99.49	99.49	99.49	99.50	99.50
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23	27.13	27.05	26.98	26.92	26.83	26.69	26.60	26.50	26.41	26.35	26.28	26.24	26.18	26.15	26.13
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.45	14.37	14.31	14.25	14.15	14.02	13.93	13.84	13.75	13.69	13.61	13.58	13.52	13.49	13.46
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.96	9.89	9.82	9.77	9.68	9.55	9.47	9.38	9.29	9.24	9.17	9.13	9.08	9.04	9.02
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.79	7.72	7.66	7.60	7.52	7.40	7.31	7.23	7.14	7.09	7.02	6.99	6.93	6.90	6.88
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.54	6.47	6.41	6.36	6.28	6.16	6.07	5.99	5.91	5.86	5.79	5.75	5.70	5.67	5.65
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.73	5.67	5.61	5.56	5.48	5.36	5.28	5.20	5.12	5.07	5.00	4.96	4.91	4.88	4.86
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.18	5.11	5.05	5.01	4.92	4.81	4.73	4.65	4.57	4.52	4.45	4.41	4.36	4.33	4.31
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.77	4.71	4.65	4.60	4.52	4.41	4.33	4.25	4.17	4.12	4.05	4.01	3.96	3.93	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.46	4.40	4.34	4.29	4.21	4.10	4.02	3.94	3.86	3.81	3.74	3.71	3.66	3.62	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.22	4.16	4.10	4.05	3.97	3.86	3.78	3.70	3.62	3.57	3.50	3.47	3.41	3.38	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	4.02	3.96	3.91	3.86	3.78	3.66	3.59	3.51	3.43	3.38	3.31	3.27	3.22	3.19	3.17
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.86	3.80	3.75	3.70	3.62	3.51	3.43	3.35	3.27	3.22	3.15	3.11	3.06	3.03	3.00
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.62	3.55	3.50	3.45	3.37	3.26	3.18	3.10	3.02	2.97	2.90	2.86	2.81	2.78	2.75
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.29	3.23	3.18	3.13	3.05	2.94	2.86	2.78	2.69	2.64	2.57	2.54	2.48	2.44	2.42
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.09	3.03	2.98	2.93	2.85	2.74	2.66	2.58	2.49	2.44	2.37	2.33	2.27	2.24	2.21
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.91	2.84	2.79	2.74	2.66	2.55	2.47	2.39	2.30	2.25	2.17	2.13	2.07	2.03	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.73	2.66	2.61	2.56	2.48	2.37	2.29	2.20	2.11	2.06	1.98	1.94	1.87	1.83	1.80
50	7.17	5.06	4.20	3.72	3.41	3.19	3.02	2.89	2.78	2.70	2.63	2.56	2.51	2.46	2.38	2.27	2.18	2.10	2.01	1.95	1.87	1.82	1.76	1.71	1.68
75	6.99	4.90	4.05	3.58	3.27	3.05	2.89	2.76	2.65	2.57	2.49	2.43	2.38	2.33	2.25	2.13	2.05	1.96	1.87	1.81	1.72	1.67	1.60	1.55	1.52
100	6.90	4.82	3.98	3.51	3.21	2.99	2.82	2.69	2.59	2.50	2.43	2.37	2.31	2.27	2.19	2.07	1.98	1.89	1.80	1.74	1.65	1.60	1.52	1.47	1.43
200	6.76	4.71	3.88	3.41	3.11	2.89	2.73	2.60	2.50	2.41	2.34	2.27	2.22	2.17	2.09	1.97	1.89	1.79	1.69	1.63	1.53	1.48	1.39	1.33	1.28
500	6.69	4.65	3.82	3.36	3.05	2.84	2.68	2.55	2.44	2.36	2.28	2.22	2.17	2.12	2.04	1.92	1.83	1.74	1.63	1.57	1.47	1.41	1.31	1.23	1.16
∞	6.58	4.61	3.79	3.33	3.02	2.81	2.64	2.52	2.41	2.32	2.25	2.19	2.13	2.08	2.00	1.88	1.79	1.70	1.59	1.52	1.42	1.36	1.25	1.15	1.00

TABLA 3D (FISHER)

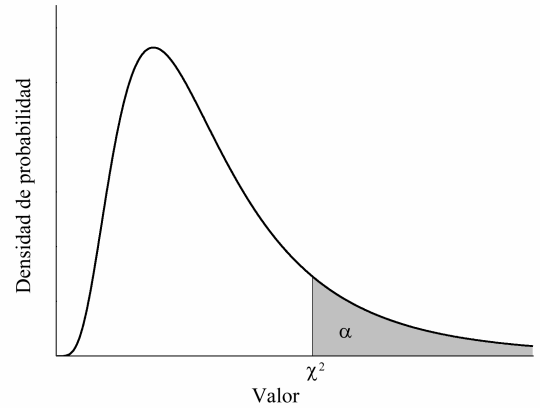
Distribución de Fisher $F(n_1, n_2)$ con n_1 y n_2 grados de libertad
Valores con probabilidad $\alpha = 0.005$ de ser superados



$n_2 \backslash n_1$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	16	20	24	30	40	50	75	100	200	500	∞
1	16210	19999	21614	22499	23056	23437	23715	23925	24091	24224	24334	24426	24504	24572	24681	24836	24940	25044	25148	25211	25295	25337	25401	25439	25464
2	198.50	199.00	199.17	199.25	199.30	199.33	199.36	199.37	199.39	199.40	199.41	199.42	199.42	199.43	199.44	199.45	199.46	199.47	199.47	199.48	199.49	199.49	199.49	199.50	199.50
3	55.55	49.80	47.47	46.19	45.39	44.84	44.43	44.13	43.88	43.69	43.52	43.39	43.27	43.17	43.01	42.78	42.62	42.47	42.31	42.21	42.09	42.02	41.93	41.87	41.83
4	31.33	26.28	24.26	23.15	22.46	21.97	21.62	21.35	21.14	20.97	20.82	20.70	20.60	20.51	20.37	20.17	20.03	19.89	19.75	19.67	19.55	19.50	19.41	19.36	19.32
5	22.78	18.31	16.53	15.56	14.94	14.51	14.20	13.96	13.77	13.62	13.49	13.38	13.29	13.21	13.09	12.90	12.78	12.66	12.53	12.45	12.35	12.30	12.22	12.17	12.14
6	18.63	14.54	12.92	12.03	11.46	11.07	10.79	10.57	10.39	10.25	10.13	10.03	9.95	9.88	9.76	9.59	9.47	9.36	9.24	9.17	9.07	9.03	8.95	8.91	8.88
7	16.24	12.40	10.88	10.05	9.52	9.16	8.89	8.68	8.51	8.38	8.27	8.18	8.10	8.03	7.91	7.75	7.64	7.53	7.42	7.35	7.26	7.22	7.15	7.10	7.08
8	14.69	11.04	9.60	8.81	8.30	7.95	7.69	7.50	7.34	7.21	7.10	7.01	6.94	6.87	6.76	6.61	6.50	6.40	6.29	6.22	6.13	6.09	6.02	5.98	5.95
9	13.61	10.11	8.72	7.96	7.47	7.13	6.88	6.69	6.54	6.42	6.31	6.23	6.15	6.09	5.98	5.83	5.73	5.62	5.52	5.45	5.37	5.32	5.26	5.21	5.19
10	12.83	9.43	8.08	7.34	6.87	6.54	6.30	6.12	5.97	5.85	5.75	5.66	5.59	5.53	5.42	5.27	5.17	5.07	4.97	4.90	4.82	4.77	4.71	4.67	4.64
11	12.23	8.91	7.60	6.88	6.42	6.10	5.86	5.68	5.54	5.42	5.32	5.24	5.16	5.10	5.00	4.86	4.76	4.65	4.55	4.49	4.40	4.36	4.29	4.25	4.23
12	11.75	8.51	7.23	6.52	6.07	5.76	5.52	5.35	5.20	5.09	4.99	4.91	4.84	4.77	4.67	4.53	4.43	4.33	4.23	4.17	4.08	4.04	3.97	3.93	3.90
13	11.37	8.19	6.93	6.23	5.79	5.48	5.25	5.08	4.94	4.82	4.72	4.64	4.57	4.51	4.41	4.27	4.17	4.07	3.97	3.91	3.82	3.78	3.71	3.67	3.65
14	11.06	7.92	6.68	6.00	5.56	5.26	5.03	4.86	4.72	4.60	4.51	4.43	4.36	4.30	4.20	4.06	3.96	3.86	3.76	3.70	3.61	3.57	3.50	3.46	3.44
16	10.58	7.51	6.30	5.64	5.21	4.91	4.69	4.52	4.38	4.27	4.18	4.10	4.03	3.97	3.87	3.73	3.64	3.54	3.44	3.37	3.29	3.25	3.18	3.14	3.11
20	9.94	6.99	5.82	5.17	4.76	4.47	4.26	4.09	3.96	3.85	3.76	3.68	3.61	3.55	3.46	3.32	3.22	3.12	3.02	2.96	2.87	2.83	2.76	2.72	2.69
24	9.55	6.66	5.52	4.89	4.49	4.20	3.99	3.83	3.69	3.59	3.50	3.42	3.35	3.30	3.20	3.06	2.97	2.87	2.77	2.70	2.61	2.57	2.50	2.46	2.43
30	9.18	6.35	5.24	4.62	4.23	3.95	3.74	3.58	3.45	3.34	3.25	3.18	3.11	3.06	2.96	2.82	2.73	2.63	2.52	2.46	2.37	2.32	2.25	2.21	2.18
40	8.83	6.07	4.98	4.37	3.99	3.71	3.51	3.35	3.22	3.12	3.03	2.95	2.89	2.83	2.74	2.60	2.50	2.40	2.30	2.23	2.14	2.09	2.01	1.96	1.93
50	8.63	5.90	4.83	4.23	3.85	3.58	3.38	3.22	3.09	2.99	2.90	2.82	2.76	2.70	2.61	2.47	2.37	2.27	2.16	2.10	2.00	1.95	1.87	1.82	1.79
75	8.37	5.69	4.63	4.05	3.67	3.41	3.21	3.05	2.93	2.82	2.74	2.66	2.60	2.54	2.45	2.31	2.21	2.10	1.99	1.92	1.82	1.77	1.68	1.63	1.59
100	8.24	5.59	4.54	3.96	3.59	3.33	3.13	2.97	2.85	2.74	2.66	2.58	2.52	2.46	2.37	2.23	2.13	2.02	1.91	1.84	1.74	1.68	1.59	1.53	1.49
200	8.06	5.44	4.41	3.84	3.47	3.21	3.01	2.86	2.73	2.63	2.54	2.47	2.40	2.35	2.25	2.11	2.01	1.91	1.79	1.71	1.60	1.54	1.44	1.37	1.31
500	7.95	5.35	4.33	3.76	3.40	3.14	2.94	2.79	2.66	2.56	2.48	2.40	2.34	2.28	2.19	2.04	1.94	1.84	1.72	1.64	1.52	1.46	1.35	1.26	1.18
∞	7.90	5.34	4.31	3.74	3.37	3.10	2.91	2.75	2.63	2.53	2.44	2.36	2.30	2.24	2.15	2.00	1.90	1.79	1.67	1.59	1.47	1.40	1.28	1.17	1.00

TABLA 4 (CHI CUADRADO)

**Distribución del chi cuadrado χ_n^2
con n grados de libertad**



Valores con probabilidad α de ser superados

$n \backslash \alpha$	0.25	0.2	0.15	0.1	0.05	0.025	0.01	0.005	0.0005
1	1.32	1.64	2.07	2.71	3.84	5.02	6.64	7.88	12.12
2	2.77	3.22	3.79	4.61	5.99	7.38	9.21	10.60	15.20
3	4.11	4.64	5.32	6.25	7.81	9.35	11.34	12.84	17.73
4	5.39	5.99	6.74	7.78	9.49	11.14	13.28	14.86	20.00
5	6.63	7.29	8.12	9.24	11.07	12.83	15.09	16.75	22.11
6	7.84	8.56	9.45	10.64	12.59	14.45	16.81	18.55	24.10
7	9.04	9.80	10.75	12.02	14.07	16.01	18.48	20.28	26.02
8	10.22	11.03	12.03	13.36	15.51	17.53	20.09	21.95	27.87
9	11.39	12.24	13.29	14.68	16.92	19.02	21.67	23.59	29.67
10	12.55	13.44	14.53	15.99	18.31	20.48	23.21	25.19	31.42
11	13.70	14.63	15.77	17.28	19.68	21.92	24.72	26.76	33.14
12	14.85	15.91	16.99	18.55	21.03	23.34	26.22	28.30	34.82
13	15.98	16.98	18.20	19.81	22.36	24.74	27.69	29.82	36.48
14	17.12	18.15	19.41	21.06	23.68	26.12	29.14	31.32	38.11
15	18.25	19.31	20.60	22.31	25.00	27.49	30.58	32.80	39.72
16	19.37	20.47	21.79	23.54	26.30	28.85	32.00	34.27	41.31
17	20.49	21.61	22.98	24.77	27.59	30.19	33.41	35.72	42.88
18	21.60	22.76	24.16	25.99	28.87	31.53	34.81	37.16	44.43
19	22.72	23.90	25.33	27.20	30.14	32.85	36.19	38.58	45.97
20	23.83	25.04	26.50	28.41	31.41	34.17	37.57	40.00	47.50
21	24.93	26.17	27.66	29.62	32.67	35.48	38.93	41.40	49.01
22	26.04	27.30	28.82	30.81	33.92	36.78	40.29	42.80	50.51
23	27.14	28.43	29.98	32.01	35.17	38.08	41.64	44.18	52.00
24	28.24	29.55	31.13	33.20	36.42	39.36	42.98	45.56	53.48
25	29.34	30.68	32.28	34.39	37.65	40.65	44.31	46.93	54.95
26	30.43	31.79	33.43	35.56	38.89	41.92	45.64	48.29	56.41
27	31.53	32.91	34.57	36.74	40.11	43.19	46.96	49.64	57.86
28	32.62	34.03	35.71	37.92	41.34	44.46	48.28	50.99	59.30
29	33.71	35.14	36.85	39.09	42.56	45.72	49.59	52.34	60.73
30	34.80	36.25	37.99	40.26	43.77	46.98	50.89	53.67	62.16
40	45.62	47.27	49.24	51.81	55.76	59.34	63.69	66.77	76.09
50	56.33	58.16	60.35	63.17	67.50	71.42	76.15	79.49	89.56
60	66.98	68.97	71.34	74.40	79.08	83.30	88.38	91.95	102.69
70	77.58	79.71	82.26	85.53	90.53	95.02	100.4	104.2	115.6
80	88.13	90.41	93.11	96.58	101.9	106.6	112.3	116.3	128.3
90	98.65	101.1	103.9	107.6	113.2	118.1	124.1	128.3	140.8
100	109.1	111.7	114.7	118.5	124.3	129.6	135.8	140.2	153.2