

MATEMÁTICAS II: NOTAS DE TEORÍA
GRADO EN QUÍMICAS

PROF. JAVIER PASTOR

DEPARTAMENTO DE MATEMÁTICA APLICADA
UNIVERSIDAD DE VALENCIA

Índice general

1. Teoría del Muestreo Estadístico	3
1.1. Introducción	3
1.2. Parámetros muestrales	4
1.3. Probabilidad	5
1.4. Variables aleatorias	7
1.4.1. Variables aleatorias discretas	8
1.4.2. Variables aleatorias continuas	10
1.5. Intervalos de confianza	13
1.5.1. Intervalos de confianza para la media μ de una distribución normal $N(\mu, \sigma)$ con nivel de confianza $1 - \varepsilon$	14
1.5.2. Intervalo de confianza para p de una distribución normal $B(1, p)$ con nivel de confianza $1 - \varepsilon$	14
1.5.3. Intervalos de confianza para varianza σ^2 de una distribución normal $N(\mu, \sigma)$ con nivel de confianza $1 - \varepsilon$	14
1.5.4. Intervalo de confianza para la diferencia de medias $\mu_1 - \mu_2$ de dos poblaciones $X_1 \rightsquigarrow N(\mu_1, \sigma_1)$ y $X_2 \rightsquigarrow N(\mu_2, \sigma_2)$ independientes, con nivel de confianza $1 - \varepsilon$	15
1.5.5. Intervalo de confianza para el cociente de varianzas σ_1^2/σ_2^2 de dos poblaciones normales $X_1 \rightsquigarrow N(\mu_1, \sigma_1)$ y $X_2 \rightsquigarrow N(\mu_2, \sigma_2)$ independientes, con nivel de confianza $1 - \varepsilon$, $0 < \varepsilon < 1$	15
1.6. Contraste de hipótesis	15
1.6.1. Contraste sobre la media μ de una distribución normal $N(\mu, \sigma)$ a un nivel de significación α	16
1.6.2. Contraste sobre la varianza σ^2 de una distribución normal $N(\mu, \sigma)$ a un nivel de significación α	16
1.6.3. Contraste sobre el parámetro p de una distribución binomial a un nivel de significación α	17
1.6.4. Contraste de las medias μ_1 y μ_2 de dos poblaciones $X_1 \rightsquigarrow N(\mu_1, \sigma_1)$ y $X_2 \rightsquigarrow N(\mu_2, \sigma_2)$ independientes a un nivel de significación α	17
1.6.5. Contraste de las varianzas σ_1^2 y σ_2^2 de dos poblaciones normales $X_1 \rightsquigarrow N(\mu_1, \sigma_1)$ y $X_2 \rightsquigarrow N(\mu_2, \sigma_2)$ independientes, con nivel de significación α	18
2. Ajuste y Regresión	19
2.1. Ajuste por mínimos cuadrados: la recta de regresión	19
2.2. Intervalos de confianza para los parámetros y la predicción	20

2.2.1.	Intervalo de confianza para α con nivel de confianza $1 - \varepsilon$	20
2.2.2.	Intervalo de confianza para β con nivel de confianza $1 - \varepsilon$	20
2.2.3.	Intervalo de confianza para σ^2 con nivel de confianza $1 - \varepsilon$	21
2.2.4.	Intervalo de predicción a nivel $1 - \varepsilon$ para el valor de la variable y correspondiente al valor x_0	21
2.3.	Ajustes no lineales reducibles a un ajuste lineal	21
2.3.1.	Ajuste hiperbólico	21
2.3.2.	Ajuste exponencial	21
2.3.3.	Ajuste potencial	21
2.4.	Ajuste parabólico	22
3.	Interpolación Polinómica	23
3.1.	Formas de Lagrange y de Newton	23
3.1.1.	Lagrange	23
3.1.2.	Newton	23
3.1.3.	Error de interpolación	24
3.2.	Interpolación segmentaria	26
3.3.	Interpolación de Hermite. Polinomio de Taylor	27
3.3.1.	Interpolación de Hermite	27
3.3.2.	Polinomio de Taylor	27
4.	Integración Numérica	28
4.1.	Fórmulas de Integración de Newton-Cotes	28
4.2.	Fórmulas compuestas	29
4.2.1.	Fórmula de los trapecios	29
4.2.2.	Fórmula de Simpson	30
4.3.	Integración de ecuaciones diferenciales	30
4.3.1.	Método de Euler	30
4.3.2.	Método de Euler mejorado	31
4.3.3.	Método de Runge-Kutta de cuarto orden	31
5.	Tablas	32
5.1.	Tabla de la normal tipificada	32
5.2.	Tabla de la distribución χ^2 de Pearson	33
5.3.	Tabla de la distribución t de Student	34
5.4.	Tablas para la F de Snedecor	35
	Bibliografía	37

Capítulo 1

Teoría del Muestreo Estadístico

1.1. Introducción

Un **experimento aleatorio** es aquel del que conocemos sus posibles resultados, pero de forma que el resultado es impredecible, fruto del azar, como contraposición a situaciones determinísticas en las que el efecto es una consecuencia cierta de las causas que lo producen.

En Estadística se denomina **población** al conjunto de todos los elementos acerca de los que se desea obtener información, a cada miembro de la población se le llama elemento o **individuo**, y se llama **variable estadística** a la característica que se observa en cada elemento de la población. Estas características pueden ser cualitativas o cuantitativas, pudiendo por tanto clasificarse estas últimas en *discretas*, si toma valores en un conjunto finito o infinito numerable, o en *continuas*, si toma valores en un intervalo.

Una **muestra** es un subconjunto extraído de la población cuyo estudio nos debe servir para sacar conclusiones válidas y tomar decisiones razonables sobre toda la población. El estudio de muestras puede resultar preferible a los censos, o estudio de toda la población, por diferentes razones: la población es muy grande (o incluso infinita), la característica de estudio de la población varía con el tiempo, el estudio de la muestra es menos costoso y más rápido, en el proceso de estudio es necesario consumir un artículo para extraer la muestra, etc.

Se llama **muestreo** a la técnica utilizada en la selección de una muestra a partir de una población. El uso de muestras para deducir fiablemente las características de la población requiere que se trate con muestras aleatorias. Si la muestra estadística considerada no constituye una muestra aleatoria, las conclusiones basadas en dicha muestra no son fiables y en general estarán sesgadas en algún aspecto. En el **muestreo aleatorio simple (m.a.s.)** todos los individuos tienen la misma probabilidad de ser seleccionados. Así, el m.a.s. en el caso de poblaciones finitas es equivalente a un proceso de extracción con reemplazamiento.

Muy resumidamente, podemos decir que nuestro objetivo en este tema es aprender a extraer, resumir y comunicar información a partir de conjuntos de datos experimentales, construir modelos para estos datos a través de variables aleatorias adecuadas, y dotarnos de las herramientas básicas nece-

sarias para poder llegar a conclusiones significativas sobre una población a partir de una muestra.

1.2. Parámetros muestrales

Sean $\{x_i\}_{i=1}^k$ una muestra de una población, es decir, un conjunto de valores observados, de forma que el valor x_i se ha observado en un número n_i de ocasiones ($1 \leq i \leq k$). Al número n_i lo llamaremos **frecuencia absoluta** del valor x_i en la muestra y al número $N = \sum_{i=1}^k n_i$ lo llamaremos tamaño de la muestra.

La **frecuencia relativa** del valor x_i , f_i , es el cociente entre la frecuencia absoluta y el número total de observaciones:

$$f_i := \frac{n_i}{N} \quad (1 \leq i \leq k).$$

Obsérvese que $\sum_{i=1}^k f_i = 1$.

La **media muestral** del conjunto de valores observados la denotaremos mediante \bar{x} y se define como la media aritmética de dichos valores:

$$\bar{x} := \frac{\sum_{i=1}^k n_i x_i}{N} = \sum_{i=1}^k f_i x_i.$$

Para medir la agrupación o dispersión global de los datos entorno a la media, se introduce la llamada **varianza muestral**

$$d^2 := \frac{\sum_{i=1}^k n_i (x_i - \bar{x})^2}{N} = \sum_{i=1}^k f_i (x_i - \bar{x})^2 = \frac{\sum_{i=1}^k n_i x_i^2}{N} - \bar{x}^2,$$

o la **desviación típica muestral**

$$d := +\sqrt{\sum_{i=1}^k f_i (x_i - \bar{x})^2},$$

que presenta la ventaja de medirse en la mismas unidades que la variable original. Por sus mejores propiedades matemáticas (véase más adelante la estimación de la varianza poblacional), suele emplearse la expresión anterior con $N - 1$ en lugar de N en los denominadores; es decir, se considera el número

$$s^2 := \frac{\sum_{i=1}^k n_i (x_i - \bar{x})^2}{N - 1},$$

que llamaremos **varianza muestral corregida** o **cuasivarianza muestral**.

La **moda** de la muestra se define como el valor(es) que presenta(n) mayor frecuencia absoluta.

La **mediana** es el valor M determinado de la siguiente forma. Supongamos que escribimos todos los valores de la muestra repitiendo cada uno de los sucesos tantas veces como indique su frecuencia absoluta y los renombramos, es decir,

$$\begin{aligned} z_1 &= z_2 = \dots = z_{n_1} := x_1 < z_{n_1+1} = z_{n_1+2} = \dots = z_{n_1+n_2} := x_2 \\ &< \dots < z_{n_1+\dots+n_{k-1}+1} = \dots = z_N = x_k \end{aligned}$$

Si $N = 2m+1$, entonces $M := z_{m+1}$, mientras que cuando $N = 2m$, entonces $M := (z_{m+1} + z_m)/2$.

El **rango** es la diferencia entre el mayor y el menor valor de la muestra.

1.3. Probabilidad

Definición 1. Un **espacio muestral** para un experimento aleatorio es un conjunto Ω tal que cada resultado posible del experimento corresponde exactamente a un elemento de Ω . El conjunto puede tener un número finito o infinito numerable de elementos en cuyo caso diremos que el espacio muestral es **discreto**, o un número infinito no numerable en cuyo caso diremos que el espacio muestral es **continuo**.

Cada subconjunto A de Ω es un **suceso** y lo indicaremos escribiendo $A \subset \Omega$. Los elementos de Ω son los **sucesos elementales** o **simples**. El propio espacio muestral Ω es un suceso llamado **suceso seguro**, mientras que el conjunto vacío, \emptyset , representa el **suceso imposible**.

Tras realizar un experimento aleatorio diremos que se ha realizado o que ha ocurrido el suceso $A \subset \Omega$ si el resultado obtenido pertenece al conjunto A .

El conjunto de todos los posibles sucesos no es más que el conjunto de las partes de Ω , o de los subconjuntos de Ω , que representaremos mediante $\mathcal{P}(\Omega)$. Si Ω tiene cardinal finito n , entonces el número de sucesos será 2^n .

Sean $A, B \subset \Omega$. Al igual que sobre se hace sobre $\mathcal{P}(\Omega)$, definimos el suceso **unión** de A y B

$$A \cup B := \{w \in \Omega : w \in A \text{ o } w \in B\},$$

el suceso **intersección** de A y B

$$A \cap B := \{w \in \Omega : w \in A \text{ y } w \in B\},$$

y el suceso **complementario** de A

$$\bar{A} := \{w \in \Omega : w \notin A\}.$$

Podemos considerar también el suceso $B \setminus A := B \cap \bar{A}$ y hablar de la **inclusión** de sucesos, $A \subset B$, que se interpreta como que si ocurre A , entonces ocurre B (recuérdese que en teoría de conjuntos, $A \subset B$ significa que si $w \in A$, entonces $w \in B$).

Recuérdense las **leyes de Morgan**: $\overline{A \cup B} = \bar{A} \cap \bar{B}$ y $\overline{A \cap B} = \bar{A} \cup \bar{B}$, para todo $A, B \subset \Omega$.

Definición 2. Sean $A, B \subset \Omega$. Diremos que A y B son **mutuamente excluyentes** si $A \cap B = \emptyset$.

Definición 3. Una **función de probabilidad** sobre un espacio muestral Ω es una aplicación

$$P : \mathcal{P}(\Omega) \rightarrow [0, 1] \\ A \rightsquigarrow P(A)$$

cumpliendo:

$$(i) P(\Omega) = 1,$$

(ii) si los sucesos $\{A_k\}_{k=1}^{\infty}$ son mutuamente excluyentes dos a dos (es decir, $A_i \cap A_j = \emptyset$, $i \neq j$, $i, j \in \mathbb{N}$), entonces

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

Si se verifican las condiciones anteriores, diremos que el par (Ω, P) es un **espacio probabilístico o de probabilidad**.

Proposición 4. Sean (Ω, P) un espacio probabilístico y $A, B \subset \Omega$. Entonces

1. $P(B) = P(B \cap \bar{A}) + P(B \cap A)$.
2. $P(\bar{A}) = 1 - P(A)$. En particular, $P(\emptyset) = 0$.
3. Si $A \subset B$, entonces $P(A) \leq P(B)$ y $P(B \setminus A) = P(B) - P(A)$.
4. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
5. **Postulado de Indiferencia:** si $\Omega = \cup_{i=1}^n \{e_i\}$, y todos los sucesos simples $\{e_i\}$ son equiprobables, entonces

$$P(\{e_i\}) = \frac{1}{n} \quad (1 \leq i \leq n),$$

y si $A = \cup_{k=1}^r \{e_{i_k}\}$, entonces

$$P(A) = \frac{r}{n} = \frac{n. \text{ casos favorables}}{n. \text{ casos posibles}} \quad (\text{ley de Laplace}).$$

Definición 5. Si A y B son sucesos de un cierto espacio de probabilidad (Ω, P) , con $P(B) > 0$, la **probabilidad de A condicionada por B** se define como

$$P(A/B) := \frac{P(A \cap B)}{P(B)}.$$

Proposición 6. Sea (Ω, P) un espacio probabilístico. Sea $B \in \Omega$ con $P(B) > 0$. Consideramos la aplicación

$$P(. / B) : \mathcal{P}(\Omega) \rightarrow [0, 1] \\ A \rightsquigarrow P(A/B)$$

Entonces $(\Omega, P(. / B))$ es un espacio probabilístico.

Nótese que, a raíz del resultado anterior, tenemos automáticamente relaciones sobre la probabilidad condicionada tales como $P(\bar{A}/B) = 1 - P(A/B)$, $P((A \cup C)/B) = P(A/B) + P(C/B) - P((A \cap C)/B)$, etc

Definición 7. Los sucesos A y B de un espacio de probabilidad (Ω, P) se dice que son **independientes** si

$$P(A \cap B) = P(A)P(B).$$

Obsérvese que, si $P(B) > 0$, entonces A y B son independientes, si, y sólo si, $P(A/B) = P(A)$.

Teorema 8 (Probabilidad Compuesta). Sean $B_i \subset \Omega$, $1 \leq i \leq k$, de forma que $P(\cap_{i=1}^{k-1} B_i) > 0$. Entonces

$$P(\cap_{i=1}^k B_i) = P(B_1)P(B_2/B_1) \dots P(B_k/\cap_{i=1}^{k-1} B_i).$$

Definición 9. Un conjunto de sucesos $\{A_i\}_{i=1}^k$ de Ω diremos que es una **partición de Ω** si son mutuamente excluyentes dos a dos y $\cup_{i=1}^k A_i = \Omega$.

Teorema 10 (Probabilidad total). Sea $\{A_i\}_{i=1}^k$ una partición de Ω , con $P(A_i) > 0$, $1 \leq i \leq k$, y sea $A \subset \Omega$. Entonces

$$P(A) = \sum_{i=1}^k P(A_i)P(A/A_i).$$

Teorema 11 (Fórmula de Bayes). Sea $\{A_i\}_{i=1}^k$ una partición de Ω , con $P(A_i) > 0$, $1 \leq i \leq k$, y sea $A \subset \Omega$ con $P(A) > 0$. Entonces

$$P(A_i/A) = \frac{P(A_i)P(A/A_i)}{\sum_{j=1}^k P(A_j)P(A/A_j)} \quad (1 \leq i \leq k).$$

1.4. Variables aleatorias

Definición 12. Sea Ω el espacio muestral asociado a un experimento aleatorio. Una **variable aleatoria (v.a.)** es una aplicación del tipo

$$X : \Omega \rightarrow \mathbb{R}.$$

Llamaremos **soporte** de X al conjunto imagen de la aplicación; es decir, al conjunto

$$S_X := \{X(s) : s \in \Omega\}.$$

La v.a. X se dice que es **discreta** si S_X se compone de un número finito o infinito numerable de números reales, mientras que si S_X es un intervalo de números reales se dice que X es **continua**.

Si sobre Ω tenemos definida una función de probabilidad P , a cada subconjunto W de números reales le podemos atribuir una probabilidad mediante

$$P_X(W) \equiv P_X(X \in W) := P(\{s \in \Omega : X(s) \in W\}).$$

En adelante no indicaremos la dependencia de la probabilidad respecto de X .

En particular, para todo $x \in \mathbb{R}$,

$$P(X \leq x) \equiv P(X \in]-\infty, x]) = P(\{s \in \Omega : X(s) \in]-\infty, x]\}).$$

Definición 13. La **función de distribución de probabilidad**, F , de una v.a. X es

$$\begin{aligned} F : \mathbb{R} &\rightarrow [0, 1] \\ x &\rightsquigarrow F(x) = P(X \leq x) \end{aligned}$$

Proposición 14. Sea F la función de distribución de una v.a. X . Sean $x, y \in \mathbb{R}$. Se cumplen:

1. si $x \leq y$, entonces $F(x) \leq F(y)$,
2. $P(X > x) = 1 - F(x)$,
3. $P(x < X \leq y) = F(y) - F(x)$,
4. $\lim_{x \rightarrow -\infty} F(x) = 0$ y $\lim_{x \rightarrow \infty} F(x) = 1$,
5. F es continua por la derecha en toda la recta real.

1.4.1. Variables aleatorias discretas

Definición 15. Sea X una v.a. discreta. Una **función de densidad** para X es cualquier función $f : \mathbb{R} \rightarrow \mathbb{R}$ cumpliendo:

1. $f(x) \geq 0$ ($x \in \mathbb{R}$),
2. $f(x) = 0$ ($x \notin S_X$),
3. $\sum_{x \in S_X} f(x) = 1$.

Determinar que función de densidad es la adecuada a cada experimento aleatorio es equivalente a proporcionar una probabilidad a cada número real:

$$P(X = x) = f(x) \quad (x \in \mathbb{R}).$$

La **función de distribución** correspondiente es

$$F(x) = P(X \leq x) = \sum_{y \leq x} f(y) \quad (x \in \mathbb{R}).$$

Nótese que F es una función escalonada con saltos en los valores de S_X .

Definición 16. Sea X una v.a. discreta con función de densidad f . La **esperanza** o **media** de X es

$$E(X) := \sum_{x \in S_X} x f(x),$$

que también indicaremos mediante μ . La **varianza** de X es

$$\text{Var}(X) := E((X - \mu)^2) = \sum_{x \in S_X} (x - \mu)^2 f(x),$$

que también indicaremos mediante σ^2 . La **desviación típica** de X es la raíz cuadrada positiva de la varianza y se denotará mediante σ .

Distribución uniforme discreta

Modeliza problemas en los que la v. a. toma un número finito de valores con la misma probabilidad. Si la v. a. toma los valores $\{x_i\}_{i=1}^k$ con igual probabilidad, diremos que X se distribuye de forma **uniforme discreta**, $X \rightsquigarrow U(\{x_i\}_{i=1}^k)$, y su función de densidad está determinada por

$$f(x) = P(X = x) := \begin{cases} 1/k, & x \in \{x_i\}_{i=1}^k, \\ 0, & x \notin \{x_i\}_{i=1}^k. \end{cases}$$

Su media estadística es $\mu = \sum_{i=1}^k x_i/k$ y su varianza $\sigma^2 = \sum_{i=1}^k (x_i - \mu)^2/k$.

Distribución Bernoulli

Un **ensayo de Bernoulli** es un experimento aleatorio en el que sólo puede producirse un suceso o el contrario (dos sucesos excluyentes, éxito-fracaso) con probabilidades $p \in [0, 1]$ y $q := 1 - p$, respectivamente. Si X es una variable aleatoria que toma el valor 0 con probabilidad q y el valor 1 con probabilidad p diremos que sigue una distribución de Bernoulli de parámetro p , y lo indicaremos mediante $X \rightsquigarrow B_e(p)$. Su función de densidad es

$$f(x) := \begin{cases} p^x q^{1-x}, & x \in \{0, 1\}, \\ 0, & x \notin \{0, 1\}. \end{cases}$$

Se tiene $\mu = p$ y $\sigma^2 = pq$.

Distribución geométrica

La distribución geométrica es un modelo adecuado para aquellos procesos en los que se repiten pruebas independientes de Bernoulli hasta la consecución del primer éxito con probabilidad de éxito p y nos interesa estudiar la v.a. X que representa el número de repeticiones realizadas antes del primer éxito. Diremos que la v.a. X se distribuye como una **geométrica de parámetro p** , $X \rightsquigarrow G(p)$, si su función de densidad es

$$f(r) := \begin{cases} pq^r, & r = 0, 1, 2, \dots, \\ 0, & r \neq 0, 1, 2, \dots \end{cases}$$

Se tiene $\mu = qp^{-1}$ y $\sigma^2 = qp^{-2}$.

Distribución binomial

Consideramos n ensayos de Bernoulli independientes de forma que la probabilidad del suceso éxito permanece inalterada de un ensayo al siguiente. Sea X la v.a. que representa el número de éxitos entre los n ensayos; así X puede tomar los valores $0, 1, \dots, n$. La función de densidad apropiada para esta variable es

$$f(r) := \begin{cases} \binom{n}{r} p^r q^{n-r}, & r = 0, 1, \dots, n, \\ 0, & r \neq 0, 1, \dots, n, \end{cases}$$

donde $\binom{n}{r} := \frac{n!}{r!(n-r)!}$. Entonces diremos que X se distribuye como una **binomial de parámetros n y p** , y lo indicaremos mediante $X \rightsquigarrow B(n, p)$. Se cumple: $\mu = np$ y $\sigma^2 = npq$.

Distribución de Poisson

La distribución de Poisson es útil para estudiar la probabilidad de que un determinado número de eventos ocurran en un determinado periodo de tiempo, dada una frecuencia media conocida e independientemente del tiempo transcurrido desde el último evento. Por ejemplo en los siguientes contextos: número de mutaciones de determinada cadena de ADN después de cierta cantidad de radiación, número de servidores web accedidos por minuto, número de núcleos atómicos inestables que decayeron durante un determinado período en una porción de sustancia radiactiva, número de estrellas en un determinado volumen de espacio, etc.

Una v.a. X se dice que sigue una distribución de **Poisson de parámetro** $\lambda > 0$ si la función de densidad asociada es

$$f(r) := \begin{cases} \frac{\lambda^r}{r!} e^{-\lambda}, & r \in \{0, 1, 2, \dots\}, \\ 0, & r \notin \{0, 1, 2, \dots\}, \end{cases}$$

y lo indicaremos escribiendo $X \rightsquigarrow P(\lambda)$. Se cumple que $\mu = \sigma^2 = \lambda$.

Si $n \rightarrow \infty$, $p = p_n \rightarrow 0$, y $np \rightarrow \lambda$, entonces $\lim_{n \rightarrow \infty} B(r, n, p) = P(r, \lambda)$. Utilizando este resultado se aproxima $B(n, p)$ por $P(np)$ cuando $n > 50$, $p < 0.1$ y $np < 5$.

1.4.2. Variables aleatorias continuas

Definición 17. Sea X una v. a. continua. Una **función de densidad** para X es aquella de la forma $f : \mathbb{R} \rightarrow \mathbb{R}$ cumpliendo

1. $f(x) \geq 0$ ($x \in \mathbb{R}$),
2. $\int_{-\infty}^{\infty} f(x) dx = 1$.

La **función de distribución** correspondiente es

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(s) ds, \quad x \in \mathbb{R}.$$

En la definición anterior se sobreentiende que las integrales existen. Nótese que $P(X = x) = 0$ ($x \in \mathbb{R}$). Además, para $a \leq b$,

$$P(a \leq X \leq b) = F(b) - F(a) = \int_a^b f(s) ds,$$

lo que se puede interpretar como el *área* encerrada por la función de densidad f , el eje x y las rectas $x = a$, $x = b$.

Definición 18. Sea X una v.a. continua con función de densidad f . La **esperanza o media** de X es

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx,$$

que también indicaremos mediante μ . La **varianza** de X es

$$\text{Var}(X) := E((X - \mu)^2) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx,$$

que también indicaremos mediante σ^2 . La **desviación típica** de X es la raíz cuadrada positiva de la varianza y se denotará mediante σ .

Distribución uniforme

Una v.a. sigue una distribución **uniforme sobre el intervalo de extremos a y b** (escribiremos $X \rightsquigarrow U(a, b)$), si su función de densidad es

$$f(x) := \begin{cases} 1/(b - a), & x \in [a, b], \\ 0, & x \notin [a, b]. \end{cases}$$

Su media estadística es $\mu = (a + b)/2$ y su varianza $\sigma^2 = (b - a)^2/12$.

Es interesante comprobar que la probabilidad de que X tome valores en dos subintervalos de $[a, b]$ de igual longitud es la misma en ambos casos.

Distribución exponencial

Una v.a. sigue una distribución **exponencial de parámetro** $\lambda > 0$ si su función de densidad es

$$f(x) := \begin{cases} \lambda e^{-\lambda x}, & x \geq 0, \\ 0, & x < 0, \end{cases}$$

y escribiremos $X \rightsquigarrow \text{Exp}(\lambda)$. La media de X es $\mu = \lambda^{-1}$ y su varianza $\sigma^2 = \lambda^{-2}$.

Distribución normal

Una v.a. X se dice que sigue una distribución **normal de parámetros** μ y σ ($\mu \in \mathbb{R}$, $\sigma > 0$) si la función de densidad de X es

$$f_{\mu,\sigma}(x) := \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (x \in \mathbb{R}).$$

Lo indicaremos escribiendo $X \rightsquigarrow N(\mu, \sigma)$. Se cumple que $E(X) = \mu$ y $\text{Var}(X) = \sigma^2$.

Si $X \rightsquigarrow N(\mu, \sigma)$, entonces $Z := (X - \mu)/\sigma \rightsquigarrow N(0, 1)$ es la llamada **normal tipificada o estándar** y está tabulada (véase pág. 32). Este proceso de pasar a considerar la $N(0, 1)$ a partir de la $N(\mu, \sigma)$ se conoce como *tipificación*. Así, $P(a \leq X \leq b)$ puede calcularse utilizando tablas para Z mediante la igualdad

$$\begin{aligned} P(a \leq X \leq b) &= P((a - \mu)/\sigma \leq Z \leq (b - \mu)/\sigma) \\ &= P(Z \geq (a - \mu)/\sigma) - P(Z \geq (b - \mu)/\sigma). \end{aligned}$$

La tabla de la normal tipificada sólo contiene valores asociados a valores positivos, pero los restantes se obtienen por simetría de la función de densidad respecto del eje y :

$$P(Z > a) = 1 - P(Z \geq -a).$$

De esta forma, para $0 < p < 1$, podemos encontrar un número real z_p tal que $P(Z < z_p) = p$.

La distribución normal se usa en el estudio de características morfológicas de individuos (estatura, perímetro bíceps, etc.), características fisiológicas (efecto de una misma dosis de un fármaco), características sociológicas (consumo de cierto producto por un mismo grupo de individuos), características psicológicas (cociente intelectual), en teoría de errores, para aproximar otras distribuciones, etc.

Distribuciones deducidas de la normal

Las siguientes distribuciones serán útiles en el estudio de los intervalos de confianza y el contraste de hipótesis.

Distribución χ^2 de Pearson Sean X_i , $1 \leq i \leq n$, v. a. independientes con $X_i \rightsquigarrow N(0, 1)$, $1 \leq i \leq n$. Entonces la v.a.

$$\chi_n^2 := \sum_{i=1}^n X_i^2$$

recibe el nombre de χ^2 (ji-cuadrado) **de Pearson con n grados de libertad**. Su función de densidad es

$$f_{\chi_n^2}(x) =: \begin{cases} 0, & x \leq 0, \\ 2^{-n/2} e^{-x/2} x^{-1+n/2} / \Gamma(n/2), & x > 0, \end{cases}$$

donde $\Gamma(p) := \int_0^\infty e^{-x} x^{p-1} dx$, $p > 0$, es la función *gamma de Euler* ($\Gamma(n+1) = n!$). Se tiene $\mu_{\chi_n^2} = n$ y $\sigma_{\chi_n^2} = \sqrt{2n}$.

En la tabla de la pág. 33 se proporciona $\chi_{n;p}^2 \geq 0$ tal que $P(\chi_n^2 < \chi_{n;p}^2) = p$ para $1 \leq n \leq 30$ y para diferentes valores de p . Para un número de grados de libertad mayor que 30 se utiliza en los cálculos que $\sqrt{2\chi_n^2} \approx N(\sqrt{2n-1}, 1)$ (o por supuesto MATLAB).

Distribución t de Student La v.a.

$$t_n := \frac{Z}{\sqrt{\frac{1}{n} \chi_n^2}}$$

se denomina **t de Student con n grados de libertad** (Z y χ_n^2 se suponen v.a. independientes). La función de densidad asociada es

$$f_{t_n}(x) := \frac{\Gamma((n+1)/2)}{\sqrt{n\pi} \Gamma(n/2)} (1 + x^2/n)^{-(n+1)/2}, \quad x \in \mathbb{R}.$$

La v.a. t_n se encuentra tabulada para valores de $n \leq 30$ (véase pág. 34) y se puede usar para determinar $t_{n;p} \geq 0$ tal que $P(t_n \leq t_{n;p}) = p$ con $p \geq 0.5$. Si $p < 0.5$, el área encerrada a la derecha del número buscado es $1-p$ y, por la simetría de la función de densidad, el punto buscado es $-t_{n,1-p}$.

Se verifica que $\mu_{t_n} = 0$ y $\sigma_{t_n}^2 = n/(n-2)$.

Para $n > 30$ la t_n se aproxima muy bien mediante $N(0, 1)$.

Distribución F de Fisher-Snedecor A una v.a. del tipo

$$F_{m,n} := \frac{\chi_m^2/m}{\chi_n^2/n}$$

se le denomina **F de Fisher-Snedecor con m y n grados de libertad** (χ_m^2 y χ_n^2 se suponen v.a. independientes).

Su función de densidad es

$$f_{m,n}(x) := \begin{cases} 0, & x \leq 0, \\ \frac{\Gamma((m+n)/2) (\frac{m}{n})^{m/2}}{\Gamma(m/2) \Gamma(n/2)} x^{-1+m/2} (1 + mx/n)^{-(m+n)/2}, & x > 0, \end{cases}$$

Se verifica $\mu_{F_{m,n}} = n/(n-2)$ y $\sigma_{F_{m,n}}^2 = 2n^2(m+n-2)/(m(n-4)(n-2)^2)$.

Obsérvese que por la definición $F_{n,m} = 1/F_{m,n}$, luego si $F_{m,n;p}$ cumple $P(F_{m,n} \leq F_{m,n;p}) = p$, entonces $F_{m,n;p} = 1/F_{n,m;1-p}$.

Podemos encontrar tablas para esta distribución en las páginas 35-36.

1.5. Intervalos de confianza

Consideremos una v.a. $X \rightsquigarrow N(\mu, \sigma)$, donde recordemos que μ es la media poblacional y σ es la desviación típica poblacional. Sean $X_i \rightsquigarrow N(\mu, \sigma)$, $1 \leq i \leq n$, v.a. independientes idénticas a X (muestra aleatoria simple de tamaño n). Consideramos las nuevas v.a.

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

media muestral, y

$$S^2 := \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1},$$

cuasivarianza muestral. Observadas las v.a. X_i para obtener una muestra aleatoria simple de tamaño n de X , tendremos valores concretos de las v.a. \bar{X} y S , que denotamos \bar{x} y s , respectivamente.

Se cumplen las siguientes propiedades:

- (i) $\bar{X} \rightsquigarrow N(\mu, \sigma/\sqrt{n})$.
- (ii) $\frac{n-1}{\sigma^2} S^2 \rightsquigarrow \mathcal{X}_{n-1}^2$.
- (iii) Las v.a. \bar{X} y $\frac{n-1}{\sigma^2} S^2$ son independientes.
- (iv) $\sqrt{n}(\bar{X} - \mu)/S \rightsquigarrow t_{n-1}$.

Otro resultado fundamental es el llamado **teorema central del límite**, que establece que si X_1, \dots, X_n son variables aleatorias independientes con la misma función de densidad desconocida, y μ y σ^2 representan la media y la varianza poblacional respectivamente, ambas finitas, entonces

$$\lim_n \sqrt{n}(\bar{X} - \mu)/\sigma = Z$$

donde $Z \rightsquigarrow N(0, 1)$.

Utilizando básicamente estas propiedades se determinan intervalos aleatorios en los que está el parámetro desconocido de la población objeto de estudio, con una probabilidad prefijada $1 - \varepsilon$, $0 < \varepsilon < 1$, que llamaremos **intervalos de confianza con un nivel de confianza $1 - \varepsilon$** . Una vez dispongamos de una muestra, obtendremos un valor concreto, no aleatorio, del intervalo de confianza, que en el $(1 - \varepsilon)\%$ de las muestras contendrá al parámetro desconocido.

En adelante usaremos la siguiente notación: para $0 < p < 1$,

- z_p es el número real cumpliendo $P(Z < z_p) = p$,
- $\mathcal{X}_{n,p}^2$ es el número real cumpliendo $P(\mathcal{X}_n^2 < \mathcal{X}_{n,p}^2) = p$,
- $t_{n,p}$ es el número real cumpliendo $P(t_n < t_{n,p}) = p$,
- $F_{m,n;p}$ es el número real cumpliendo $P(F_{m,n} < F_{m,n;p}) = p$.

1.5.1. Intervalos de confianza para la media μ de una distribución normal $N(\mu, \sigma)$ con nivel de confianza $1 - \varepsilon$

1. σ conocida:

$$\left[\bar{X} - \frac{\sigma}{\sqrt{n}} z_{1-\varepsilon/2}, \bar{X} + \frac{\sigma}{\sqrt{n}} z_{1-\varepsilon/2} \right],$$

es decir,

$$P\left(\bar{X} - \frac{\sigma}{\sqrt{n}} z_{1-\varepsilon/2} \leq \mu \leq \bar{X} + \frac{\sigma}{\sqrt{n}} z_{1-\varepsilon/2}\right) = 1 - \varepsilon.$$

Si tenemos una muestra x_1, \dots, x_n , obtendríamos el intervalo no aleatorio

$$\left[\bar{x} - \frac{\sigma}{\sqrt{n}} z_{1-\varepsilon/2}, \bar{x} + \frac{\sigma}{\sqrt{n}} z_{1-\varepsilon/2} \right].$$

2. σ desconocida:

$$\left[\bar{X} - \frac{S}{\sqrt{n}} t_{n-1, 1-\varepsilon/2}, \bar{X} + \frac{S}{\sqrt{n}} t_{n-1, 1-\varepsilon/2} \right] \quad (n \leq 30)$$

$$\left[\bar{X} - \frac{S}{\sqrt{n}} z_{1-\varepsilon/2}, \bar{X} + \frac{S}{\sqrt{n}} z_{1-\varepsilon/2} \right] \quad (n > 30)$$

1.5.2. Intervalo de confianza para p de una distribución normal $B(1, p)$ con nivel de confianza $1 - \varepsilon$

Recordando que una binomial $B(1, p)$ tiene media p y desviación típica \sqrt{pq} , para muestras de tamaño grande ($n > 30$), se toma como intervalo de confianza del parámetro p al intervalo

$$\left[\hat{p} - \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} z_{1-\varepsilon/2}, \hat{p} + \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} z_{1-\varepsilon/2} \right],$$

donde \hat{p} es la v.a. frecuencia relativa del suceso éxito. En la obtención de este intervalo están involucrados el Teorema Central del Límite y la aproximación de una binomial por una normal.

1.5.3. Intervalos de confianza para varianza σ^2 de una distribución normal $N(\mu, \sigma)$ con nivel de confianza $1 - \varepsilon$

1. μ desconocida:

$$\left[\frac{(n-1)}{\chi_{n-1, 1-\varepsilon/2}^2} S^2, \frac{(n-1)}{\chi_{n-1, \varepsilon/2}^2} S^2 \right]$$

2. μ conocida:

$$\left[\frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi_{n, 1-\varepsilon/2}^2}, \frac{\sum_{i=1}^n (X_i - \mu)^2}{\chi_{n, \varepsilon/2}^2} \right]$$

1.5.4. Intervalo de confianza para la diferencia de medias $\mu_1 - \mu_2$ de dos poblaciones $X_1 \rightsquigarrow N(\mu_1, \sigma_1)$ y $X_2 \rightsquigarrow N(\mu_2, \sigma_2)$ independientes, con nivel de confianza $1 - \varepsilon$

Se supone que disponemos de una muestra de X_1 y otra de X_2 de tamaño n_1 y n_2 respectivamente.

1. σ_1, σ_2 conocidas:

$$(\bar{X}_1 - \bar{X}_2) \mp \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} z_{1-\varepsilon/2}$$

2. $\sigma_1 = \sigma_2$ desconocidas:

$$(\bar{X}_1 - \bar{X}_2) \mp \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}} t_{n_1 + n_2 - 2, 1-\varepsilon/2}$$

3. σ_1, σ_2 desconocidas y $n_1, n_2 > 30$:

$$(\bar{X}_1 - \bar{X}_2) \mp \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} z_{1-\varepsilon/2}$$

1.5.5. Intervalo de confianza para el cociente de varianzas σ_1^2/σ_2^2 de dos poblaciones normales $X_1 \rightsquigarrow N(\mu_1, \sigma_1)$ y $X_2 \rightsquigarrow N(\mu_2, \sigma_2)$ independientes, con nivel de confianza $1 - \varepsilon$, $0 < \varepsilon < 1$.

Se supone disponemos de una muestra de X_1 y otra de X_2 de tamaños n_1 y n_2 respectivamente, y que μ_1 y μ_2 son desconocidas, que es la situación más frecuente.

$$\left[\frac{S_1^2/S_2^2}{F_{n_1-1, n_2-1; 1-\varepsilon/2}}, \frac{S_1^2/S_2^2}{F_{n_1-1, n_2-1; \varepsilon/2}} \right].$$

1.6. Contraste de hipótesis

El contraste de hipótesis consiste en analizar si una determinada afirmación sobre una propiedad poblacional (por ejemplo acerca de un parámetro de la distribución asociada a la v.a.) es confirmada o invalidada por los datos recogidos en una muestra.

El contraste de hipótesis comienza con el planteamiento de una **hipótesis nula** H_0 y de la **hipótesis alternativa** H_1 , no siempre complementaria de la anterior. Para rechazar H_0 se utiliza el llamado **nivel de significación** α (los valores más habituales son 0.1, 0.05 ó 0.01), de forma que si la ‘compatibilidad’ de la muestra con H_0 es menor que α , rechazaremos dicha hipótesis de partida y daremos como válida la hipótesis alternativa H_1 . El nivel de significación, α , mide, en términos probabilísticos, el riesgo de rechazar H_0 cuando es cierta.

Para facilitar comparaciones entre diferentes niveles de significación, se introduce el concepto de **p-valor** que se define como el nivel de significación más pequeño bajo el que la hipótesis nula se rechaza.

1.6.1. Contraste sobre la media μ de una distribución normal $N(\mu, \sigma)$ a un nivel de significación α .

Suponemos que disponemos de una muestra aleatoria simple de la v.a. de tamaño n con media muestral \bar{x} y cuasivarianza muestral s^2 .

La hipótesis nula es

$$H_0 : \mu = \mu_0$$

que contrastamos frente a varias alternativas

1. σ conocida:

H_1	Rechazar H_0
$\mu \neq \mu_0$	$ \bar{x} - \mu_0 \frac{\sqrt{n}}{\sigma} > z_{1-\alpha/2}$
$\mu > \mu_0$	$(\bar{x} - \mu_0) \frac{\sqrt{n}}{\sigma} > z_{1-\alpha}$
$\mu < \mu_0$	$(\bar{x} - \mu_0) \frac{\sqrt{n}}{\sigma} < z_\alpha$

2. σ desconocida: para $n \leq 30$,

H_1	Rechazar H_0
$\mu \neq \mu_0$	$ \bar{x} - \mu_0 \frac{\sqrt{n}}{s} > t_{n-1, 1-\alpha/2}$
$\mu > \mu_0$	$(\bar{x} - \mu_0) \frac{\sqrt{n}}{s} > t_{n-1, 1-\alpha}$
$\mu < \mu_0$	$(\bar{x} - \mu_0) \frac{\sqrt{n}}{s} < t_{n-1, \alpha}$

mientras que para $n > 30$ se sustituye $t_{n-1, p}$ por z_p para $p = \alpha/2, \alpha, 1-\alpha$.

1.6.2. Contraste sobre la varianza σ^2 de una distribución normal $N(\mu, \sigma)$ a un nivel de significación α .

Disponemos de una muestra aleatoria simple de tamaño n con media muestral \bar{x} y cuasivarianza muestral s^2 .

La hipótesis nula es

$$H_0 : \sigma^2 = \sigma_0^2$$

que contrastamos frente a varias alternativas:

1. μ desconocida:

H_1	Rechazar H_0
$\sigma^2 \neq \sigma_0^2$	$\frac{(n-1)}{\sigma_0^2} s^2 \notin [\chi_{n-1, \alpha/2}^2, \chi_{n-1, 1-\alpha/2}^2]$
$\sigma^2 > \sigma_0^2$	$\frac{(n-1)}{\sigma_0^2} s^2 > \chi_{n-1, 1-\alpha}^2$
$\sigma^2 < \sigma_0^2$	$\frac{(n-1)}{\sigma_0^2} s^2 < \chi_{n-1, \alpha}^2$

2. μ conocida:

H_1	Rechazar H_0
$\sigma^2 \neq \sigma_0^2$	$\frac{\sum_{i=1}^n (x_i - \mu)^2}{\sigma_0^2} \notin [\chi_{n-1, \alpha/2}^2, \chi_{n-1, 1-\alpha/2}^2]$
$\sigma^2 > \sigma_0^2$	$\frac{\sum_{i=1}^n (x_i - \mu)^2}{\sigma_0^2} > \chi_{n-1, 1-\alpha}^2$
$\sigma^2 < \sigma_0^2$	$\frac{\sum_{i=1}^n (x_i - \mu)^2}{\sigma_0^2} < \chi_{n-1, \alpha}^2$

1.6.3. Contraste sobre el parámetro p de una distribución binomial a un nivel de significación α .

Disponemos de una muestra aleatoria simple de tamaño $n > 30$ de forma que el éxito tiene frecuencia relativa en la muestra \hat{p} .

La hipótesis nula es

$$H_0 : p = p_0$$

que contrastamos frente a varias alternativas:

H_1	Rechazar H_0
$p \neq p_0$	$ \hat{p} - p_0 \sqrt{\frac{n}{p_0(1-p_0)}} > z_{1-\alpha/2}$
$p > p_0$	$(\hat{p} - p_0) \sqrt{\frac{n}{p_0(1-p_0)}} > z_{1-\alpha}$
$p < p_0$	$(\hat{p} - p_0) \sqrt{\frac{n}{p_0(1-p_0)}} < z_\alpha$

1.6.4. Contraste de las medias μ_1 y μ_2 de dos poblaciones $X_1 \rightsquigarrow N(\mu_1, \sigma_1)$ y $X_2 \rightsquigarrow N(\mu_2, \sigma_2)$ independientes a un nivel de significación α .

Disponemos de una muestra de X_1 y otra de X_2 de tamaños n_1 y n_2 , medias \bar{x}_1 y \bar{x}_2 , y cuasivarianzas muestrales s_1^2 y s_2^2 , respectivamente.

La hipótesis nula es

$$H_0 : \mu_1 = \mu_2$$

que contrastamos frente a varias alternativas:

1. σ_1, σ_2 conocidas: si $k := \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$, entonces

H_1	Rechazar H_0
$\mu_1 \neq \mu_2$	$\frac{ \bar{x}_1 - \bar{x}_2 }{k} > z_{1-\alpha/2}$
$\mu_1 > \mu_2$	$\frac{\bar{x}_1 - \bar{x}_2}{k} > z_{1-\alpha}$
$\mu_1 < \mu_2$	$\frac{\bar{x}_1 - \bar{x}_2}{k} < z_\alpha$

2. $\sigma_1 = \sigma_2$ desconocidas: si $k := \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}$, entonces

H_1	Rechazar H_0
$\mu_1 \neq \mu_2$	$\frac{ \bar{x}_1 - \bar{x}_2 }{k} > t_{n_1+n_2-2, 1-\alpha/2}$
$\mu_1 > \mu_2$	$\frac{\bar{x}_1 - \bar{x}_2}{k} > t_{n_1+n_2-2, 1-\alpha}$
$\mu_1 < \mu_2$	$\frac{\bar{x}_1 - \bar{x}_2}{k} < t_{n_1+n_2-2, \alpha}$

3. σ_1, σ_2 desconocidas y $n_1, n_2 > 30$: si $k := \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$, entonces

H_1	Rechazar H_0
$\mu_1 \neq \mu_2$	$\frac{ \bar{x}_1 - \bar{x}_2 }{k} > z_{1-\alpha/2}$
$\mu_1 > \mu_2$	$\frac{\bar{x}_1 - \bar{x}_2}{k} > z_{1-\alpha}$
$\mu_1 < \mu_2$	$\frac{\bar{x}_1 - \bar{x}_2}{k} < z_\alpha$

1.6.5. Contraste de las varianzas σ_1^2 y σ_2^2 de dos poblaciones normales $X_1 \rightsquigarrow N(\mu_1, \sigma_1)$ y $X_2 \rightsquigarrow N(\mu_2, \sigma_2)$ independientes, con nivel de significación α .

Disponemos de una muestra de X_1 y otra de X_2 de tamaños n_1 y n_2 , y cuasivarianzas muestrales s_1^2 y s_2^2 , respectivamente. Suponemos que μ_1 y μ_2 son desconocidas, que es la situación más frecuente.

La hipótesis nula es

$$H_0 : \sigma_1^2 = \sigma_2^2$$

que contrastamos frente a varias alternativas

H_1	Rechazar H_0
$\sigma_1^2 \neq \sigma_2^2$	$s_1^2/s_2^2 \notin [F_{n_1-1, n_2-1; \alpha/2}, F_{n_1-1, n_2-1; 1-\alpha/2}]$
$\sigma_1^2 > \sigma_2^2$	$s_1^2/s_2^2 > F_{n_1-1, n_2-1; 1-\alpha}$
$\sigma_1^2 < \sigma_2^2$	$s_1^2/s_2^2 < F_{n_1-1, n_2-1; \alpha}$

Capítulo 2

Ajuste y Regresión

2.1. Ajuste por mínimos cuadrados: la recta de regresión

Consideremos un conjunto de puntos del plano $\{(x_i, y_i)\}_{i=1}^m$. Nos planteamos el problema de encontrar una recta $y = \alpha x + \beta$ que contenga dichos puntos. Es evidente que este problema no tiene solución a menos que los puntos estén alineados. Los parámetros $\alpha, \beta \in \mathbb{R}$ que determinan la recta han de ser solución del sistema de ecuaciones lineales

$$y_i = \alpha x_i + \beta, \quad 1 \leq i \leq m.$$

El problema de *mínimos cuadrados* asociado al sistema anterior es

$$\min_{\alpha, \beta \in \mathbb{R}} \sum_{i=1}^m (y_i - \alpha x_i - \beta)^2,$$

cuya solución es

$$\alpha^* = \frac{m \sum_{i=1}^m x_i y_i - \sum_{i=1}^m y_i \sum_{i=1}^m x_i}{m \sum_{i=1}^m x_i^2 - (\sum_{i=1}^m x_i)^2} = \frac{S_{xy}}{S_{xx}}$$

y

$$\beta^* = \frac{\sum_{i=1}^m x_i^2 \sum_{i=1}^m y_i - \sum_{i=1}^m x_i y_i \sum_{i=1}^m x_i}{m \sum_{i=1}^m x_i^2 - (\sum_{i=1}^m x_i)^2} = \bar{y} - \bar{x} \frac{S_{xy}}{S_{xx}},$$

donde

$$S_{xy} = \frac{1}{m} \sum_{i=1}^m (x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{m} \sum_{i=1}^m x_i y_i - \bar{x} \bar{y},$$

es la *covarianza muestral* de x e y , y

$$S_{yy} = \frac{1}{m} \sum_{i=1}^m (y_i - \bar{y})^2 = \frac{1}{m} \sum_{i=1}^m y_i^2 - \bar{y}^2,$$

es decir, S_{yy} es la *varianza muestral* de la variable y . Así, la recta obtenida, llamada *recta de regresión de y sobre x* , es

$$y - \bar{y} = \frac{S_{xy}}{S_{xx}} (x - \bar{x})$$

2.2 Intervalos de confianza para los parámetros y la predicción 20

Nota 19. De forma análoga la recta de regresión de x sobre y es

$$x - \bar{x} = \frac{S_{xy}}{S_{yy}}(y - \bar{y})$$

Nótese que es una recta en general distinta a la recta de regresión de y sobre x , ya que en este caso se están minimizando las distancias horizontales de los puntos a la recta en lugar de las distancias verticales.

Utilizando que $\beta^* = \bar{y} - \bar{x}\alpha^*$, el valor del mínimo es

$$\sum_{i=1}^m (y_i - \alpha^* x_i - \beta^*)^2 = m S_{yy} \left(1 - \frac{S_{xy}^2}{S_{yy} S_{xx}} \right) \equiv m S_{yy} (1 - r^2),$$

donde

$$r := \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}},$$

es el llamado *coeficiente de correlación lineal*. Nótese que $|r| \leq 1$ y que cuando $r = \pm 1$ los puntos están alineados y viceversa.

2.2. Intervalos de confianza para los parámetros y la predicción

En los resultados que se exponen a continuación se supone que las variables aleatorias Y_i , cuyo valor observado tras fijar x_i es y_i , se distribuyen como una normal de media $\alpha x_i + \beta$, $1 \leq i \leq m$, y varianza común para todas ellas σ^2 , y que son v.a. independientes.

Consideremos la notación introducida en el tema de intervalos de confianza. Sea

$$\sigma^* := \left(\frac{1}{m} \sum_{i=1}^m (y_i - \alpha^* x_i - \beta^*)^2 \right)^{1/2} = \sqrt{S_{yy} (1 - r^2)}.$$

Para $0 < \varepsilon < 1$ se tienen los siguientes intervalos de confianza y de predicción:

2.2.1. Intervalo de confianza para α con nivel de confianza $1 - \varepsilon$

$$\alpha^* \mp t_{m-2, 1-\varepsilon/2} \frac{\sigma^*}{\sqrt{(m-2)S_{xx}}}$$

2.2.2. Intervalo de confianza para β con nivel de confianza $1 - \varepsilon$

$$\beta^* \mp t_{m-2, 1-\varepsilon/2} \sigma^* \sqrt{\frac{\sum_{i=1}^m x_i^2}{m(m-2)S_{xx}}}$$

2.2.3. Intervalo de confianza para σ^2 con nivel de confianza $1 - \varepsilon$

$$\left[\frac{m\sigma^{*2}}{\chi^2_{m-2, 1-\varepsilon/2}}, \frac{m\sigma^{*2}}{\chi^2_{m-2, \varepsilon/2}} \right]$$

2.2.4. Intervalo de predicción a nivel $1 - \varepsilon$ para el valor de la variable y correspondiente al valor x_0

$$(\alpha^*x_0 + \beta^*) \mp t_{m-2, 1-\varepsilon/2} \sigma^* \sqrt{\frac{1}{m-2} \left(m+1 + \frac{(x_0 - \bar{x})^2}{Sxx} \right)}$$

2.3. Ajustes no lineales reducibles a un ajuste lineal

2.3.1. Ajuste hiperbólico

Por el tipo de datos que queremos ajustar, puede ser razonable ajustarlos mediante una función de tipo hiperbólico

$$y = a + \frac{b}{x}.$$

Buscaremos los valores a y b , que determinan que curva elegimos, de manera que ésta se ajuste lo mejor posible a los puntos $\{(x_i, y_i)\}_{i=1}^m$. A la vista de la estructura de la función, realizamos un ajuste lineal a los datos $\{(1/x_i, y_i)\}_{i=1}^m$ para obtener los parámetros a y b .

2.3.2. Ajuste exponencial

Ahora el objetivo consiste en encontrar una curva del tipo exponencial

$$y = ae^{bx},$$

que mejor se ajuste a los puntos $\{(x_i, y_i)\}_{i=1}^m$; es decir, hay que determinar los parámetros a y b que la definen para los que el ajuste es óptimo. Tomando logaritmos en la ecuación que define la curva, se observa que

$$\log y = \log a + bx.$$

Esta relación expresa una relación lineal entre x y $\log y$, que conduce a aplicar un ajuste lineal a los datos $\{(x_i, \log y_i)\}_{i=1}^m$ para obtener los valores óptimos de $\log a$ y b , de donde, tomando exponentiales, se obtendría también el valor óptimo de a .

2.3.3. Ajuste potencial

Si se desea ajustar los datos $\{(x_i, y_i)\}_{i=1}^m$ mediante una función del tipo

$$y = ax^b,$$

realizamos un ajuste lineal entre los datos $\{(\log x_i, \log y_i)\}_{i=1}^m$ para obtener $\log a$ y b , puesto que, tomando logaritmos en la ecuación de la función, tenemos la relación

$$\log y = \log a + b \log x.$$

Conocido el valor de $\log a$ sólo nos queda tomar exponenciales para obtener a .

2.4. Ajuste parabólico

Ajustamos los datos $\{(x_i, y_i)\}_{i=1}^m$ mediante una función polinómica de segundo grado del tipo $y = \gamma x^2 + \alpha x + \beta$ en el sentido de mínimos cuadrados. Los parámetros $\alpha, \beta, \gamma \in \mathbb{R}$ se determinan buscando la solución del sistema

$$A \begin{bmatrix} \gamma \\ \alpha \\ \beta \end{bmatrix} = y, \quad (2.1)$$

donde

$$A := \begin{bmatrix} x_1^2 & x_1 & 1 \\ x_2^2 & x_2 & 1 \\ \vdots & \vdots & \vdots \\ x_m^2 & x_m & 1 \end{bmatrix}, \quad y := \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix},$$

en el sentido de mínimos cuadrados. Dicha solución se obtiene resolviendo el sistema de ecuaciones de tamaño 3×3 , llamado *sistema de ecuaciones normales* asociado al sistema (2.1), siguiente

$$A' A \begin{bmatrix} \gamma \\ \alpha \\ \beta \end{bmatrix} = A' y,$$

donde A' representa la matriz transpuesta de A .

Capítulo 3

Interpolación Polinómica

Consideremos un conjunto de puntos del plano $\{(x_i, y_i)\}_{i=0}^m$, que pueden corresponderse con puntos de la gráfica de cierta función f , es decir, $f(x_i) = y_i$, $0 \leq i \leq m$.

El problema general de interpolación consiste en encontrar, si es posible, una función g , de cierta clase de funciones como por ejemplo la de los polinomios, de forma que

$$g(x_i) = y_i \quad (0 \leq i \leq m);$$

es decir, cuya gráfica contenga a los puntos prefijados. Se dirá entonces que g interpola los datos de partida.

Existencia y unicidad del polinomio de interpolación

Si los nodos x_0, x_1, \dots, x_m , son todos distintos, entonces para cualesquiera y_0, y_1, \dots, y_m , existe un polinomio $P_m(x)$, de grado menor o igual que m , único, cumpliendo

$$P_m(x_i) = y_i \quad (0 \leq i \leq m).$$

3.1. Formas de Lagrange y de Newton

3.1.1. Lagrange

$$P_m(x) = \sum_{i=0}^m y_i L_i(x),$$

donde

$$\begin{aligned} L_i(x) & : = \frac{(x - x_0) \dots (x - x_{i-1})(x - x_{i+1}) \dots (x - x_m)}{(x_i - x_0) \dots (x_i - x_{i-1})(x_i - x_{i+1}) \dots (x_i - x_m)} \\ & = \prod_{j=0, j \neq i}^m \frac{x - x_j}{x_i - x_j} \quad (i = 0, 1, \dots, m). \end{aligned}$$

3.1.2. Newton

$$P_m(x) = f[x_0] + \sum_{k=1}^m \left(f[x_0, x_1, \dots, x_k] \prod_{j=0}^{k-1} (x - x_j) \right)$$

donde los coeficientes a determinar son las llamadas *diferencias divididas* que están definidas mediante la fórmula de recurrencia

$$\begin{aligned} f[x_i] & : = y_i \quad (i = 0, 1, \dots, m), \\ f[x_0, x_1, \dots, x_k] & : = \frac{f[x_1, x_2, \dots, x_k] - f[x_0, x_1, \dots, x_{k-1}]}{x_k - x_0} \quad (k = 1, 2, \dots, m). \end{aligned} \tag{3.1}$$

Su cálculo se puede organizar con facilidad en una tabla:

x_i	$f[.]$	$f[.,.]$	$f[.,.,.]$	\dots	$f[\overbrace{., \dots, .}^{m+1}]$
x_0	$f[x_0]$				
x_1	$f[x_1]$	$f[x_0, x_1]$			
x_2	$f[x_2]$	$f[x_1, x_2]$	$f[x_0, x_1, x_2]$		
\vdots	\vdots	\vdots	\vdots	\ddots	
x_{m-1}	$f[x_{m-1}]$	$f[x_{m-2}, x_{m-1}]$	$f[x_{m-3}, x_{m-2}, x_{m-1}]$	\dots	
x_m	$f[x_m]$	$f[x_{m-1}, x_m]$	$f[x_{m-2}, x_{m-1}, x_m]$	\dots	$f[x_0, \dots, x_m]$

Además, esta forma del polinomio de interpolación permite añadir un dato con comodidad a partir del cálculo ya realizado. En concreto, si añadimos el dato (x_{m+1}, y_{m+1}) , entonces

$$P_{m+1}(x) = P_m(x) + f[x_0, x_1, \dots, x_{m+1}] \prod_{j=0}^m (x - x_j).$$

3.1.3. Error de interpolación

Sea $f : [a, b] \rightarrow \mathbb{R}$ es $m + 1$ veces derivable con derivadas continuas, y sea $M = \max_{x \in [a, b]} |f^{(m+1)}(x)|$. Sea $P_m(x)$ el polinomio de interpolación correspondiente a los datos $(x_i, f(x_i))$, $0 \leq i \leq m$, donde $x_i \in [a, b]$, $0 \leq i \leq m$. Entonces

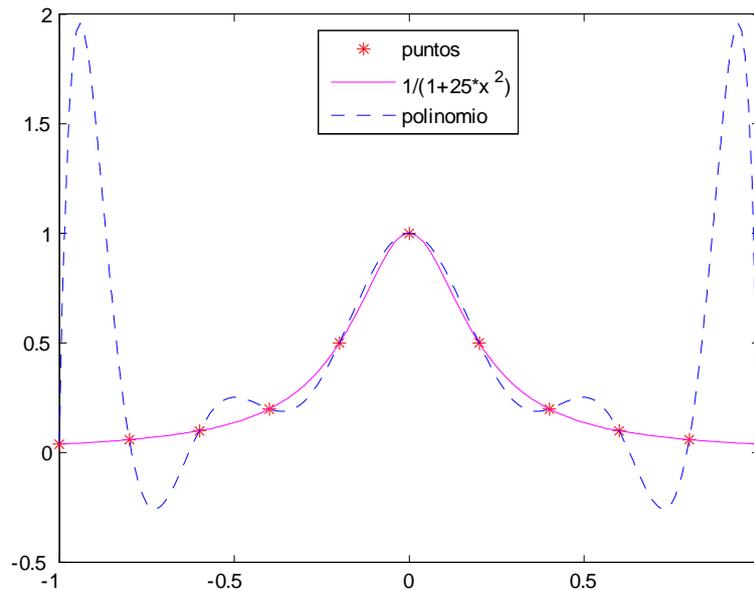
$$|f(x) - P_m(x)| \leq \frac{M}{(m + 1)!} \prod_{i=0}^m |x - x_i| \quad (x \in [a, b]).$$

Cuando interpolamos con un polinomio de grado elevado con nodos igualmente espaciados se puede producir el llamado **fenómeno de oscilación de Runge** en los extremos del intervalo. Este comportamiento del polinomio interpolador pone de manifiesto que no es siempre aconsejable trabajar con polinomios de grado elevado y particiones uniformes del intervalo de trabajo, prefiriendo el uso de los llamados nodos de Chebyshev en lugar de nodos equidistantes, la interpolación a trozos con polinomios de grado inferior, el uso de otro tipo de interpoladores como los splines, etc.

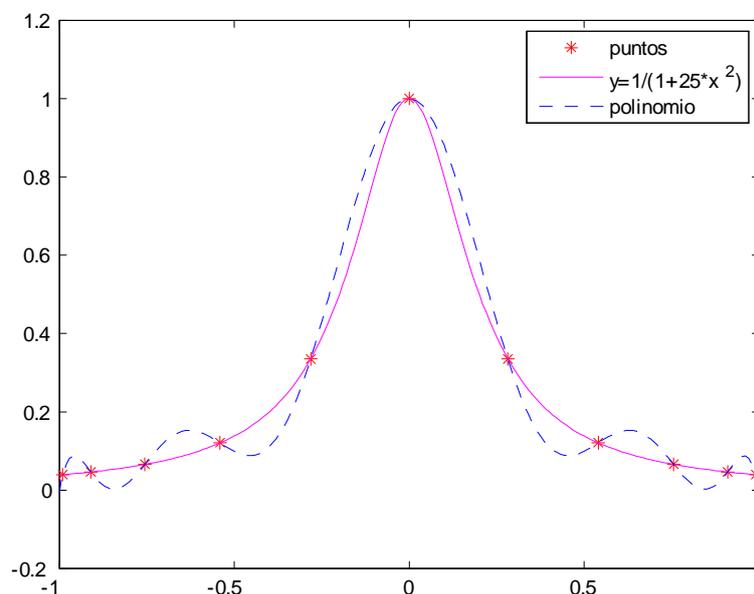
Consideremos la llamada función de Runge

$$f(x) := \frac{1}{1 + 25x^2} \quad (x \in [-1, 1]).$$

Tomando una partición uniforme del intervalo $[-1, 1]$ en 10 subintervalos, e interpolando utilizando los correspondientes puntos de la gráfica de f por un polinomio de grado menor o igual que 10, se obtiene un polinomio que comparamos con f en la gráfica siguiente. Nótese la oscilación del polinomio en las cercanías de los extremos del intervalo.



En la siguiente gráfica se muestra el polinomio de interpolación obtenido al usar los llamados nodos de Chebyshev ($x_i := \cos(\frac{2i+1}{m+1} \frac{\pi}{2})$ ($i = 0, 1, \dots, 10$)). Nótese la mejora en la aproximación de la función en los extremos del intervalo usando dichos nodos no equidistantes:



3.2. Interpolación segmentaria

La **interpolación segmentaria** consiste en dividir el intervalo donde se quiere aproximar la función en subintervalos y aplicar interpolación de grado menor a cada uno de ellos.

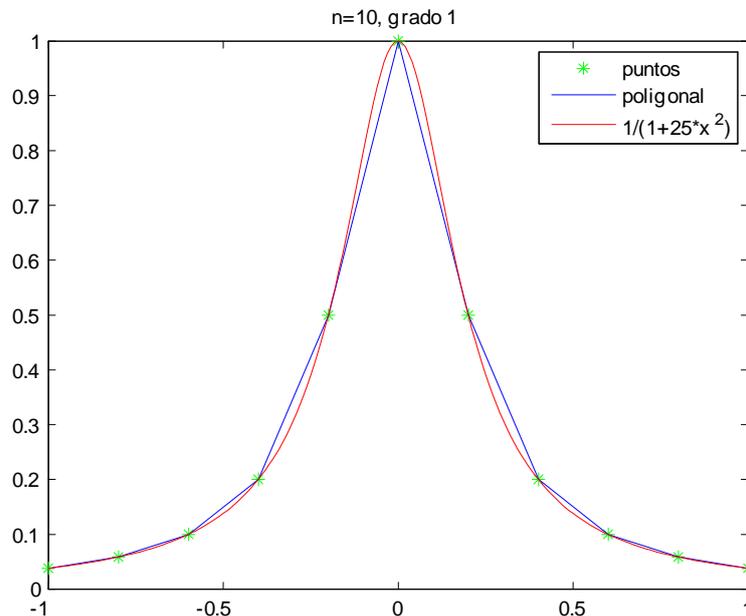
Sea $f : [a, b] \rightarrow \mathbb{R}$ una función dos veces derivable con derivada continua. Para $n \in \mathbb{N}$, consideramos una **partición uniforme** del intervalo $[a, b]$ en n subintervalos de igual longitud $h := (b - a)/n$, determinada por los puntos

$$x_i := a + ih \quad (i = 0, 1, \dots, n).$$

Sea $P(x) : [a, b] \rightarrow \mathbb{R}$ la poligonal obtenida al considerar en cada subintervalo de la partición $[x_i, x_{i+1}]$ la recta que une $(x_i, f(x_i))$ con $(x_{i+1}, f(x_{i+1}))$ (**interpolación segmentaria lineal**). Entonces

$$|f(x) - P(x)| \leq \frac{\max_{x \in [a, b]} |f''(x)|}{2} \frac{h^2}{4} \quad (x \in [a, b]).$$

La interpolación segmentaria lineal con $n = 10$ aplicada al ejemplo de Runge se representa en el gráfico siguiente.



La cota del error para interpolación segmentaria con polinomios de segundo grado es

$$\frac{\max_{x \in [a, b]} |f'''(x)|}{6} \frac{2h^3}{4} \quad (x \in [a, b]),$$

mientras que para interpolación segmentaria cúbica (polinomios de grado tres) la cota del error es

$$\frac{\max_{x \in [a, b]} |f^{(iv)}(x)|}{24} \frac{3h^4}{2} \quad (x \in [a, b]),$$

donde h es la longitud común de cada subintervalo de la partición uniforme.

3.3. Interpolación de Hermite. Polinomio de Taylor

3.3.1. Interpolación de Hermite

Sean $x_0, x_1 \in \mathbb{R}$ abscisas distintas y sean $y_i, y'_i \in \mathbb{R}$ ($i = 0, 1$). Nos plantearíamos el problema de obtener un polinomio $P(x)$, de grado menor o igual que 3, que cumpla:

$$\begin{aligned} P(x_0) &= y_0, P'(x_0) = y'_0, \\ P(x_1) &= y_1, P'(x_1) = y'_1. \end{aligned}$$

Es sencillo comprobar que dicho polinomio existe y es único. Si generalizamos las diferencias divididas, incorporando las relaciones

$$f[x_i, x_i] := y'_i \quad (i = 0, 1)$$

en (3.1), entonces el polinomio se puede expresar en términos de diferencias divididas como sigue:

$$\begin{aligned} P(x) &= f[x_0] + f[x_0, x_0](x - x_0) + f[x_0, x_0, x_1](x - x_0)^2 \\ &\quad + f[x_0, x_0, x_1, x_1](x - x_0)^2(x - x_1). \end{aligned}$$

Se puede probar la siguiente estimación del error: Sea $f : [x_0, x_1] \rightarrow \mathbb{R}$, 4 veces derivable con derivadas continuas, y sea $M = \max_{x \in [x_0, x_1]} |f^{(4)}(x)|$. Sea $P(x)$ el polinomio de interpolación correspondiente a los datos $(x_i, f(x_i))$, $(x_i, f'(x_i))$, $i = 0, 1$. Entonces

$$|f(x) - P(x)| \leq \frac{M}{4!} |x - x_0|^2 |x - x_1|^2 \quad (x \in [x_0, x_1]).$$

Se puede hablar de **interpolación segmentaria de Hermite** de forma análoga a la presentada anteriormente para el caso de disponer solo de información de la función en los nodos de una partición y no de las derivadas. Para aproximar una función $f : [a, b] \rightarrow \mathbb{R}$ derivable consideramos una partición del intervalo $[a, b]$ y en cada subintervalo obtenemos el polinomio de Hermite determinado por los valores de f y de f' en los extremos del mismo. Obtendríamos así una función aproximante de f que es derivable con derivada continua.

3.3.2. Polinomio de Taylor

Sea $f : [a, b] \rightarrow \mathbb{R}$ una función n veces derivable. El **polinomio de Taylor de grado n asociado a f en el punto a** es

$$P_n(x) := f(a) + f'(a)(x - a) + \frac{f''(a)}{2!}(x - a)^2 + \dots + \frac{f^{(n)}(a)}{n!}(x - a)^n$$

Si f es $n + 1$ veces derivable en $]a, b[$, entonces para cada $x \in [a, b]$ existe $\rho \in]a, x[$ tal que

$$f(x) = P_n(x) + \frac{f^{(n+1)}(\rho)}{(n+1)!}(x - a)^{n+1}.$$

Capítulo 4

Integración Numérica

Sea $f : [a, b] \rightarrow \mathbb{R}$ una función integrable Riemann. El cálculo de la integral definida

$$\int_a^b f(x) dx,$$

es un problema con gran número de aplicaciones. Cuando disponemos de una primitiva del integrando; es decir, de una función $F : [a, b] \rightarrow \mathbb{R}$ derivable con $F'(x) = f(x)$, $x \in [a, b]$, la *regla de Barrow* nos da el valor de I :

$$\int_a^b f(x) dx = F(b) - F(a).$$

No obstante, no siempre podemos obtener una primitiva del integrando, por lo que se impone recurrir a métodos discretos que aproximen la integral mediante sumas finitas de valores del integrando f en ciertos puntos del intervalo $[a, b]$. La misma idea se aplica cuando no conocemos la expresión matemática del integrando f pero disponemos de información sobre f en el intervalo de integración. A este respecto, recuérdese cómo se define la integral de Riemann como límite de sumas de Riemann.

4.1. Fórmulas de Integración de Newton-Cotes

Las fórmulas de Integración de Newton-Cotes se obtienen reemplazando el integrando $f(x)$ por un polinomio de interpolación $P_m(x)$ adecuado.

Para $m \in \mathbb{N}$, consideramos una **partición uniforme** del intervalo $[a, b]$ en m subintervalos de igual longitud $h := (b - a)/m$, determinada por los puntos

$$x_i := a + ih \quad (i = 0, 1, \dots, m).$$

Usaremos la notación

$$y_i := f(x_i) \quad (i = 0, 1, \dots, m).$$

Sea $P_m(x)$ el polinomio de interpolación de grado menor o igual que m tal que

$$P_m(x_i) = y_i \quad (i = 0, 1, \dots, m).$$

Por la fórmula de Lagrange sabemos que

$$P_m(x) = \sum_{i=0}^m y_i \left(\prod_{j=0, j \neq i}^m \frac{x - x_j}{x_i - x_j} \right).$$

Así,

$$\begin{aligned} \int_a^b P_m(x) dx &= \sum_{i=0}^m y_i \int_a^b \left(\prod_{j=0, j \neq i}^m \frac{x - x_j}{x_i - x_j} \right) dx \\ &\stackrel{x=a+ht}{=} h \sum_{i=0}^m y_i \int_0^m \left(\prod_{j=0, j \neq i}^m \frac{t - j}{i - j} \right) dt \end{aligned}$$

donde se ha realizado el cambio de variables indicado. Introduciendo la notación

$$\alpha_i := \int_0^m \left(\prod_{j=0, j \neq i}^m \frac{t - j}{i - j} \right) dt \quad (i = 0, 1, \dots, m),$$

se llega a

$$\int_a^b f(x) dx \simeq \int_a^b P_m(x) dx = h \sum_{i=0}^m y_i \alpha_i.$$

Nótese que los pesos α_i dependen únicamente de m .

Cálculos sencillos conducen a las fórmulas de Newton-Cotes que tabulamos a continuación.

m	α_i	Nombre
1	$\frac{1}{2} \quad \frac{1}{2}$	<i>Trapezios</i>
2	$\frac{1}{3} \quad \frac{4}{3} \quad \frac{1}{3}$	<i>Simpson</i>
3	$\frac{3}{8} \quad \frac{9}{8} \quad \frac{9}{8} \quad \frac{3}{8}$	<i>Regla $\frac{3}{8}$</i>

4.2. Fórmulas compuestas

Las fórmulas compuestas que vamos a obtener se basan en dividir el intervalo de integración en subintervalos de longitud más pequeña y en utilizar en ellos las fórmulas de Newton-Cotes asociadas a polinomios de interpolación de grado pequeño.

Consideramos de nuevo las particiones uniformes del intervalo de integración $[a, b]$ en $n \in \mathbb{N}$ subintervalos iguales de longitud $h := (b - a)/n$:

$$x_i := a + ih \quad (i = 0, 1, \dots, n).$$

4.2.1. Fórmula de los trapezios

En cada subintervalo de la partición, $[x_i, x_{i+1}]$, $i = 0, 1, \dots, n - 1$, aplicamos la fórmula de Newton-Cotes con $m = 1$; es decir, en dicho intervalo aproximamos la función f por el polinomio de grado uno que une los puntos (x_i, y_i) y (x_{i+1}, y_{i+1}) y lo integramos. Tras sumar el valor obtenido para cada subintervalo (aditividad de la integral respecto del intervalo de integración), da lugar a la fórmula

$$T_n := \frac{h}{2} \sum_{i=0}^{n-1} (y_i + y_{i+1}) = \frac{h}{2} (y_0 + 2y_1 + \dots + 2y_{n-1} + y_n).$$

Si $f \in C^2([a, b])$, entonces se verifica la siguiente cota del error:

$$\left| \int_a^b f(x) dx - T_n \right| \leq \max_{a \leq x \leq b} |f''(x)| \frac{(b-a)}{12} h^2.$$

4.2.2. Fórmula de Simpson

Se consideran una particiones de $[a, b]$ en un número *par* de subintervalos, es decir, $n = 2k$. En cada subintervalo de la forma $[x_{2i}, x_{2i+2}]$, $i = 0, 1, \dots, k-1$, aplicamos la fórmula de Newton-Cotes con $m = 2$ (aproximamos la función f por el polinomio de grado dos que pasa por los puntos (x_{2i}, y_{2i}) , (x_{2i+1}, y_{2i+1}) y (x_{2i+2}, y_{2i+2}) e integramos). Tras sumar las integrales correspondientes a cada subintervalo del tipo indicado se obtiene la fórmula

$$\begin{aligned} S_{2k} &:= \frac{h}{3} \sum_{i=0}^{k-1} (y_{2i} + 4y_{2i+1} + y_{2i+2}) \\ &= \frac{h}{3} (y_0 + 4(y_1 + \dots + y_{2k-1}) + 2(y_2 + \dots + y_{2k-2}) + y_{2k}). \end{aligned}$$

Si $f \in C^4([a, b])$, entonces

$$\left| \int_a^b f(x) dx - S_{2k} \right| \leq \max_{a \leq x \leq b} |f^{(iv)}(x)| \frac{(b-a)}{180} h^4.$$

Nota 20 (Integrales Impropias). *Las fórmulas anteriores pueden emplearse para aproximar integrales cuyo integrando presente algún problema en algún punto del intervalo de integración o que correspondan a integrales sobre intervalos de integración no acotados. Previamente realizamos un cambio de variable o aproximamos el intervalo de integración, etc.*

4.3. Integración de ecuaciones diferenciales

Sea $F : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ y sea $(a, c) \in \mathbb{R} \times \mathbb{R}$. Consideramos el siguiente **problema de Cauchy** (o problema de valores iniciales):

$$\begin{cases} y' = F(x, y) \\ y(a) = c \end{cases} \quad (4.1)$$

Una solución del problema anterior es una función $\varphi : I \rightarrow \mathbb{R}$ definida sobre un intervalo I , con $a \in I$, que es derivable en I y cumple

$$\varphi'(x) = F(x, \varphi(x)) \quad (x \in I); \quad \varphi'(a) = c.$$

Teorema 21 (Existencia y unicidad de solución). *Si F es una función continua que admite parcial respecto de y , F_y , continua, entonces existe una solución de (4.1). Además, la solución de (4.1) es única.*

Suponemos en adelante que estamos en las condiciones del teorema anterior. Sea φ la solución de (4.1). El objetivo es intentar aproximar el valor de φ en un punto $b > a$. Para ello consideramos una partición del intervalo $[a, b]$ en m subintervalos de igual longitud $h = (b-a)/m$ como en la sección anterior.

4.3.1. Método de Euler

Mediante la fórmula de recurrencia

$$y_0 = c; \quad y_{i+1} = y_i + hF(x_i, y_i) \quad (i = 0, 1, \dots, m-1),$$

se obtiene y_m que se tomará como aproximación de $\varphi(b)$.

4.3.2. Método de Euler mejorado

Se corresponde con la fórmula de recurrencia

$$y_0 = c; \quad y_{i+1} = y_i + \frac{h}{2}(F(x_i, y_i) + F(x_{i+1}, y_i + hF(x_i, y_i))) \quad (i = 0, 1, \dots, m-1).$$

El valor y_m se tomará como aproximación de $\varphi(b)$.

4.3.3. Método de Runge-Kutta de cuarto orden

La fórmula de recurrencia

$$y_0 = c; \quad y_{i+1} = y_i + \frac{1}{6}(k_{i1} + 2k_{i2} + 2k_{i3} + k_{i4}) \quad (i = 0, 1, \dots, m-1),$$

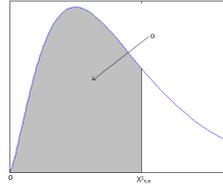
donde

$$\begin{aligned} k_{i1} &= hF(x_i, y_i) \\ k_{i2} &= hF\left(x_i + \frac{h}{2}, y_i + \frac{k_{i1}}{2}\right) \\ k_{i3} &= hF\left(x_i + \frac{h}{2}, y_i + \frac{k_{i2}}{2}\right) \\ k_{i4} &= hF(x_i + h, y_i + k_{i3}) \end{aligned}$$

proporciona y_m que se tomará como aproximación de $\varphi(b)$.

5.2. Tabla de la distribución χ^2 de Pearson

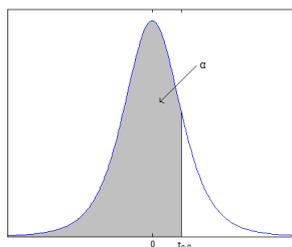
Los valores en el interior de la tabla nos proporcionan el número real $\chi_{n,\alpha}^2 \geq 0$ cumpliendo $P(\chi_n^2 < \chi_{n,\alpha}^2) = \alpha$. La siguiente figura ilustra la relación mencionada.



$n \backslash \alpha$	0.005	0.010	0.025	0.050	0.100	0.250	0.500	0.750	0.900	0.950	0.975	0.990	0.995
1	0.0000	0.0002	0.0010	0.0039	0.0158	0.1015	0.4549	1.3233	2.7055	3.8415	5.0239	6.6349	7.8794
2	0.0100	0.0201	0.0506	0.1026	0.2107	0.5754	1.3863	2.7726	4.6052	5.9915	7.3778	9.2103	10.5966
3	0.0717	0.1148	0.2158	0.3518	0.5844	1.2125	2.3660	4.1083	6.2514	7.8147	9.3484	11.3449	12.8382
4	0.2070	0.2971	0.4844	0.7107	1.0636	1.9226	3.3567	5.3853	7.7794	9.4877	11.1433	13.2767	14.8603
5	0.4117	0.5643	0.8312	1.1455	1.6103	2.6746	4.3515	6.6257	9.2364	11.0705	12.8325	15.0863	16.7496
6	0.6757	0.8721	1.2373	1.6354	2.2041	3.4546	5.3481	7.8408	10.6446	12.5916	14.4494	16.8119	18.5476
7	0.9893	1.2390	1.6899	2.1673	2.8331	4.2549	6.3458	9.0371	12.0170	14.0671	16.0128	18.4753	20.2777
8	1.3444	1.6465	2.1797	2.7326	3.4895	5.0706	7.3441	10.2189	13.3616	15.5073	17.5345	20.0902	21.9550
9	1.7349	2.0879	2.7004	3.3251	4.1082	5.8988	8.3428	11.3888	14.6837	16.9190	19.0228	21.6660	23.5894
10	2.1559	2.5582	3.2470	3.9403	4.8652	6.7372	9.3418	12.5489	15.9872	18.3070	20.4832	23.2093	25.1882
11	2.6032	3.0635	3.8157	4.5748	5.5778	7.5841	10.3410	13.7007	17.2750	19.6751	21.9200	24.7250	26.7568
12	3.0738	3.5706	4.4038	5.2260	6.3038	8.4384	11.3403	14.8454	18.5493	21.0261	23.3367	26.2170	28.2995
13	3.5650	4.1069	5.0088	5.8919	7.0415	9.2991	12.3398	15.9839	19.8119	22.3620	24.7356	27.6882	29.8195
14	4.0747	4.6604	5.6287	6.5706	7.7895	10.1653	13.3393	17.1169	21.0641	23.6848	26.1189	29.1412	31.3193
15	4.6009	5.2293	6.2621	7.2609	8.5468	11.0365	14.3389	18.2451	22.3071	24.9958	27.4884	30.5779	32.8013
16	5.1422	5.8122	6.9077	7.9616	9.3122	11.9122	15.3385	19.3689	23.5418	26.2962	28.8454	31.9999	34.2672
17	5.6972	6.4078	7.5642	8.6718	10.0852	12.7919	16.3382	20.4887	24.7690	27.5871	30.1910	33.4087	35.7185
18	6.2648	7.0149	8.2307	9.3905	10.8649	13.6753	17.3379	21.6049	25.9894	28.8693	31.5264	34.8053	37.1565
19	6.8440	7.6327	8.9065	10.1170	11.6509	14.5620	18.3377	22.7178	27.2036	30.1435	32.8523	36.1909	38.5823
20	7.4338	8.2604	9.5908	10.8508	12.4426	15.4518	19.3374	23.8277	28.4120	31.4104	34.1696	37.5662	39.9968
21	8.0337	8.8972	10.2829	11.5913	13.2396	16.3444	20.3372	24.9348	29.6151	32.6706	35.4789	38.9322	41.4011
22	8.6427	9.5425	10.9823	12.3380	14.0415	17.2396	21.3370	26.0393	30.8133	33.9244	36.7807	40.2894	42.7957
23	9.2604	10.1957	11.6886	13.0905	14.8480	18.1373	22.3369	27.1413	32.0069	35.1725	38.0756	41.6384	44.1813
24	9.8862	10.8564	12.4012	13.8484	15.6587	19.0373	23.3367	28.2412	33.1962	36.4150	39.3641	42.9798	45.5585
25	10.5197	11.5240	13.1197	14.6114	16.4734	19.9393	24.3366	29.3389	34.3816	37.6525	40.6465	44.3141	46.9279
26	11.1602	12.1981	13.8439	15.3792	17.2919	20.8434	25.3365	30.4346	35.5632	38.8851	41.9232	45.6417	48.2899
27	11.8076	12.8785	14.5734	16.1514	18.1139	21.7494	26.3363	31.5284	36.7412	40.1133	43.1945	46.9629	49.6449
28	12.4613	13.5647	15.3079	16.9279	18.9392	22.6572	27.3362	32.6205	37.9159	41.3371	44.4608	48.2782	50.9934
29	13.1211	14.2565	16.0471	17.7084	19.7677	23.5666	28.3361	33.7109	39.0875	42.5570	45.7223	49.5879	52.3356
30	13.7867	14.9535	16.7908	18.4927	20.5992	24.4776	29.3360	34.7997	40.2560	43.7730	46.9792	50.8922	53.6720

5.3. Tabla de la distribución t de Student

Los valores en el interior de la tabla nos proporcionan el número real $t_{n,\alpha} \geq 0$ cumpliendo $P(t_n < t_{n,\alpha}) = \alpha$. La siguiente figura ilustra la relación mencionada.



$n \backslash \alpha$	0.7500	0.9000	0.9500	0.9750	0.9900	0.9950	0.9995
1	1.0000	3.0777	6.3138	12.7062	31.8205	63.6567	636.6192
2	0.8165	1.8856	2.9200	4.3027	6.9646	9.9248	31.5991
3	0.7649	1.6377	2.3534	3.1824	4.5407	5.8409	12.9240
4	0.7407	1.5332	2.1318	2.7764	3.7469	4.6041	8.6103
5	0.7267	1.4759	2.0150	2.5706	3.3649	4.0321	6.8688
6	0.7176	1.4398	1.9432	2.4469	3.1427	3.7074	5.9588
7	0.7111	1.4149	1.8946	2.3646	2.9980	3.4995	5.4079
8	0.7064	1.3968	1.8595	2.3060	2.8965	3.3554	5.0413
9	0.7027	1.3830	1.8331	2.2622	2.8214	3.2498	4.7809
10	0.6998	1.3722	1.8125	2.2281	2.7638	3.1693	4.5869
11	0.6974	1.3634	1.7959	2.2010	2.7181	3.1058	4.4370
12	0.6955	1.3562	1.7823	2.1788	2.6810	3.0545	4.3178
13	0.6938	1.3502	1.7709	2.1604	2.6503	3.0123	4.2208
14	0.6924	1.3450	1.7613	2.1448	2.6245	2.9768	4.1405
15	0.6912	1.3406	1.7531	2.1314	2.6025	2.9467	4.0728
16	0.6901	1.3368	1.7459	2.1199	2.5835	2.9208	4.0150
17	0.6892	1.3334	1.7396	2.1098	2.5669	2.8982	3.9651
18	0.6884	1.3304	1.7341	2.1009	2.5524	2.8784	3.9216
19	0.6876	1.3277	1.7291	2.0930	2.5395	2.8609	3.8834
20	0.6870	1.3253	1.7247	2.0860	2.5280	2.8453	3.8495
21	0.6864	1.3232	1.7207	2.0796	2.5176	2.8314	3.8193
22	0.6858	1.3212	1.7171	2.0739	2.5083	2.8188	3.7921
23	0.6853	1.3195	1.7139	2.0687	2.4999	2.8073	3.7676
24	0.6848	1.3178	1.7109	2.0639	2.4922	2.7969	3.7454
25	0.6844	1.3163	1.7081	2.0595	2.4851	2.7874	3.7251
26	0.6840	1.3150	1.7056	2.0555	2.4786	2.7787	3.7066
27	0.6837	1.3137	1.7033	2.0518	2.4727	2.7707	3.6896
28	0.6834	1.3125	1.7011	2.0484	2.4671	2.7633	3.6739
29	0.6830	1.3114	1.6991	2.0452	2.4620	2.7564	3.6594
30	0.6828	1.3104	1.6973	2.0423	2.4573	2.7500	3.6460

5.4. Tablas para la F de Snedecor

$P(F_{m,n} < F_{m,n;0.95}) = 0.95$

$m \backslash n$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	243.91	245.95	248.01	249.05	250.1	251.14	252.2	253.25
2	18.513	19.000	19.164	19.247	19.296	19.33	19.353	19.371	19.385	19.396	19.413	19.429	19.446	19.454	19.462	19.471	19.479	19.487
3	10.128	9.5521	9.2766	9.1172	9.0135	8.9406	8.8867	8.8452	8.8123	8.7855	8.7446	8.7029	8.6602	8.6385	8.6166	8.5944	8.5720	8.5494
4	7.7086	6.9443	6.5914	6.3882	6.2561	6.1631	6.0942	6.0410	5.9988	5.9644	5.9117	5.8578	5.8025	5.7744	5.7459	5.7170	5.6877	5.6581
5	6.6079	5.7861	5.4095	5.1922	5.0503	4.9503	4.8759	4.8183	4.7725	4.7351	4.6777	4.6188	4.5581	4.5272	4.4957	4.4638	4.4314	4.3985
6	5.9874	5.1433	4.7571	4.5337	4.3874	4.2839	4.2067	4.1468	4.0990	4.0600	3.9999	3.9381	3.8742	3.8415	3.8082	3.7743	3.7398	3.7047
7	5.5914	4.7374	4.3468	4.1203	3.9715	3.8660	3.7870	3.7257	3.6767	3.6365	3.5747	3.5107	3.4445	3.4105	3.3758	3.3404	3.3043	3.2674
8	5.3177	4.4590	4.0662	3.8379	3.6875	3.5806	3.5005	3.4381	3.3881	3.3472	3.2839	3.2184	3.1503	3.1152	3.0794	3.0428	3.0053	2.9669
9	5.1174	4.2565	3.8625	3.6331	3.4817	3.3738	3.2927	3.2296	3.1789	3.1373	3.0729	3.0061	2.9365	2.9005	2.8637	2.8259	2.7872	2.7475
10	4.9646	4.1028	3.7083	3.4780	3.3258	3.2172	3.1355	3.0717	3.0204	2.9782	2.9130	2.8450	2.7740	2.7372	2.6996	2.6609	2.6211	2.5801
12	4.7472	3.8853	3.4903	3.2592	3.1059	2.9961	2.9134	2.8486	2.7964	2.7534	2.6866	2.6169	2.5436	2.5055	2.4663	2.4259	2.3842	2.3410
15	4.5431	3.6823	3.2874	3.0556	2.9013	2.7905	2.7066	2.6408	2.5876	2.5437	2.4753	2.4033	2.3275	2.2878	2.2468	2.2043	2.1601	2.1141
20	4.3512	3.4928	3.0984	2.8661	2.7109	2.5990	2.5140	2.4471	2.3928	2.3479	2.2776	2.2033	2.1242	2.0825	2.0391	1.9938	1.9464	1.8963
24	4.2597	3.4028	2.9988	2.7663	2.6207	2.5082	2.4226	2.3551	2.3002	2.2547	2.1834	2.1077	2.0267	1.9838	1.9390	1.8920	1.8424	1.7896
30	4.1709	3.3158	2.9223	2.6896	2.5336	2.4205	2.3343	2.2662	2.2107	2.1646	2.0921	2.0148	1.9317	1.8874	1.8409	1.7918	1.7396	1.6835
40	4.0847	3.2317	2.8387	2.6060	2.4495	2.3359	2.2490	2.1802	2.1240	2.0772	2.0035	1.9245	1.8389	1.7929	1.7444	1.6928	1.6373	1.5766
60	4.0012	3.1504	2.7581	2.5252	2.3683	2.2541	2.1665	2.0970	2.0401	1.9926	1.9174	1.8364	1.7480	1.7001	1.6491	1.5943	1.5343	1.4673
120	3.9201	3.0718	2.6802	2.4472	2.2899	2.1750	2.0868	2.0164	1.9588	1.9105	1.8337	1.7505	1.6587	1.6084	1.5543	1.4952	1.4290	1.3519

$P(F_{m,n} < F_{m,n;0.99}) = 0.99$

$m \backslash n$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120
1	4052.2	4999.5	5403.4	5624.6	5763.7	5859.0	5928.4	5981.1	6022.5	6055.9	6106.3	6157.3	6208.7	6234.6	6260.7	6286.8	6313.0	6339.4
2	98.503	99.000	99.166	99.249	99.299	99.333	99.357	99.374	99.388	99.399	99.416	99.432	99.449	99.448	99.458	99.474	99.483	99.491
3	34.116	30.816	29.457	28.71	28.237	27.911	27.672	27.489	27.345	27.229	27.051	26.872	26.690	26.598	26.505	26.411	26.316	26.221
4	21.198	18.000	16.694	15.977	15.522	15.207	14.976	14.799	14.659	14.546	14.374	14.198	14.020	13.929	13.838	13.745	13.652	13.558
5	16.258	13.274	12.060	11.392	10.967	10.672	10.456	10.289	10.1578	10.051	9.8883	9.7222	9.526	9.4665	9.3793	9.2912	9.2020	9.1118
6	13.745	10.925	9.7795	9.1483	8.7459	8.4661	8.2600	8.1017	7.9761	7.8741	7.7183	7.5590	7.3958	7.3127	7.2285	7.1432	7.0567	6.9690
7	12.246	9.5466	8.4513	7.8466	7.4604	7.1914	6.9928	6.8400	6.7188	6.6201	6.4691	6.3143	6.1554	6.0743	5.9920	5.9084	5.8236	5.7373
8	11.259	8.6491	7.5910	7.0061	6.6318	6.3707	6.1776	6.0289	5.9106	5.8143	5.6667	5.5151	5.3591	5.2793	5.1981	5.1156	5.0316	4.9461
9	10.561	8.0215	6.9919	6.4221	6.0569	5.8018	5.6129	5.4671	5.3511	5.2565	5.1114	4.9621	4.8080	4.7290	4.6486	4.5666	4.4831	4.3978
10	10.044	7.5594	6.5523	5.9943	5.6363	5.3858	5.2001	5.0567	4.9434	4.8491	4.7059	4.5581	4.4054	4.3269	4.2469	4.1653	4.0819	3.9965
12	9.3302	6.9266	5.9525	5.4120	5.0643	4.8206	4.6395	4.4994	4.3875	4.2961	4.1533	4.0096	3.8584	3.7805	3.7008	3.6192	3.5355	3.4494
15	8.6831	6.3589	5.4170	4.8932	4.5556	4.3183	4.1415	4.0045	3.8948	3.8049	3.6622	3.5222	3.3719	3.2940	3.2141	3.1319	3.0471	2.9595
20	8.0960	5.8489	4.9382	4.4307	4.1027	3.8714	3.6987	3.5644	3.4567	3.3682	3.2311	3.0880	2.9377	2.8594	2.7785	2.6947	2.6077	2.5168
24	7.8229	5.6136	4.7181	4.2184	3.8951	3.6667	3.4959	3.3629	3.2560	3.1681	3.0316	2.8887	2.7380	2.6594	2.5773	2.4923	2.4035	2.3100
30	7.5625	5.3903	4.5097	4.0179	3.6990	3.4735	3.3045	3.1726	3.0665	2.9791	2.8431	2.7002	2.5487	2.4689	2.3860	2.2992	2.2079	2.1108
40	7.3141	5.1785	4.3126	3.8283	3.5138	3.2910	3.1238	2.9930	2.8876	2.8005	2.6648	2.5216	2.3689	2.2880	2.2034	2.1142	2.0194	1.9172
60	7.0771	4.9774	4.1259	3.6490	3.3389	3.1187	2.9530	2.8233	2.7185	2.6318	2.4961	2.3523	2.1978	2.1154	2.0285	1.9360	1.8363	1.7263
120	6.8509	4.7865	3.9491	3.4795	3.1735	2.9559	2.7918	2.6629	2.5586	2.4721	2.3363	2.1915	2.0346	1.9500	1.8600	1.7628	1.6557	1.5330

Bibliografía

- [1] Amat, S.; Aràndiga, F.; y otros, *Aproximació Numèrica*, Publicacions de la Universitat de València., 2002.
- [2] Burden, R. L.; Faires, J. D., *Análisis Numérico*, Thomson Learning, 2003.
- [3] Canavos, George C., *Probabilidad y Estadística. Aplicaciones y Métodos*, McGraw-Hill, México, 1988.
- [4] Conte, S. D.; Boor, C. De, *Análisis Numérico Elemental*, McGraw-Hill, México, 1974.
- [5] Dahlquist, G.; Björck, A., *Numerical Methods*, Prentice-Hall, Inc., 1974.
- [6] Devore, Jay L., *Probabilidad y Estadística para Ingeniería y Ciencias*, Ed. Thomson, México, 2001.
- [7] García, A.; Hernández V.; otros autores, *Estadística I*, Universidad Nacional de Educación a Distancia, 1994.
- [8] Milton, J. S.; Tsokos, J. O., *Estadística para Biología y Ciencias de la Salud*, McGraw Hill Interamericana, 1989.
- [9] Quesada, V.; Isidoro, A.; López, L. A., *Curso y Ejercicios de Estadística*, Alhambra Universidad, 1989.
- [10] Spiegel, M. R., *Teoría y Problemas de Estadística*, McGraw-Hill, 1970.