

# Índice

<b>1. Conceptos básicos</b>	<b>1</b>
1.1. Introducción . . . . .	1
1.2. Concepto de población . . . . .	4
1.3. Concepto de muestra aleatoria . . . . .	7
1.4. Estadísticos y su distribución en el muestreo . . . . .	15
<b>2. Distribución muestral</b>	<b>33</b>
2.1. Introducción . . . . .	33
2.2. Distribución muestral y sus características . . . . .	34
2.3. Distribución en el muestreo y comportamiento asintótico . . . . .	36
2.4. Distribución en el muestreo de los momentos muestrales . . . . .	39
2.5. Comportamiento asintótico de los cuantiles muestrales . . . . .	41
2.6. Comportamiento asintótico de los momentos muestrales . . . . .	45
<b>3. Poblaciones normales</b>	<b>59</b>
3.1. Introducción . . . . .	59
3.2. Distribución de la media y varianza muestrales. Distribución $\chi^2$ de Pearson . . . . .	60
3.3. Estadístico de Student. Distribución $t$ de Student . . . . .	67
3.4. Distribución de la diferencia de medias muestrales . . . . .	71
3.5. Distribución del cociente de cuasivarianzas muestrales. Distribución $F$ de Snedecor . . . . .	75
3.6. El coeficiente de correlación muestral . . . . .	79
<b>4. Intervalos de confianza</b>	<b>91</b>
4.1. Introducción . . . . .	91
4.2. Definición de intervalo de confianza . . . . .	92
4.3. Método de la cantidad pivotal para la construcción de intervalos de confianza . . . . .	96

4.4.	Método de Neyman para la construcción de intervalos de confianza . . . . .	101
4.5.	Intervalos de confianza para los parámetros de distribuciones normales . . . . .	114
4.6.	Intervalos de confianza basados en distribuciones asintóticas .	120
<b>5.</b>	<b>Estimación Puntual</b>	<b>139</b>
5.1.	Introducción . . . . .	139
5.2.	Comparación de estimadores: Riesgo y error cuadrático medio	142
5.3.	Propiedades deseables de los estimadores . . . . .	151
5.3.1.	Estimadores insesgados . . . . .	151
5.3.2.	Estimadores consistentes . . . . .	155
5.3.3.	Estimadores invariantes . . . . .	157
5.4.	Estadísticos suficientes . . . . .	161
5.4.1.	Estadísticos suficientes minimales . . . . .	170
5.5.	Criterios de selección de estimadores . . . . .	175
5.5.1.	Estimadores minimax . . . . .	176
5.5.2.	Estimadores Bayes . . . . .	179
<b>6.</b>	<b>Estimadores de mínimo riesgo</b>	<b>209</b>
6.1.	Introducción . . . . .	209
6.2.	Estimadores centrados de mínima varianza . . . . .	210
6.3.	Acotaciones de la varianza de un estimador . . . . .	223
6.4.	Estimación de parámetros de posición y escala . . . . .	238
6.4.1.	Parámetros de posición . . . . .	238
6.4.2.	Parámetros de escala . . . . .	244
6.4.3.	Parámetros de posición y escala . . . . .	249
<b>7.</b>	<b>Métodos de estimación</b>	<b>279</b>
7.1.	Introducción . . . . .	279
7.2.	El método de los momentos . . . . .	279
7.3.	El método de la máxima verosimilitud . . . . .	281
7.4.	Propiedades asintóticas de los estimadores de máxima verosimilitud . . . . .	292
7.5.	Estimación Bayesiana . . . . .	303
7.6.	Estimación mínimo cuadrática . . . . .	306
7.7.	Conclusión . . . . .	308

<b>8. Contraste de hipótesis</b>	<b>327</b>
8.1. Introducción . . . . .	327
8.2. Planteamiento general de los contrastes de hipótesis . . . . .	328
8.3. Contraste de hipótesis simple frente a simple . . . . .	336
8.4. Contraste de hipótesis unilaterales y bilaterales . . . . .	347
<b>9. Métodos de contraste</b>	<b>387</b>
9.1. Introducción . . . . .	387
9.2. Tests de razón de verosimilitudes . . . . .	388
9.3. Relación entre estimación confidencial y contraste de hipótesis	400
<b>10. Contrastes globales</b>	<b>429</b>
10.1. Introducción . . . . .	429
10.2. Contrastes $\chi^2$ de bondad del ajuste . . . . .	431
10.2.1. Hipótesis simple . . . . .	431
10.2.2. Hipótesis compuesta . . . . .	438
10.3. Contraste $\chi^2$ de homogeneidad . . . . .	445
10.4. Contraste $\chi^2$ de independencia . . . . .	450
10.5. Contraste de Kolmogorov-Smirnov de bondad del ajuste . . .	454
10.6. Contraste de homogeneidad de Kolmogorov-Smirnov . . . . .	461
10.7. Contrastes de posición . . . . .	467
10.7.1. El test de los signos . . . . .	468
10.7.2. El test de Wilcoxon de los rangos signados . . . . .	470
10.7.3. El test de la mediana . . . . .	473
10.7.4. El test de Mann-Whitney . . . . .	476
10.8. Contrastes de independencia . . . . .	478
10.8.1. El test $\tau$ de Kendall . . . . .	478
10.8.2. El test del coeficiente de correlación entre rangos de Spearman . . . . .	481
10.8.3. El test de rachas . . . . .	483

### 3.3. Estadístico de Student. Distribución $t$ de Student

Saber que  $\bar{X}$  tiene distribución en el muestreo  $N(\mu, \sigma/\sqrt{n})$ , o equivalentemente que

$$\sqrt{n} \frac{\bar{X} - \mu}{\sigma} \quad \text{tiene distribución } N(0, 1)$$

puede resultar de poca utilidad si la varianza poblacional  $\sigma^2$  es un parámetro desconocido de la distribución teórica, ya que no se puede, entonces, usar esta conclusión para hacer previsiones acerca de la diferencia  $\bar{X} - \mu$ . En tal caso, cabe pensar que el resultado no será muy distinto si se sustituye  $\sigma$  por el valor de la desviación típica muestral,  $s$ , puesto que, al menos para muestras grandes,  $\sigma^2$  y  $s^2$  tendrán valores similares. Tal idea llevó a Student a considerar el estadístico:

$$t = \sqrt{n-1} \frac{\bar{X} - \mu}{s}.$$

La sustitución de  $\sqrt{n}$  por  $\sqrt{n-1}$  quedará explicada un poco más adelante, aunque de momento, si se quiere reforzar la analogía, se puede expresar

$$t = \sqrt{n} \frac{\bar{X} - \mu}{S},$$

donde  $S^2$  es la cuasivarianza muestral ( $S^2 = ns^2/(n-1)$ ) que, a efectos de dar una aproximación de  $\sigma^2$ , tiene tantas ventajas, si no más, que  $s^2$ .

En cualquier caso, para poder usar el estadístico  $t$  en sustitución de  $\sqrt{n}(\bar{X} - \mu)/\sigma$ , la cuestión se centra en determinar su distribución en el muestreo (confiando en que el resultado no dependa de  $\sigma^2$ , pues, de lo contrario, no se habría adelantado nada). Para ello, parece natural prescindir de  $\mu$  y adoptar el siguiente planteamiento:

Sean  $X, X_1, X_2, \dots, X_n$ ,  $n+1$  variables aleatorias independientes, con distribución  $N(0, \sigma)$ ; sea  $Y = \sum_{i=1}^n X_i^2$ , y  $U = \sqrt{Y/n}$ . Se pretende hallar la distribución de  $t = X/U$ .

La distribución de  $Y/\sigma^2$  es  $\chi_n^2$ , así que la densidad de  $Y$  será

$$\frac{1}{2^{n/2} \Gamma(\frac{n}{2}) \sigma^n} y^{\frac{n}{2}-1} e^{-y/(2\sigma^2)} \quad \text{para } y > 0,$$

con lo cual la densidad de  $U$  es

$$\frac{n^{n/2}}{2^{(n-2)/2} \Gamma(\frac{n}{2}) \sigma^n} u^{n-1} e^{-nu^2/(2\sigma^2)} \quad \text{para } u > 0.$$

Como  $X$  y  $U$  son independientes, su densidad conjunta será

$$K e^{-x^2/(2\sigma^2)} u^{n-1} e^{-nu^2/(2\sigma^2)} = K u^{n-1} e^{-(x^2+nu^2)/(2\sigma^2)}$$

para  $x \in \mathbb{R}$  y  $u > 0$ , siendo  $K = \frac{1}{\sqrt{2\pi}} \frac{n^{n/2}}{2^{(n-2)/2} \Gamma(\frac{n}{2}) \sigma^{n+1}}$ .

El cambio bidimensional  $z = U$ ,  $t = X/U$ , da como densidad de  $(z, t)$

$$K z^n e^{-z^2(t^2+n)/(2\sigma^2)} \quad \text{para } t \in \mathbb{R} \text{ y } z > 0;$$

con lo cual la densidad marginal de  $t$  resulta ser

$$\begin{aligned} K \int_0^\infty z^n e^{-z^2(t^2+n)/(2\sigma^2)} dz &= \frac{K}{2} \int_0^\infty v^{(n-1)/2} e^{-v(t^2+n)/(2\sigma^2)} dv \\ &= \frac{K}{2} \left( \frac{t^2+n}{2\sigma^2} \right)^{-(n+1)/2} \Gamma\left(\frac{n+1}{2}\right) = \frac{1}{\sqrt{n\pi}} \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \left(1 + \frac{t^2}{n}\right)^{-(n+1)/2} \end{aligned}$$

para  $t \in \mathbb{R}$  (después de sustituir la constante  $K$  y simplificar).

Nuevamente la situación merece un nombre propio:

Si  $X, X_1, X_2, \dots, X_n$  son variables aleatorias independientes y con distribución  $N(0, \sigma)$ , la distribución de

$$\frac{X}{\sqrt{\frac{1}{n} \sum_{i=1}^n X_i^2}},$$

de densidad

$$\frac{1}{\sqrt{n\pi}} \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \left(1 + \frac{t^2}{n}\right)^{-(n+1)/2} \quad \text{para } t \in \mathbb{R},$$

se denomina distribución  $t$  de Student con  $n$  grados de libertad.

Suele decirse que la distribución  $t_n$  es la de una  $N(0, 1)$  dividida por la raíz cuadrada de una  $\chi_n^2$  dividida por sus grados de libertad, ambas independientes entre sí. Lo cual refleja la propiedad fundamental de la distribución de Student: su independencia de la varianza  $\sigma^2$  de las variables que intervienen en su definición.

Antes de examinar otras propiedades de la distribución  $t$ , se puede ver que resuelve el problema que ha suscitado su definición.

**Teorema de Student:** Si  $(X_1, X_2, \dots, X_n)$  es una muestra aleatoria simple de una población  $N(\mu, \sigma)$ , el estadístico de Student:

$$\sqrt{n-1} \frac{\bar{X} - \mu}{s} = \sqrt{n} \frac{\bar{X} - \mu}{S}$$

tiene distribución  $t$  de Student con  $n-1$  grados de libertad.

En efecto,

$$\sqrt{n-1} \frac{\bar{X} - \mu}{s} = \frac{\sqrt{n}(\bar{X} - \mu)/\sigma}{\sqrt{\frac{1}{n-1} ns^2/\sigma^2}}$$

y basta observar que  $\sqrt{n}(\bar{X} - \mu)/\sigma$  tiene distribución  $N(0, 1)$ , mientras que, según el teorema de Fisher,  $ns^2/\sigma^2$  tiene la misma distribución que  $\sum_{i=1}^{n-1} X_i'^2$ , siendo  $X_1', \dots, X_{n-1}'$  variables  $N(0, 1)$ , independientes entre sí y de la media muestral  $\bar{X}$ . Por tanto,  $\sqrt{n-1}(\bar{X} - \mu)/s$  tiene distribución  $t_{n-1}$ .

Cuando  $\sigma^2$  sea desconocida, el resultado permite juzgar que valores son previsibles para  $\bar{X} - \mu$ , utilizando  $S$  en lugar de  $\sigma$  y la distribución  $t_{n-1}$  en lugar de la  $N(0, 1)$ . Véase el ejercicio 3.3.

Acerca de la distribución  $t$  de Student, cabe señalar que, al igual que la distribución  $\chi^2$ , depende de un único parámetro  $n$  (el número de sumandos que intervienen en la suma del denominador) que nuevamente se denomina “grados de libertad” de la distribución. En cambio, la distribución de Student tiene como soporte el intervalo  $(-\infty, \infty)$ , siendo su densidad simétrica respecto al origen, de manera que su representación gráfica (para  $n = 10$ ) es de la forma:

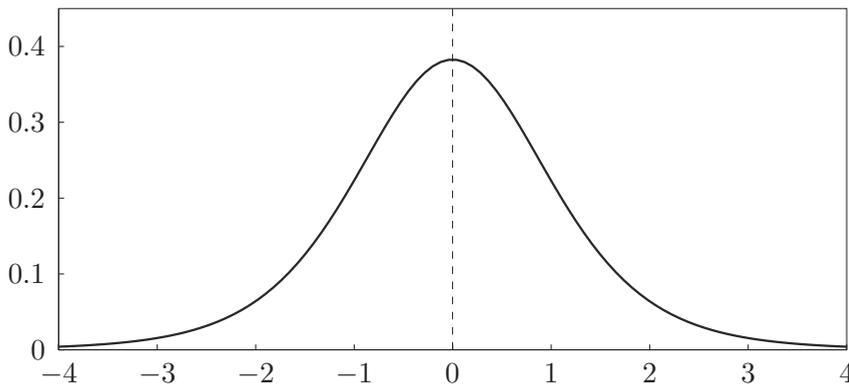


Figura 3.2: Densidad  $t$  de Student