Tema 6. Contrastes paramétricos (I)

Dr. David Castilla Espino

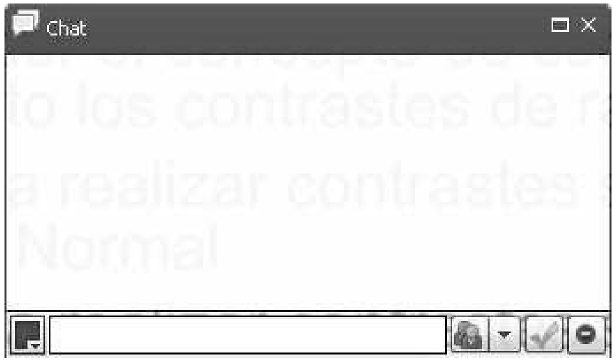
CA UNED de Huelva, "Profesor Dr. José Carlos Vílchez Martín"

Introducción

- Bienvenida
- **Objetivos** pedagógicos:
  - Conocer el concepto de hipótesis paramétrica
  - Comprender el concepto de contraste de significación y en concreto los contrastes de razón de verosimilitud.
  - Aprender a realizar contrastes…
    - sobre la media de una población Normal
    - sobre la varianza de una población Normal
    - sobre la proporción

Introducción

- Comunicación:
  - Online: Chat de webconferencia 20/12/2012.



  - Offline: Foro del grupo de tutoría 5.
- Referencias:
  - Casas, J.M. & P. Gutiérrez (2011): Estadística II: Inferencia Estadística. Editorial Universitaria Ramón Areces, Madrid. 231-299.
  - Novales (1997), Ruiz Maya & Martín Pliego (1999),...

Introducción

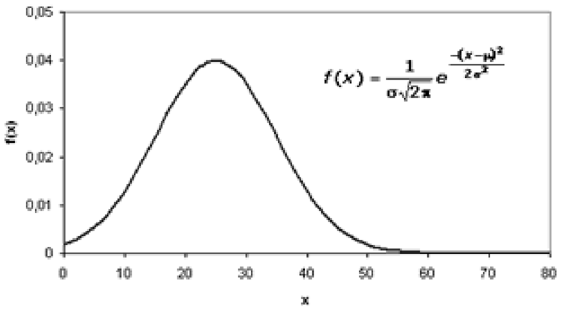
- Antecedentes
  - Inferencia estadística (Tema 1). Utiliza datos muestrales para llevar a cabo estimaciones, tomar decisiones, realizar predicciones, comprobar hipótesis u otras generalizaciones acerca de un conjunto de datos más grande denominado población.

Control de calidad

Dr. David Castilla Espino
CA UNED Huelva

Introducción

- Antecedentes:
  - Procedimientos de la inferencia estadística:
    - Estimación de parámetros poblacionales (Temas 2-4).
    - Contrastación de hipótesis estadística acerca de la población (Temas 5-7).
  - Contrastes de hipótesis (Tema 5)
    - Paramétricos. Contrastan hipótesis sobre el valor que toman los parámetros de distribuciones poblacionales conocidas (Tema 6). Ej. Contrastar hipótesis sobre μ y σ de una familia de distribuciones Normal.



Introducción

- Antecedentes:
    - No paramétricos. Contrastan hipótesis sobre otras características de la distribución distintas de los parámetros: forma, localización, aleatoriedad,… (Tema 7).
- Contrastes de hipótesis óptimos (Tema 5)
    - Lema de Neyman-Pearson: Contraste de hipótesis simples de máxima potencia.
    - Contrastes uniformemente más potentes de hipótesis compuestas.

Contrastes de significación

- Se basan para la definición de la regla de decisión en el tamaño de la discrepancia existente entre el valor del parámetro formulado bajo al hipótesis nula y el valor que le correspondería según la información que proporciona la muestra.
- Su procedimiento de resolución sigue en general las fases de realización de un contraste establecidas en el Tema 5.
- Considérese el ejemplo de un contraste de hipótesis sobre la media de una población Normal con varianza desconocida.
  - Formulación de las hipótesis

    H_0: $\mu \leq 4$

    H_1: $\mu > 4$
  - Obtención del estadístico de prueba

$$D = \frac{\bar{x} - \mu_0}{S / \sqrt{n}}$$

Contrastes de significación

- Selección del nivel de significación: α=0,05
- Determinación de la región crítica (n = 20)

  dado que:

$$\frac{\bar{x}-\mu_0}{S/\sqrt{n}} \rightarrow t_{n-1}$$

  se cumple que bajo el supuesto de la H_0

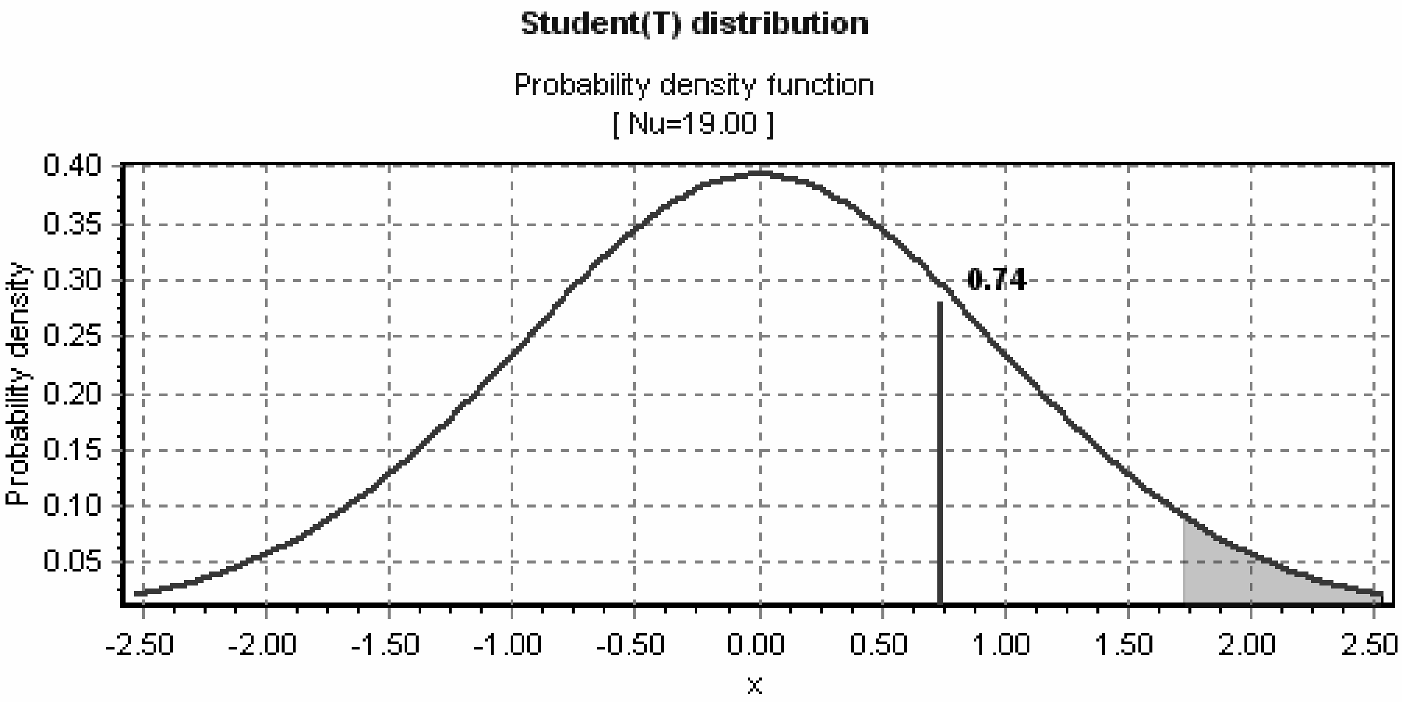
$$P\left(D_{\exp} > d_{\alpha}\right) = P\left(t_{\exp} > t_{0,05}\right) = 0,05 \Rightarrow t_{0,05} = 1,729 \Rightarrow RC : t_{\exp} > 1,729$$

- Selección aleatoria de la muestra y cálculo del estadístico de prueba (media=4,5 y cuasidesviación típica = 3)

$$D_{\exp} = \frac{\bar{x}-\mu_0}{S/\sqrt{n}} = \frac{4,5-4}{3/\sqrt{20}} = 0,745$$

- Decisión: $D_{exp} < d_{\alpha}$ ➔ Acepto H_0

Contrastes de significación



Contrastes de significación

- Los **contrastes de razón de verosimilitud** son un caso particular de los contrastes de significación que miden la discrepancia mediante la ratio de distintas verosimilitudes.
- Los contrastes de razón verosimilitud presentan buenas propiedades cuando el tamaño muestral es grande.
- Este tipo de contrastes serán los que empleemos a lo largo de este curso.
- Ejemplo:
  Sea (X_1,…, X_n) una muestra aleatoria de una población Normal (μ; σ_0) y deseamos contrastar las hipótesis:

$$H_0: \mu = \mu_0$$

$$H_1: \mu \neq \mu_0$$

Contrastes de significación

La función de verosimilitud para una muestra de tamaño n viene dada por:

$$L(x;\mu,\sigma_0) = \left(2\pi\sigma_0^2\right)^{-n/2} e^{-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}(x_i-\mu)^2}$$

El estadístico de razón de verosimilitud [λ(x)] se define por medio del siguiente cociente:

$$\lambda(x) = \frac{L_0^*(x;\mu,\sigma_0)}{L^*(x;\mu,\sigma_0)} = \frac{\max_{\mu=\mu_0} L(x;\mu,\sigma_0)}{\max_{\mu} L(x;\mu,\sigma_0)} = e^{-\frac{n}{2\sigma_0^2}(\bar{x}-\mu_0)^2}$$

dado que:

$$L_0^*(x;\mu,\sigma_0) = \left(2\pi\sigma_0^2\right)^{-n/2} e^{-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}(x_i-\mu_0)^2}$$

y que:

$$L^*(x;\mu,\sigma_0) = \left(2\pi\sigma_0^2\right)^{-n/2} e^{-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

Contrastes de significación

Para determinar la región crítica:

$$P\left[\lambda(x) < \lambda_0\right] = \alpha \Rightarrow P\left[e^{-\frac{n}{2\sigma_0^2}(\bar{x}-\mu_0)^2} < \lambda_0\right] = \alpha \Rightarrow$$

$$P\left[-\frac{n}{2\sigma_0^2}(\bar{x}-\mu_0)^2 < \ln\lambda_0\right] = \alpha \Rightarrow P\left[\left(\frac{\bar{x}-\mu_0}{\sigma_0/\sqrt{n}}\right)^2 < \underbrace{-2\ln\lambda_0}_{\lambda_0'}\right] = \alpha \Rightarrow$$

$$\Rightarrow P\left(\left|\frac{\bar{x}-\mu_0}{\sigma_0/\sqrt{n}}\right| < \sqrt{\lambda_0'}\right) = \alpha$$

Contrastes de significación

dado que:

$$\frac{\bar{x}-\mu_0}{\sigma_0/\sqrt{n}} \to N(0,1)$$

se verifica que:

$$P\left(|Z| < \sqrt{\lambda_0'}\right) = \alpha$$

de modo que por simetría de la normal se cumple cuando H_0 es cierta que la región crítica (RC) es:

$$P\left(|Z_{\exp}| < Z_{\alpha/2}\right) = \alpha \Rightarrow RC: Z_{\exp} > Z_{\alpha/2} \text{ y } Z_{\exp} < -Z_{\alpha/2}.$$

Contrastes de significación

- El esquema de aplicación del contraste de máxima verosimilitud presentado en las diapositivas anteriores se puede repetir de forma análogo para otros casos como se puende comprobar en el manual al que me remito.
- Los resultados de la aplicación de este esquema para los distintos contrastes abordados en el tema 6 se encuentran resumidos en el **cuadro resumen** disponible de las páginas **294-299**.

Contrastes sobre la media

Contraste la H_0 ($\mu = 3$) frente a la alternativa H_1 ($\mu \neq 3$) en una distribución $N(\mu,\sigma)$, con un nivel de significación (α) igual a 0,01, mediante una muestra aleatoria simple de tamaño 18, cuyo resultado es:

4,19	3,98	3,65	3,25	3,86	4,03	3,93	3,83	3,98
4,08	3,93	3,75	3,93	3,98	3,84	4,01	4,16	4,01

Dr. David Castilla Espino
CA UNED Huelva

Contrastes sobre la media

SOLUCIÓN:

1- Formulación de hipótesis

H_0: μ = 3

H_1: μ ≠ 3

2- Elegir y calcular el estadístico de prueba

$$t_{\exp} = \frac{\bar{x} - \mu_0}{S / \sqrt{n}} = \frac{3{,}911 - 3}{0{,}04486 / \sqrt{18}} = 18{,}247$$

3- Determinación de la región crítica

$$P\left(\left|t_{17}\right| > t_{\alpha/2}\right) = 0{,}01 \Rightarrow t_{0,005} = 2{,}898 \Rightarrow t_{\exp} > 2{,}898 \text{ y } t_{\exp} < -2{,}898.$$

4- Decisión: $t_{exp} > t_{0,005}$ ➔ Rechazo la H_0

Contrastes sobre la varianza

Contraste la H_0 ($\sigma^2 \geq 6$) frente a la alternativa H_1 ($\sigma^2 < 6$) en una distribución $N(\mu,\sigma)$, con un nivel de significación (α) igual a 0,1, mediante una muestra aleatoria simple de tamaño 7, cuyo resultado es:

7,1	5,3	4,7	8,0	9,9	3,4	3,6

Contrastes sobre la varianza

SOLUCIÓN:

1- Formulación de hipótesis

H_0: $\sigma^2 \geq 6$

H_1: $\sigma^2 < 6$

2- Elegir y calcular el estadístico de prueba

$$\chi^2_{\exp} = \frac{(n-1)S^2}{\sigma_0^2} = \frac{(7-1)5{,}8533}{6} = 5{,}8533$$

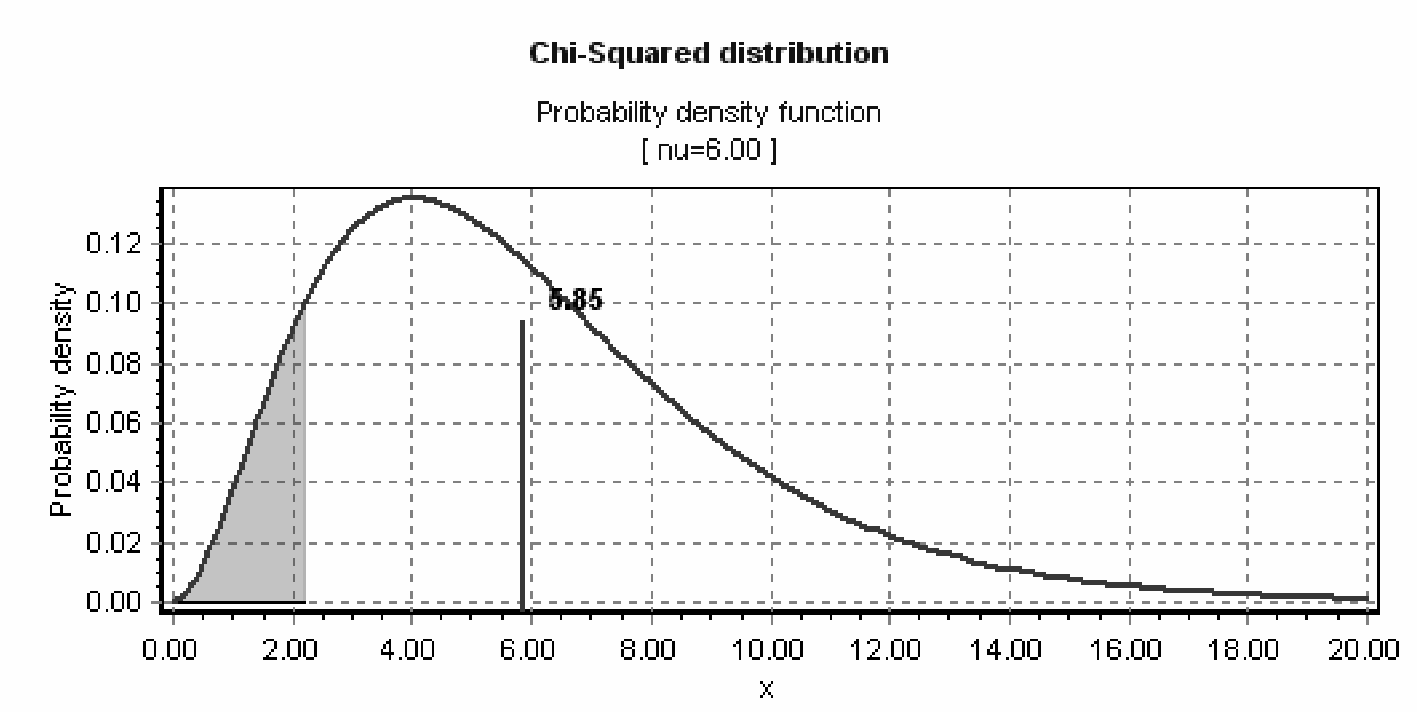
3- Determinación de la región crítica

$$P\left(\chi^2_{\exp} < \chi^2_{\alpha}\right) = 0{,}1 \Rightarrow \chi^2_{0{,}1} = 2{,}204 \Rightarrow \chi^2_{\exp} < 2{,}204.$$

4- Decisión: $X^2_{exp} > X^2_{0,1}$ ➔ Acepto la H_0

Dr. David Castilla Espino
CA UNED Huelva

Contrastes sobre la varianza



Contrastes sobre la proporción

En un proceso de fabricación se detecta que el 17% de las piezas son defectuosas. Mediante una puesta a punto se espera que esta cifra haya bajado. Contrastar esta hipótesis con un nivel de significación de 0,001, en una muestra aleatoria simple de tamaño 100 en la que se han obtenido 6 piezas defectuosas.

Contrastes sobre la proporción

SOLUCIÓN:

1- Formulación de hipótesis

H_0: $p \geq 0{,}17$

H_1: $p < 0{,}17$

2- Elegir y calcular el estadístico de prueba

$$Z_{exp} = \frac{p - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0{,}06 - 0{,}17}{\sqrt{\frac{0{,}17(1-0{,}17)}{100}}} = -2{,}9284$$

3- Determinación de la región crítica

$$P\left(Z_{exp} < -Z_{\alpha}\right) = 0{,}001 \Rightarrow Z_{0,001} = 3{,}08 \Rightarrow RC : Z_{exp} < -3{,}08.$$

4- Decisión: $Z_{exp} > Z_{0,001}$ ➔ Acepto la H_0

Sumario

- Se han introducido los contrastes paramétricos en el marco de la inferencia estadística como procedimiento para corroborar el cumplimiento de hipótesis estadísticas sobre los parámetros de distribuciones conocidas en contraposición con los contrastes no paramétricos.
- Se han introducido los contrastes de significación como aquellos que se basan a los efectos de aceptar o rechazar la H_0 en la discrepancia entre el valor del parámetro formulado bajo la hipótesis nula y el valor que le correspondería según la información que proporciona la muestra.
- Se han desarrollado los contrastes de razón de verosimilitud como aquellos contrastes de significación que miden la discrepancia por medio de la comparación de funciones de verosimilitud.
- Se han ejemplificado algunos problemas de resolución de contrastes de hipótesis paramétricos sobre la media, la varianza y la proporción.