

# Bioestadística

## Práctica de Ordenadores 7

José Aurelio Pina Romero

Ja.pina@ua.es

Bioestadística – Grado Enfermería

UA- Departamento de Enfermería

# ¿Qué vamos hacer?

- Repaso de Estadística Descriptiva
- Tablas de contingencia
- Intervalos para una media.
- Chi-cuadrado
- Conceptos de Inferencia (IC proporción)
- Ejercicio 4.1 (mano)

# Repaso de Estadística descriptiva

- **1 variable**

- **Cualitativa** ( $f_i$ ,  $f_{ri}$ ,  $F_i$ ,  $F_{ri}$ , %, gráficos de sectores, gráficos de barras)
- **Cuantitativa** (Medidas de tendencia central, medidas de dispersión, medidas de forma, gráficos de barras, histogramas)

# Repaso de estadística descriptiva

A partir de la base se datos LITIABIS.sav

- i. Calcule los estadísticos descriptivos de tendencia central, dispersión y forma para las variables EDAD, TALLA, PESO y TRIGLICERIDOS para cada sexo.
- ii. Calcule los percentiles 5,25,75 y 95 para cada sexo. Interpreta los resultados.
- iii. Calcule un diagrama de sectores para las variables Sexo, glucosa, cálculos y obesidad por sexos.

# Repaso de Estadística descriptiva

- **2 variables**
  - **Cuantitativa us Cuantitativa** (Diagramas de dispersión, coeficiente de correlación lineal de Pearson, Modelos regresión lineal)
  - **Cualitativa us cualitativa** (tablas de contingencia, Ji-cuadrado)

## Repaso de Estadística descriptiva

Construya el diagrama de dispersión para las variables Peso y Colesterol tanto para hombres como para mujeres. (Se sugiere representar ambos sexos sobre el mismo diagrama de dispersión distinguiendo mediante distintos colores).

Interprete los resultados.

**RECUERDO:**

**Gráficos [?] Cuadros de diálogo antiguos [?]**  
**Dispersión/Puntos**

## Repaso de Estadística descriptiva

Calcule la **correlación lineal** entre las variables peso y colesterol.

- i. Para hombres y mujeres conjuntamente
- ii. Para hombres y mujeres por separado
- iii. Compare los resultados
- iv. Construya un modelo de regresión lineal

## Repaso de Estadística descriptiva

Calcule la **correlación lineal** entre las variables peso y colesterol.

- i. Para hombres y mujeres conjuntamente
- ii. Para hombres y mujeres por separado
- iii. Compare los resultados
- iv. Construya un modelo de regresión lineal



# Tablas de contingencia

Relación entre dos variables cualitativas

Explora la distribución que posee una variable cualitativa entre diferentes muestras.

Calcule las tablas de contingencia para:

- i. Sexo us Obesidad
- ii. Sexo us Glucosa

# Independencia entre dos variables

## Chi- cuadrado

- Consiste en que la **distribución de una de las dos variables es similar sea cual sea el nivel que examinemos de la otra**
  - frecuencia de las filas ( o columnas) son aproximadamente proporcionales.
- Se contrasta que las variables sean independientes us 1 variable se distribuye diferente para diversos niveles de la otra

# Independencia entre dos variables

## Chi- cuadrado (LIMITACIONES)

- Frecuencias esperadas no muy pequeñas ( $n < 5$ ), no más del 20% de las celdas → Prueba exacta de Fisher.
- Para muestras grandes no es aconsejable utilizar esta prueba.
- No es aconsejable si la variable tienen muchos niveles de estudio.

# Independencia entre dos variables

## Estudio entre escolares

**Tabla de contingencia Sexo \* Prioridad**

Recuento

		Prioridad			Total
		Deportes	Notas	Popular	
Sexo	Niña	17	101	75	<b>193</b>
	Niño	51	95	38	<b>184</b>
Total		<b>68</b>	<b>196</b>	<b>113</b>	<b>377</b>

**Tabla de contingencia Sexo \* Prioridad**

% de Sexo

		Prioridad		
		Deportes	Notas	Popular
Sexo	Niña	9%	52%	39%
	Niño	28%	52%	21%

Contrasta si las diferencias observadas en los dos grupos son atribuibles al azar. No influye sexo en la prioridad de elección de actividades.

# Independencia entre dos variables

	Deportes	Notas	Popularidad	
Niña	17 (9%)	101 (52%)	75 (39%)	193
Niño	51 (28%)	95 (52%)	38 (21%)	184

	Deportes	Notas	Popularidad
Niña	(193/377)*68	(193/377)*196	(193/377)*113
Niño	(184/377)*68	(184/377)*196	(184/377)*113

**NO ASOCIACIÓN**

	Deportes	Notas	Popularidad
Niña	35 (18%)	100 (52%)	58 (30%)
Niño	33 (18%)	96 (52%)	55 (30%)

$$\chi^2 = \sum_{i=1}^I \sum_{j=1}^J \frac{(o_{ij} - e_{ij})^2}{e_{ij}} ; e_{ij} = \frac{o_{i+} \cdot o_{+j}}{n}$$

Pruebas de chi-cuadrado

	Valor	gl	Sig. asintótica (bilateral)
Chi-cuadrado de Pearson	29,100 <sup>a</sup>	2	,000

a. 0 casillas (.0%) tienen una frecuencia esperada inferior a 5. La frecuencia mínima esperada es 33,19.

Las preferencias no se distribuyen del mismo modo entre niñas y niños. Y no es atribuible al azar.

# Independencia entre dos variables

Calcule las tablas de contingencia para:

- i. Sexo us Obesidad
- ii. Sexo us Glucosa
  
- Calcule el valor de la chi-cuadrado

# Tablas de contingencia

The image shows the SPSS software interface. The 'Análisis' menu is open, and 'Tablas de contingencia...' is selected. A dialog box titled 'Tablas de contingencia' is displayed. The 'Filas' field contains 'sexo' and the 'Columnas' field contains 'Nivel de obesidad [obesidad]'. The 'Exact' options are highlighted with a red box, and a red arrow points from the menu to the dialog.

**Menú de Análisis:**

- Informes
- Estadísticos descriptivos
- Tablas
- Comparar medias
- Modelo lineal general
- Modelos lineales generalizados
- Modelos mixtos
- Correlaciones
- Regresión
- Loglineal
- Redes neuronales
- Clasificar
- Reducción de dimensionalidad
- Escala
- Pruebas no paramétricas
- Predicciones
- Supervivencia
- Respuesta múltiple
- Análisis de varianzas
- Imputación múltiple
- Muestras complejas
- Simulación
- Control de calidad
- Curva COR...

**Diálogo de Tablas de contingencia:**

- Filas:** sexo
- Columnas:** Nivel de obesidad [obesidad]
- Capa 1 de 1:** Anterior, Siguiente
- Mostrar variables de capa en capas de tabla
- Mostrar los gráficos de barras agrupadas
- Suprimir tablas
- Botones: Ayuda, Restablecer, Pegar, Cancelar, Aceptar

**Opciones Exactas (destacadas):**

- Exacta
- Estadísticos...
- Casillas...
- Formato...
- Bootstrap...

# Tablas de contingencia – Chi cuadrado

The image shows the SPSS 'Tablas de contingencia' dialog box with the following settings:

- Filas:** sexo
- Columnas:** Nivel de obesidad [obesidad]
- Estadísticos:**  Chi-cuadrado
- Casillas:**  Observado,  Esperado

Other options visible in the 'Estadísticos' dialog include:

- Correlaciones
- Coeficiente de contingencia
- Phi y V de Cramer
- Lambda
- Coeficiente de incertidumbre
- Gamma
- d de Somers
- Tau-b de Kendall
- Tau-c de Kendall
- Kappa
- Riesgo
- McNemar
- Estadísticos de Cochran y Mantel-Haenszel

The 'Estadísticos de Cochran y Mantel-Haenszel' section includes the text 'Probar que la razón de ventajas común equivale a:' followed by a text box containing the value '1'.



# Intervalo de confianza para una proporción(p)

$$IC_{1-a}(p) = \hat{p} \pm z_{1-a/2} \sqrt{\frac{\hat{p} \cdot (1 - \hat{p})}{n}}$$

$$z = \frac{\bar{X} - m}{\left(\frac{s}{\sqrt{n}}\right)} \sim N(0,1)$$

$$\begin{aligned} n \cdot \hat{p} &\geq 5 \\ n \cdot (1 - \hat{p}) &\geq 5 \end{aligned}$$

## Ejercicio 4.1

En un estudio realizado para determinar el estado de salud de una comunidad se entrevistó a **82 personas**, preguntándoles acerca de su actividad física habitual. De las 82 personas encuestadas, 36 de ellas declararon practicar algún deporte de forma regular.

**Construya el intervalo de confianza al nivel 0.95 para la proporción poblacional de práctica de algún deporte de forma regular.**

$$\begin{aligned} n \cdot \hat{p} &\geq 5 \\ n \cdot (1 - \hat{p}) &\geq 5 \end{aligned}$$

$$IC_{1-\alpha}(p) = \hat{p} \pm z_{1-\alpha/2} \sqrt{\frac{\hat{p} \cdot (1 - \hat{p})}{n}}$$

## Ejercicio 4.1

A partir de la información proporcionada por el intervalo de confianza anterior, ¿puede ser admisible que tal proporción sea de 0.60?