



PRÁCTICA 3: INFERENCIA ESTADÍSTICA

1. Abre el fichero *mercurio.sav*. En este ejercicio calcularemos intervalos de confianza y haremos contrastes de hipótesis para la media y la diferencia de medias.

a) Estima mediante un intervalo de confianza del 95% la concentración media de **mercurio**. Calcula también el intervalo de confianza al 90%. ¿Cuál tiene mayor longitud de los dos?

Procedimiento: *Analizar + Estadísticos descriptivos + Explorar*.
En *Dependientes* introducimos la variable mercurio.

b) Se dice que estos peces no son aptos para el consumo si tienen una concentración de mercurio mayor o igual a 1 parte por millón. ¿Podemos considerar que la concentración media de mercurio es igual a 1 parte por millón? En caso contrario, ¿crees que será mayor o menor?

Procedimiento: *Analizar + Comparar medias + Prueba T para una muestra*.
En *Contrastar variables* ponemos la variable mercurio y en *Valor de prueba* ponemos 1.

c) Calcula ahora un intervalo de confianza del 95% para la concentración media de mercurio en cada uno de los ríos.

Procedimiento: Igual que en el (a) metiendo ahora en *Factores* la variable río.

d) Estudia si es significativamente diferente la concentración media de mercurio del río Lumber a la del río Wacamaw.

Procedimiento: *Analizar + Comparar medias + Prueba T para muestras independientes*.
En *Contrastar variables* ponemos mercurio.
En *Variable de agrupación* ponemos río.

e) Estudia si es significativamente diferente la concentración media de mercurio de la estación 1 a la de la estación 9. Haz lo mismo para las estaciones 2 y 8.

Procedimiento: El mismo del apartado anterior poniendo en *Variable de agrupación* la variable estación.

2. En este ejercicio nos centraremos en el experimento consistente en el lanzamiento de una moneda. El SPSS nos ayudará a simular repetidas veces el lanzamiento de una moneda ficticia. Para ello consideramos la variable aleatoria X que vale 1 si sale cara y 0 cuando sale cruz. Esta variable es una variable de Bernoulli de parámetro p , donde p es la probabilidad de que X tome el valor 1, o, lo que es lo mismo, la probabilidad de obtener cara.

a) Supón que la moneda no está trucada. ¿Cuánto vale en este caso el parámetro p ? Teniendo en cuenta este valor de p , genera 100 valores de X .

Procedimiento: Define la variable X en *Vista de variables*.
Posiciónate en la celda número 100. Introduce en esa celda cualquier valor.
Transformar + Calcular.
En *Variable de destino* pon la variable X y en *Expresión numérica* introduce $RV.BERNOULLI(p)$.

b) Calcula la media para los datos que has obtenido.

Procedimiento: *Analizar + Estadísticos descriptivos + Descriptivos.*

c) Supón que sólo conocemos los datos que tienes en el fichero y sabemos que son 100 valores correspondientes al lanzamiento de una moneda, pero no sabemos si la moneda está trucada o no. Calcula un intervalo de confianza del 95% para la probabilidad de cara, es decir, para p .

Procedimiento: *Analizar + Comparar medias + Prueba T para una muestra.*

En *Opciones* aparece por defecto el nivel de confianza 0.95.

Lo dejamos entonces tal y como está.

¿Contiene el intervalo que hemos calculado al valor verdadero de p ?

d) Construye ahora un intervalo de confianza al 99%. Compáralo con el del 95%. ¿Cuál crees que tiene mayor probabilidad de contener el valor verdadero de p ?

e) Supón que queremos contrastar si la moneda es correcta. Para ello vamos a suponer entonces que la hipótesis nula consiste en suponer que es correcta, es decir, $H_0: p=1/2$. De esta forma, tenemos el siguiente contraste de hipótesis: $H_0: p=1/2$ y $H_1: p \neq 1/2$. A la vista del contraste de hipótesis, ¿Rechazaremos la hipótesis nula?

Procedimiento: *Analizar + Comparar medias + Prueba T para una muestra.*

Marca en *Valor de prueba* el valor 0.5.

3. En este ejercicio vamos a calcular intervalos de confianza y hacer contrastes de hipótesis para **muestras relacionadas**. Para ello consideramos el siguiente ejercicio:

Se tienen los pesos de 10 parejas antes y después de 6 meses de matrimonio:

Hombres antes	72	69	81	71	88	78	68	76	86	95
Hombres después	77	68.5	85	74.5	90.5	76	71	75	87.5	101
Mujeres antes	52	56	61	49	57	63	66	59	67	51
Mujeres después	54	55	58	50	55	61	64	56	70	50

a) Introduce los datos que tienes en esta tabla en un fichero al que le llamarás *matrimonio.sav* y cuyas variables van a ser **pesoha**, **pesohd**, **pesoma** y **pesomd**.

b) ¿Tiene alguna influencia el matrimonio sobre el peso de los hombres?

Procedimiento: *Analizar + Comparar medias + Prueba T para muestras relacionadas.*

En *Variables* introducimos las variables *pesoh1* y *pesoh2*.

c) ¿Tiene alguna influencia el matrimonio sobre el peso de las mujeres?

4. Abre el fichero de datos *alumnos.sav* en la carpeta *Estadística*. En este ejercicio vamos a hacer un contraste de asociación para una tabla de contingencia.

a) En este apartado vamos a hacer el contraste de independencia para las variables **sexo** y **mate**.

Procedimiento: *Analizar + Estadísticos descriptivos + Tablas de contingencia.*

Poner en *Filas* la variable *mate* y en *Columnas* la variable *sexo*.

En *Estadísticos* marcamos *Chi-cuadrado*.

En *Casillas* marcamos *Frecuencias Observadas* y *Frecuencias Esperadas* y *Residuos no tipificados*.

A la vista de los resultados, ¿existen evidencias significativas de dependencia entre el sexo y estudiar matemáticas en el bachillerato?

b) Hacer lo mismo que en el apartado anterior para las variables **sexo** y **ojos**.

5. Abre el fichero de datos *Mundo95.sav* en la carpeta del SPSS. En este ejercicio vamos a trabajar con las variables **región** (región económica), **espvidaf** (esperanza de vida femenina) y **espvidam** (esperanza de vida masculina).

a) Crea una nueva variable llamada **evfrecod** que tome el valor 1 si la esperanza de vida de las mujeres es menor que 70, el valor 2 si la esperanza de vida de las mujeres oscila entre 70 y 77 años y el valor 3 si la esperanza de vida es mayor que 77.

b) ¿La esperanza de vida de una mujer es independiente de la región económica a la que pertenece su país? Utiliza para verlo la variable evfrecod.

Procedimiento: *Analizar + Estadísticos descriptivos + Tablas de contingencia.*

En *filas* selecciona la variable *región* y en *columnas* la variable evfrecod.

En *Estadísticos* selecciona *Chi-cuadrado*.

En *Casillas* selecciona *Frecuencias esperadas* y *Residuos no tipificados*.

Como algunas frecuencias esperadas son inferiores a 5 recodificamos la variable *región* en la variable **regrecod**. Esta nueva variable se obtendrá agrupando las clases 2 y 3 de la variable *región*.

c) Haz un estudio similar para el caso de los hombres.