

PRÁCTICA 5: INFERENCIA ESTADÍSTICA

T E O R I A

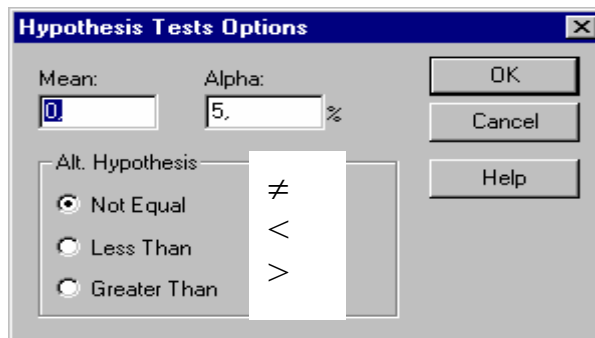
En esta práctica trabajaremos los intervalos de confianza y contrastes de hipótesis. Veremos únicamente las opciones más básicas que podemos calcular con el Statgraphics. Empezaremos por el estudio de una muestra, después seguiremos con el estudio para dos y por último para más de dos. A lo largo de la práctica, únicamente trataremos variables continuas medidas a nivel de intervalo (contrastos en tablas de contingencia, pueden encontrarse en **Describe > Categorical Data**).

Estudio de una sola muestra:

Describe > Numeric Data > One variable analysis: en **Data** seleccionamos el nombre de la variable sobre la que deseamos realizar el estudio.

Con **Tabular options** (botón amarillo), podemos estudiar:

- **Confidence Intervals** (intervalos de confianza) para la media (mean) y la desviación típica (standard deviation). Con **Pane options** (botón derecho del ratón), podemos cambiar el grado de confianza.
- **Hypothesis Test** (contrastos de hipótesis). Nos devuelve varios contrastes. A nosotros nos interesa, el contraste t. La hipótesis nula (H_0) es $\mu = 0$, y la alternativa (H_1) es $\mu \neq 0$. Podemos cambiar, tanto el valor de H_0 , como la hipótesis alternativa y el nivel de significación, mediante **Pane options** (botón derecho del ratón). Devuelve el valor del estadístico y el p-valor (o



nivel crítico = α), que nos indicará si rechazar o no la hipótesis nula.

Con **Graphical options** (botón azul) podemos obtener los gráficos, como vimos en la práctica 1.

Estudio de dos muestras:

Compare > Two Samples

1) Si las muestras son independientes: > **Two Sample Comparison**

En Sample 1 y Sample 2 colocamos las dos variables a estudiar.

Con **Tabular options** (botón amarillo) podemos seleccionar:

- **Summary statistics:** devuelve medidas estadísticas para ambas variables.
- **Comparison of means:** devuelve los intervalos de confianza para la media de ambas variables y su diferencia (suponiendo varianzas iguales y distintas). Además, se contrasta la igualdad de medias frente a distintas alternativas (suponiendo varianzas iguales y distintas). Con **Pane options** (botón derecho del ratón) podemos cambiar el grado de confianza.
- **Comparison of standard deviation:** devuelve los intervalos de confianza de las desviaciones típicas de cada muestra, además del cociente entre ambas varianzas. Conjuntamente, calcula diversos contrastes de igualdad de varianzas (tanto el valor de F como el p-valor = α). Nuevamente, con **Pane options** (botón derecho del ratón) podemos cambiar el grado de confianza.

Con **Graphical options** (botón azul) podemos obtener distintos gráficos, entre ellos el diagrama de cajas (**Box-and-Whisker Plot**).

2) Si las muestras son dependientes o apareadas: > **Paired Sample Comparison**

En Sample 1 y Sample 2 colocamos las dos variables a estudiar.

Con **Tabular options** (botón amarillo) podemos seleccionar:

- **Summary statistics:** devuelve medidas estadísticas para la variable Diferencia (Sample1 – Sample 2).
- **Confidence Intervals** (intervalos de confianza) para la media (mean) y la desviación típica (standard deviation) de la Diferencia. Con **Pane options** (botón derecho del ratón), podemos cambiar el grado de confianza.
- **Hypothesis Test** (contrastos de hipótesis). Nos devuelve varios contrastes. A nosotros nos interesa, el contraste t. La hipótesis nula (H_0) es $\mu_1 - \mu_2 = 0$, y la alternativa (H_1) es $\mu_1 - \mu_2 \neq 0$. Podemos cambiar, tanto el valor de H_0 , como la hipótesis alternativa y el nivel de significación, mediante **Pane options** (botón derecho del ratón). Devuelve el valor del estadístico, así como el p-valor = α , que nos indicará si rechazar o no la hipótesis nula.

Estudio de varias muestras:

Compare > Analysis of Variance > One Way ANOVA

En Dependent Variable colocaremos la variable dependiente (sobre la que queremos contrastar la igualdad de medias) y en Factor la variable que indica los grupos o muestras distintas que tenemos.

Con **Tabular options** (botón amarillo) podemos seleccionar:

- **ANOVA Table:** devuelve la tabla ANOVA.
- **Kruskal-Wallis test:** es un contraste no paramétrico, alternativo al ANOVA de una vía, que usaremos cuando las hipótesis de normalidad e igualdad de varianzas no se cumplan. El p-valor nos indicará si rechazar o no la hipótesis nula.

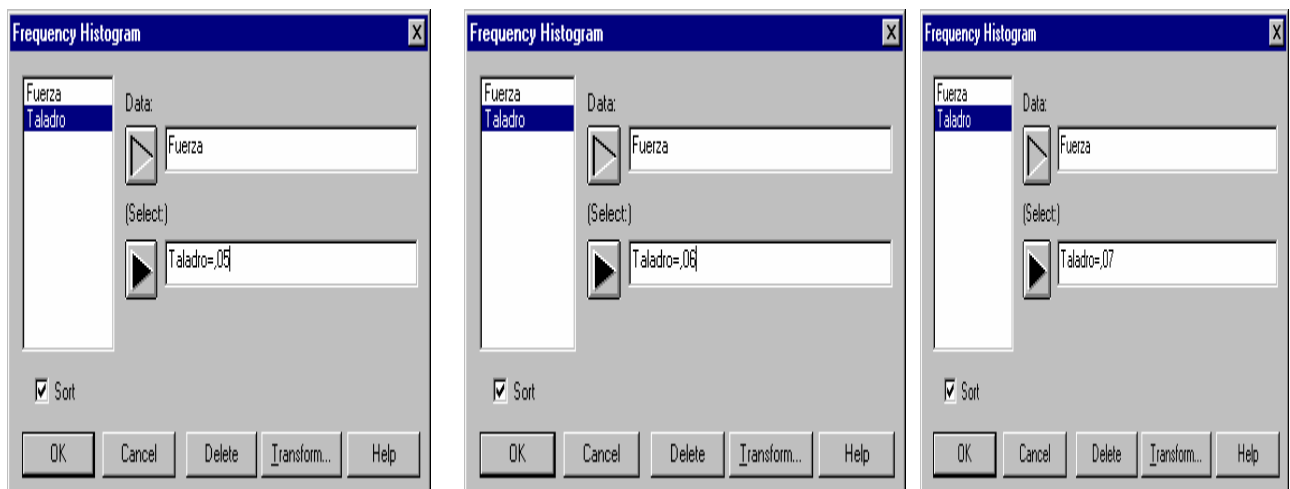
Con **Graphical options** (botón azul) podemos obtener distintos gráficos, entre ellos el diagrama de cajas (**Box-and-Whisker Plot**).

Histogramas de varias muestras:

Puedes realizar distintos histogramas de unos datos, según el valor que adopte el factor. Por ejemplo,

Si tu variable dependiente fuera *Fuerza* y tu factor fuese *Taladro*, (que tiene 3 niveles), podrías realizar 3 histogramas de la siguiente forma:

>Plot >Exploratory Plots > Frequency Histograms



Práctica5. Nombre:**Grupo:**

1. De la variable **precio** del archivo “**impresoras**”, calcula el intervalo de confianza al nivel de significación $\alpha = 0'01$ para la μ . * Media: $\mu \in [\quad , \quad]$

* Desviación típica: $\sigma \in [\quad , \quad]$ Confianza = $1 - \alpha = \quad$.

2. En el archivo “**encuesta23**”, construir las variables: $\text{edadpsi} = \text{edad con titulacion}=1$ $\text{edaditis} = \text{edad con titulacion}=0$

a) Calcula el intervalo de confianza al 95 % para cada una de las medias.

$\alpha = \quad$; $\mu \in [\quad , \quad]$; $\mu \in [\quad , \quad]$

b) Calcula el intervalo de confianza para las desviaciones típicas con $\alpha = 0'05$

$\sigma \in [\quad , \quad]$; $\sigma \in [\quad , \quad]$

c) Calcula el intervalo para el cociente (ratio) de varianzas: $[\quad , \quad]$

d) Dibuja un diagrama de caja con las dos variables. (Detrás)

e) Calcula el valor de F para comparar la varianza de las dos variables ($\alpha = 0'05$):

$F = \quad$; $\alpha = \quad$ $\alpha = 0'05$ ¿son diferentes las varianzas? \quad .

f) Calcula el intervalo de confianza al 95 % para la diferencia de las medias, suponiendo varianzas iguales: $[\quad , \quad]$

g) ¿Podemos considerar las medias diferentes? \quad Porque \quad ;

Contraste: $t = \quad$, $\alpha = \quad$ $\alpha = 0'05$

3. Del archivo “**encuesta23**”, considera: **notamas = X** **notamen = Y**

Son variables \quad .

a) Calcula la media y varianza para cada una de las dos variables y de la variable diferencia: $x = \quad$, $s^2 = \quad$; $y = \quad$, $s^2 = \quad$; $D = \quad$, $s^2 = \quad$.

b) Calcula el intervalo de confianza para la media de la variable diferencia al 99 %.

$\mu_D \in [\quad , \quad]$

c) ¿Podemos suponer una diferencia de 4 puntos entre las notas? $H_0: \mu_D = 4$

Contraste: $t = \quad$, $\alpha = \quad$ $\alpha = 0'05 \Rightarrow \quad$.