Seguimos con el ejercicio y pasamos a una fase muy interesante.

Vamos a comparar medias utilizando los intervalos de confianza y las tablas ANOVA que nos van a venir muy bien para realizar el ejercicio.

Vamos a analizar la primera de las situaciones. La comparación de medias con SPSS y las tablas ANOVA.

Vamos a comparar la fertilidad en función de los si el país es o no del OECD.

Para hacerlo sigue el ejemplo de la zona tech Zen.

Y utiliza la guía del contraste de hipótesis:

<https://s3.eu-west-3.amazonaws.com/analiza-tus-datos/5-ANALIZA+TUS+DATOS/1-THE+BASICS/%5BAD%5D+La+plantilla+del+Contraste+de+Hipotesis.pdf>

Aviso: el ejercicio lo vamos a hacer en Excel y SPSS. Pero puedes utilizar RStudio, RCommander o cualquier otro software. No hay ningún problema.

El ejercicio trata de entender conceptos estadísticos de forma práctica y entender un método de trabajo ☺

Puedes conseguir una copia de prueba SPSS durante un mes en esta página:

<https://www.ibm.com/analytics/es/es/technology/spss/spss-trials.html>

O utilizar PSPP que es el software libre y gratuito casi igual que SPSS:

<https://www.gnu.org/software/pspp/>

Excel supongo que lo tendrás instalado ☺

¡A por ello!

# Analizando la comparación de medias con SPSS

## Variable Fertilidad

Mira el vídeo de la zona tech Zen de SPSS de la página del ejercicio y mira después los resultados que te pongo aquí.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **ANOVA** | | | | | |
| fertility | | | | | |
|  | Suma de cuadrados | gl | Media cuadrática | F | Sig. |
| Entre grupos | 167,100 | 2 | 83,550 | 85,957 | ,000 |
| Dentro de grupos | 184,680 | 190 | ,972 |  |  |
| Total | 351,780 | 192 |  |  |  |

**Pruebas post hoc**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Comparaciones múltiples** | | | | | | |
| Variable dependiente: fertility | | | | | | |
| HSD Tukey | | | | | | |
| (I) group | (J) group | Diferencia de medias (I-J) | Error estándar | Sig. | Intervalo de confianza al 95% | |
| Límite inferior | Límite superior |
| oecd | africa | -2,507473\* | ,223712 | ,000 | -3,03593 | -1,97901 |
| other | -,595759\* | ,200477 | ,009 | -1,06933 | -,12219 |
| africa | oecd | 2,507473\* | ,223712 | ,000 | 1,97901 | 3,03593 |
| other | 1,911714\* | ,165918 | ,000 | 1,51978 | 2,30365 |
| other | oecd | ,595759\* | ,200477 | ,009 | ,12219 | 1,06933 |
| africa | -1,911714\* | ,165918 | ,000 | -2,30365 | -1,51978 |
| \*. La diferencia de medias es significativa en el nivel 0.05. | | | | | | |

Conclusión: tenemos diferencias significativas entre todos los grupos ya que el p-valor es significativo y el 0 no aparece dentro de los intervalos de confianza al 95%.

Replica el ejercicio con las variables infantMortality y lifeExpF:

## Variable “infantMortality”

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| infantMortality por grupo  H1: Existen diferencias significativas en la mortalidad infantil por grupos  A priori se observan medias muy distintas por grupo.      Al ver el box plot observamos que los datos parecen tener una distribución normal. Igualmente se observa claramente que hay diferencias en la mortalidad infantil dependiendo del grupo  El diagrama de Error o de medias nos muestra o recalca las diferencias. Se nota que no hay cero entre las medias por lo que puedo concluir que hay diferencias entre grupos.  ANOVA    ¿Qué conclusiones puedes obtener?  Al revisar el ANOVA vemos que en efecto tanto el anova del grupo completo dio < 0.05 y luego el ANOVA por pares en todos los casos da un p value < 0.05 por lo que la H1 es cierta y se cumple.   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **ANOVA** | | | | | | | infantMortality | | | | | | |  | Suma de cuadrados | gl | Media cuadrática | F | Sig. | | Entre grupos | 91934,999 | 2 | 45967,500 | 120,862 | ,000 | | Dentro de grupos | 72262,802 | 190 | 380,331 |  |  | | Total | 164197,802 | 192 |  |  |  |   Pruebas post hoc   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | **Comparaciones múltiples** | | | | | | | | Variable dependiente: infantMortality | | | | | | | | HSD Tukey | | | | | | | | (I) group | (J) group | Diferencia de medias (I-J) | Error estándar | Sig. | Intervalo de confianza al 95% | | | Límite inferior | Límite superior | | oecd | africa | -60,430672\* | 4,425243 | ,000 | -70,88410 | -49,97724 | | other | -16,785930\* | 3,965637 | ,000 | -26,15367 | -7,41819 | | africa | oecd | 60,430672\* | 4,425243 | ,000 | 49,97724 | 70,88410 | | other | 43,644743\* | 3,282011 | ,000 | 35,89188 | 51,39760 | | other | oecd | 16,785930\* | 3,965637 | ,000 | 7,41819 | 26,15367 | | africa | -43,644743\* | 3,282011 | ,000 | -51,39760 | -35,89188 | | \*. La diferencia de medias es significativa en el nivel 0.05. | | | | | | | |

¿Qué conclusiones puedes obtener?

|  |
| --- |
|  |

Replica el ejercicio con las variables infantMortality y lifeExpF:

## Variable “lifeExpF”

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| H1: Existen difencias significativas en la expectativa de vida entre los grupos  A priori la media muestra diferencias marcadas entre Africa y los demás ocde y otros. Entre estos dos últimos hay diferencias pero son más cercanas las medias      El boxplot nos dice muy claramente que aparentemente hay diferencias entre ocde y otros. Africa tiene una expectativa de vida mucho menor que todos  El diagrama de error sigue mostrando las diferencias. Ya se ve que los rangos de las medias de cada grupo no incluyen al cero por lo que se va concluyendo que la H1 se cumple. Y que en los tres grupos hay diferencias.  ANOVA        Qué conclusiones puedes obtener?  Al revisar el ANOVA vemos que en efecto tanto el anova del grupo completo dio < 0.05 y luego el ANOVA por pares en todos los casos da un p value < 0.05 por lo que la H1 es cierta y se cumple.   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **ANOVA** | | | | | | | lifeExpF | | | | | | |  | Suma de cuadrados | gl | Media cuadrática | F | Sig. | | Entre grupos | 12694,486 | 2 | 6347,243 | 167,221 | ,000 | | Dentro de grupos | 7211,880 | 190 | 37,957 |  |  | | Total | 19906,366 | 192 |  |  |  |   Pruebas post hoc   |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | | **Comparaciones múltiples** | | | | | | | | Variable dependiente: lifeExpF | | | | | | | | HSD Tukey | | | | | | | | (I) group | (J) group | Diferencia de medias (I-J) | Error estándar | Sig. | Intervalo de confianza al 95% | | | Límite inferior | Límite superior | | oecd | africa | 23,024721\* | 1,397990 | ,000 | 19,72235 | 26,32709 | | other | 7,304856\* | 1,252794 | ,000 | 4,34547 | 10,26424 | | africa | oecd | -23,024721\* | 1,397990 | ,000 | -26,32709 | -19,72235 | | other | -15,719864\* | 1,036829 | ,000 | -18,16909 | -13,27064 | | other | oecd | -7,304856\* | 1,252794 | ,000 | -10,26424 | -4,34547 | | africa | 15,719864\* | 1,036829 | ,000 | 13,27064 | 18,16909 | | \*. La diferencia de medias es significativa en el nivel 0.05. | | | | | | | |

¿Qué conclusiones puedes obtener?

|  |
| --- |
|  |

¡Has hecho un gran trabajo de análisis!

Estas captando la esencia de la práctica estadística.

Es muy común realizar este tipo de análisis en la práctica en investigación.

¡Sigue así!

Ciao ciao!