

Regresión y ANOVA: Analgésicos Infantiles *

García Prado, Sergio
sergio@garciparedes.me

15 de noviembre de 2017

1. Contexto y Conjunto de Datos

En este trabajo se va a realizar un estudio acerca de la diferencia de medias sobre el conjunto de datos **painkillers**, el cual se refiere a un experimento sobre *Analgésicos Infantiles*. Para ello, se utilizará la técnica de *Análisis de la Varianza (ANOVA)*. Una contextualización más detallada del experimento se describe a partir del siguiente enunciado:

“El departamento de pediatría de un hospital desea analizar la eficacia de cuatro analgésicos infantiles ante las cefaleas. Para ello, realiza un experimento en el que se seleccionan aleatoriamente cinco grupos de cuatro pacientes, de manera que en cada grupo se da un cefalea distinto. A continuación se suministra, también de forma aleatoria, cada analgésico a uno de los pacientes de cada grupo, y se observa el tiempo de remisión de la cefalea, en minutos. Se registran los datos siguientes, en cada uno de los cinco grupos (**tiempo de remisión**, **analgésico** y **cefalea**).”

2. Cuestiones

En esta sección se incluyen una serie de cuestiones que serán resueltas mediante el estudio del conjunto de datos a partir de la técnica *ANOVA*.

2.1. Estudia el tipo de diseño adecuado para esta situación, e identifica las variables, factores y parámetros

Tras analizar el contexto del experimento, se sabe que la variable respuesta Y que se utilizará para el *análisis de la varianza* es **tiempo de remisión**, dado que es la que se utiliza para cuantificar la calidad del tratamiento. En cuanto a las variables a partir de las cuales se pretende explicar el **tiempo de remisión**, estas son el **analgésico** y el **cefalea**.

Sin embargo, estas no han sido seleccionadas de la misma manera, dado que tal y como se indica en el enunciado, se han fijado 4 muestras de *a priori* 5 **tipos de cefalea**, sobre los cuales aplicar los 4 tipos de **analgésico**. Por tanto, podemos interpretar dicha situación diciendo que el factor **tipos de cefalea** representa un **bloque** (que denotaremos por B_β siendo $\beta \in \{1, 2, 3, 4, 5\}$ el identificador de cada uno de los niveles del bloque), y asumiento que el factor **analgésico** representa un **tratamiento** (que denotaremos por T_α siendo $\alpha \in \{A, B, C, D\}$ el identificador de cada uno de los niveles del tratamiento).

$$Y_{ij} = \mu + T_\alpha + B_\beta + \epsilon_{ij} \tag{1}$$

*URL: <https://github.com/garciparedes/anova-painkillers>

Por dichas razones, utilizaremos el modelo de *1 factor + 1 bloque*, el cual se muestra en la ecuación (1). A este modelo se le ha añadido además la componente ϵ_{ij} , que representa el error aleatorio y sigue una distribución $N(0, \sigma^2)$, asumiendo que σ^2 es la misma para todas las observaciones.

[TODO]

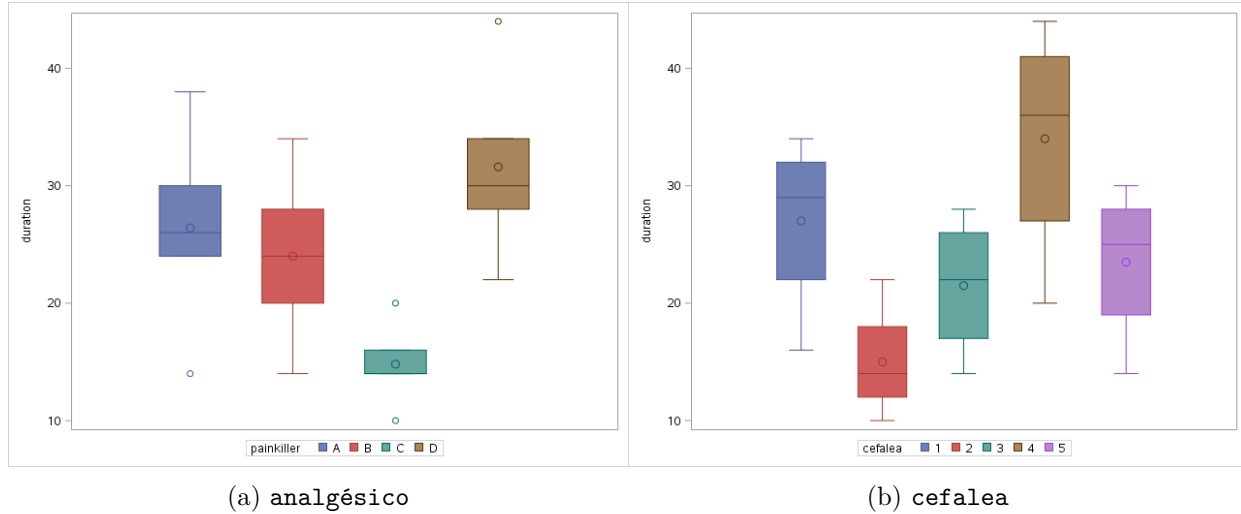


Figura 1: Diagramas de cajas

2.2. ¿Es adecuado usar las cefaleas como bloques?. ¿Existen diferencias significativas entre los tiempos de remisión de las cefaleas para los distintos analgésicos?

[TODO]

2.3. Estudia gráficamente la existencia de interacción entre analgésico y cefalea.

[TODO]

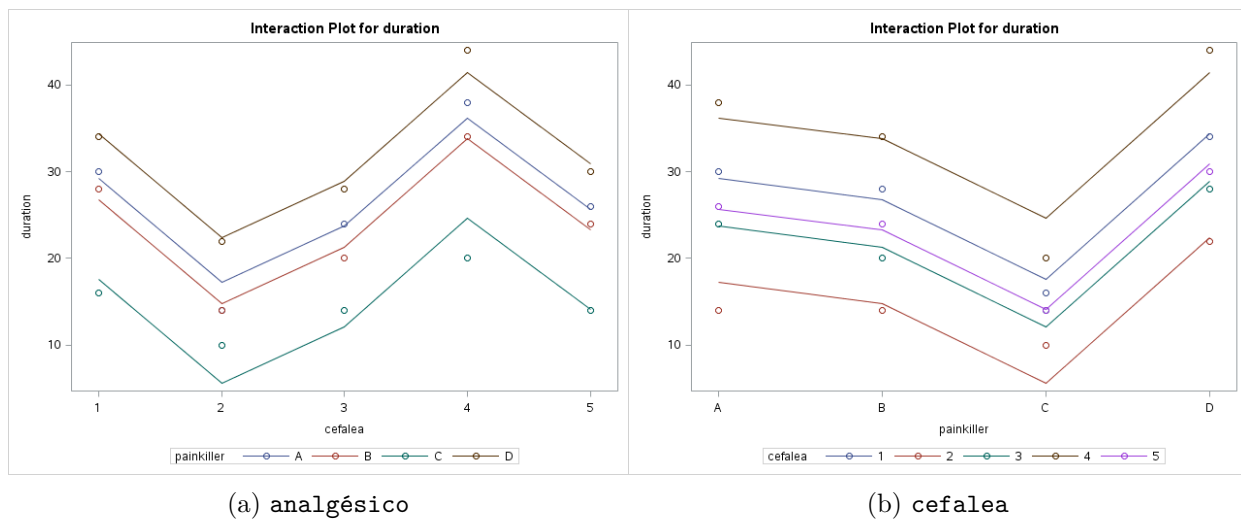


Figura 2: Diagramas de Interacción

2.4. Determina mediante comparaciones múltiples cuál de los analgésicos es más eficaz.

[TODO]

2.5. Suponiendo que el analgésico A es un placebo, realiza el test de Dunnett

[TODO]

2.6. Haz un análisis gráfico de los residuos.

[TODO]

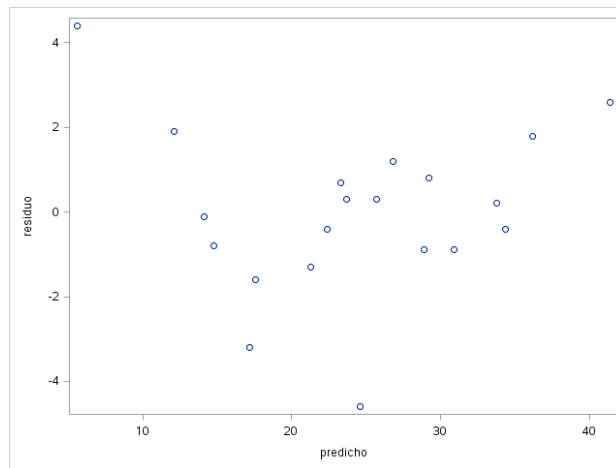


Figura 3

2.7. Si los analgésicos se hubieran elegido al azar entre todos los existentes, plantea el modelo adecuado y estima las componentes de la varianza.

[TODO]

2.8. Plantea el modelo como diseño unifactorial completamente aleatorizado y compara los resultados.

[TODO]

3. Código Fuente

[TODO]

```
data painkillers;  
  input duration painkiller$ cefalea;  
  datalines;  
30 A 1  
28 B 1  
16 C 1  
34 D 1  
14 A 2  
14 B 2  
10 C 2  
22 D 2  
24 A 3  
20 B 3  
14 C 3  
28 D 3  
38 A 4  
34 B 4  
20 C 4  
44 D 4  
26 A 5  
24 B 5  
14 C 5  
30 D 5  
;  
run;
```

Figura 4: *Código SAS*: Lectura del conjunto de datos.

Referencias

- [1] BARBA ESCRIBÁ, L. Regresión y ANOVA, 2017/18. Facultad de Ciencias: Departamento de Estadística.
- [2] SAS® SOFTWARE INSTITUTE. Sas. <https://www.sas.com/>.