

2. Un gobierno provincial realiza una evaluación de opción múltiple a estudiantes de quinto año en 5 localidades, calificándolos entre 0 y 100. Desea saber si hay diferencias significativas entre los desempeños de las distintas localidades, con el objeto de reforzar el presupuesto educativo en aquellas con peor desempeño. A pesar de su buena voluntad a futuro, no tiene presupuesto asignado para pagar las correcciones de estos exámenes, así que el ministro elige apenas 20 exámenes de cada localidad al azar entre las respectivas pilas y se pone a corregir.

a) (4p) Describir las variables aleatorias involucradas, los parámetros y los supuestos para realizar un test F de Análisis de la varianza. ¿Cuáles son las hipótesis del test?

.....
.....
.....
.....
.....
.....
.....
.....
.....
.....

A continuación se detallan las salidas obtenidas en R:

```
> anovanotas<-aov(notas~localidad)
> shapiro.test(anovanotas$res)
```

Shapiro-Wilk normality test

```
data: anovanotas$res
W = 0.9837, p-value = 0.2559
```

```
> levene.test(notas,localidad)
```

modified robust Brown-Forsythe Levene-type test based on the absolute deviations from the median

```
data: notas
Test Statistic = 0.4128, p-value = 0.799
```

```
> summary(anovanotas)
          Df Sum Sq Mean Sq F value    Pr(>F)
localidad  4  4796  1199.1    5.298 0.000678 ***
Residuals 95 21501   226.3
```

Signif. codes: 0 *** 0.001 ** 0.01 * 0.05 . 0.1 1

```
> TukeyHSD(anovanotas)
  Tukey multiple comparisons of means
    95% family-wise confidence level
```

```
Fit: aov(formula = notas ~ localidad)
```

```
$localidad
      diff      lwr      upr      p adj
2-1 -6.75 -19.9796199  6.47962 0.6171542
3-1 10.00  -3.2296199 23.22962 0.2278144
4-1 12.50  -0.7296199 25.72962 0.0734032
5-1  4.10  -9.1296199 17.32962 0.9100341
3-2 16.75   3.5203801 29.97962 0.0058434
4-2 19.25   6.0203801 32.47962 0.0009858
5-2 10.85  -2.3796199 24.07962 0.1602573
4-3  2.50 -10.7296199 15.72962 0.9845509
5-3 -5.90 -19.1296199  7.32962 0.7279397
5-4 -8.40 -21.6296199  4.82962 0.3995901
```

A continuación se calculan las medias muestrales de notas por localidad

```
> tapply(notas,localidad,mean)
      1      2      3      4      5
63.95 57.20 73.95 76.45 68.05
```

b) (4p) ¿Se cumplen los supuestos del modelo para aplicar el test Anova? Fundamente brevemente:

.....
.....
.....

c) (4p) ¿Qué conclusión puede sacar el ministro?

.....
.....
.....

d) (4p) ¿Existen pares de localidades para las cuales se puede decir que la media de sus notas es significativamente diferente con nivel simultáneo 5%? Fundamente brevemente

.....
.....

.....
e) (4p) Si se decide subir el presupuesto en el mayor grupo de localidades de peores resultados en el que no se encuentren diferencias significativas a nivel simultáneo 5%. ¿En qué localidades deberían subir el presupuesto? Fundamente brevemente

.....
.....
.....