Didáctica de la Estadística

Miguel R. Wilhelmi*

Resumen

En este taller proponemos medios para la enseñanza y el aprendizaje de la combinatoria, la probabilidad y la estadística y proponemos una vía para la valoración de propuestas de enseñanza.

1. Conexiones Matemáticas y Atomización de la Enseñanza

Las matemáticas son esencialmente relacionales. Un objeto matemático aislado pierde su sentido. El NCTM (2000) propone por ello las *conexiones* como uno de lo estándares de proceso fundamentales. En el resumen ejecutivo para los *Principios y Estándares para la Educación Matemática* (www.nctm.org/ standards/) se explicita:

> "*Conexiones*. Las matemáticas no es un conjunto separado de ejes temáticos o estándares, aún cuando sean presentados a menudo de esta manera. Por el contrario, las matemáticas son un campo de estudio integrado. Cuando los estudiantes relacionan las ideas matemáticas, su comprensión y entendimiento acerca de ellas se hacen profundos y son más permanentes, y pueden percibir las matemáticas como un todo coherente. Ellos pueden visualizar las conexiones matemáticas en gran interacción con otros tópicos matemáticos, los contextos que relacionan las matemáticas con otros temas, y sus propios intereses y experiencias. En una enseñanza que enfatiza en la interrelación de las ideas matemáticas, los estudiantes

* Universidad Pública de Navarra - España

> no sólo aprenden matemáticas sino también acerca de la utilidad de las matemáticas.”

En la institución educativa hay, sin embargo, una tendencia a limitar las conexiones matemáticas y a presentar las nociones y procedimientos de manera aislada bajo la ilusión de que este mecanismo didáctico facilita su aprendizaje.

> “Al intentar proteger al alumno de toda desconcentración y evitarle el encuentro con los sucesivos obstáculos epistemológicos, se fracciona el proceso de enseñanza hasta hacerlo desaparecer como proceso. Se pretende de esta manera paliar las dificultades que comporta toda actividad matemática sostenida y compleja. La enseñanza se convierte en un conjunto atomizado de actividades matemáticas aisladas, de ‘anécdotas’ matemáticas encadenadas arbitrariamente e independientes entre sí que no permiten al alumno llegar a dominar ninguna técnica y lo convierten, de hecho, en un ‘incompetente’. Este tipo de enseñanza tiende a lo que podríamos denominar una ‘enseñanza instantánea’: [… los] objetos matemáticos […] se ‘enseñan’ y ‘aprenden’ casi al mismo tiempo […] Desaparecen los objetivos a largo plazo a favor de los objetivos relativos al funcionamiento diario de la clase […] Esta situación tiene consecuencias paradójicas porque, intentando proteger a los alumnos de toda desconcentración, los lleva a un estado de desconcentración permanente y los sitúa definitivamente fuera del contrato didáctico.” (Chevallard, Bosch y Gascón, 1997, 285)

La enseñanza y aprendizaje de las matemáticas deben ser fieles a la naturaleza esencialmente relacional de las matemáticas. La labor del profesor en la determinación de situaciones lo suficientemente complejas que permitan la actividad matemática sostenida en procesos de estudio matemático y que a su vez determinen condiciones de corresponsabilidad matemática (profesor-estudiantes) en el funcionamiento y control de los sistemas didácticos.

En la sección 2 mostramos una situación para introducir la combinatoria, la probabilidad y la estadística de manera relacionada. En la sección 3, se analiza una situación real analizada mediante un “contraste de hipótesis” y se propone elaborar una transposición didáctica de la noción para el primer ciclo de secundaria. Estas dos secciones terminan con unas cuestiones que guiarán el seminario-taller, en el cual se irán fundamentando la necesidad de nociones y métodos de didáctica de las matemáticas para la elaboración de situaciones de enseñanza. Este proceso llevará a la proposición de un instrumento para el análisis de situaciones (sección 4), que se desarrollará en la última sesión del taller.

2. La Carrera

Wilhelmi (2004) propone la situación *La carrera* para introducir de forma intuitiva y relacionada los conceptos clave de recuento sistemático, frecuencia, probabilidad y estadística (recogida, organización, visualización y análisis de datos). Se trata de un juego para dos personas. Se necesita un tablero como el que se muestra en la figura 1 (con once filas numeradas del 2 al 12 y 11 columnas, la última de las cuales está marcada con la palabra meta), 10 fichas de dos colores distintos (5 de cada color) y dos dados (numerados del 1 al 6).

<table>
<tr><td>2</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td rowspan="11">M
E
T
A</td></tr>
<tr><td>3</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>4</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>5</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>6</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>7</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>8</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>9</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>10</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>11</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>12</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
</table>

Figura 1. Tablero de la situación *La carrera*

Reglas de juego

1. Alternativamente, cada uno de los contrincantes, escoge un número comprendido entre 2 y 12 (posibles resultados en la suma de un par de dados), colocando una ficha en la casilla correspondiente. Una vez distribuidos 10 de los 11 números, se empieza a jugar.

2. Por turno, lanzan los dados cada uno de los jugadores. Si la suma de los dados es uno de los números escogidos por el lanzador, se adelanta la ficha correspondiente una casilla.

3. Si la suma de los dados es el número que no ha sido escogido por ninguno de los dos adversarios, el jugador del turno escoge una de sus fichas (la que quiera) y adelanta una casilla.

4. Si la suma de los dados es un número del adversario, las fichas quedan como están.

5. Gana el jugador que consigue llevar una de sus fichas hasta la meta.

La realización del juego, con intervenciones explícitas del profesor y una organización de la clase adecuada permiten introducir, de manera intuitiva, objetos fundamentales tales como:

- *Combinatoria o recuento sistemático*. Estrategias de control sobre el recuento de los casos, para no contar por *exceso* (contar más de una vez un mismo caso) ni por *defecto* (omitir casos). Por ejemplo, las duplas (n, m) y (n + 1, m–1) suman lo mismo.
- *Análisis de datos*. Toma, ordenación, síntesis y visualización de datos, cuyo fin es inferir alguna conjetura, que tendrá que ser validada por un análisis experimental o teórico.
- *Probabilidad*. Esperanza de que un suceso ocurra, para, en particular, tomar decisiones justificadas.
- *Frecuencia absoluta*. Número de veces que ha ocurrido un determinado suceso, repetido un experimento aleatorio un número finito de veces en las mismas circunstancias.

- *Frecuencia relativa*. Relación entre la frecuencia absoluta y el número de veces que se ha realizado el experimento.
- *Ley del azar*. Si un experimento aleatorio se repite "muchas" veces, las frecuencias relativas de un suceso determinado se aproximan a la probabilidad (teórica) de que dicho suceso ocurra. Esta ley permite predecir resultados y, por lo tanto, es un indicativo para la toma de decisiones. Así, la ley del azar es un "puente tendido" entre el pasado y el futuro, entre las frecuencias (lo que ha ocurrido) y las probabilidades (lo que se espera suceda).

Cuestiones

- Si la *combinatoria es el arte de contar sin enumerar directamente todos los casos*, ¿qué situaciones podríamos utilizar para aprender técnicas de ordenación, colocación, selección, etc., de objetos? ¿Es necesario y pertinente introducir las nociones de permutaciones, variaciones y combinaciones antes de proponer situaciones de recuento sistemático?
- ¿Es necesario introducir la probabilidad de manera axiomática (álgebra de Boole, Kolmogorov, etc.)?

3. Contraste de Hipótesis

> The good schoolmaster is known by the number of valuable subjects that he declines to teach[1]. (*The Future in Education*, Livingstone, 1941)

La recta interpretación de la máxima de Livingstone no es conformista y de renuncia. Muy al contrario, exige la búsqueda de una *transposición didáctica* (Chevallard, 1991) adecuada para la enseñanza de una determinada noción, proceso o significado matemático.

---

[1] Se reconoce al buen maestro por el número de temas valiosos que se abstiene de enseñar.

Supongamos la siguiente situación. Se pregunta a un grupo de personas que estimen la edad de una persona (P). Una vez hecha la estimación, se les pide que vuelvan a estimarla conociendo una información que se presume relevante (S). En el anexo se muestran los datos y todas las órdenes que justifican los resultados que explicitaremos a continuación, editadas con el *software* R de libre distribución[2].

En la figura 2 se puede observar una exploración de los datos obtenidos de la muestra de 134 personas. En la figura 2a) se observa que los diagramas de cajas están a diferentes alturas, lo que sugiere la existencia de diferencias significativas entre las medianas de las edades con y sin la información relevante. En particular, la mediana de las edades en la estimación S es superior a la P; sin embargo, en la estimación P se observa una menor simetría. En la figura 2b) se representan las edades medias en los extremos de la barra vertical y en la figura 2c) se representan las medianas. A simple vista, tanto las medias como las medianas están bastante separadas.

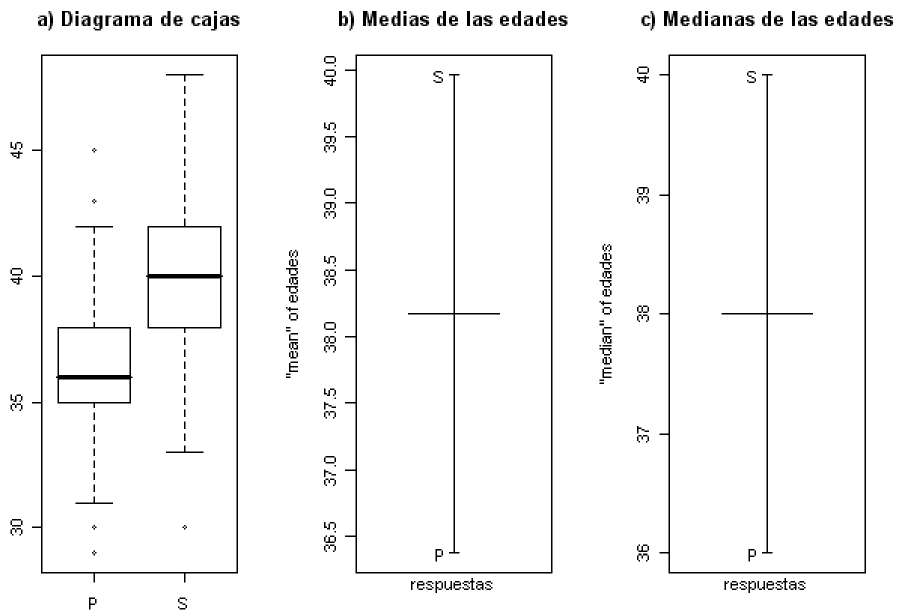

Figura 2. Exploración de los datos

---

[2] Recuperable en: http://www.r-project.org/

El análisis exploratorio de los datos debe ser confirmado. Para ello planteamos un modelo con estructura unifactorial en diseño completamente aleatorizado: $y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$, donde μ es la media de los k niveles (en nuestro caso, 2: P y S); α_i, el efecto sobre la respuesta debido al i-ésimo tratamiento (desconocimiento o conocimiento de la información); y, por último, ε_{ij} es el error experimental para la j-ésima observación bajo el i-ésimo tratamiento, tal que $\varepsilon_{ij} \sim N\ (0, \sigma^2)$. Para este modelo planteamos las hipótesis:

$$\begin{cases} H_0: & \mu_P = \mu_S \quad \text{o} \quad \alpha_\text{i} = 0, \forall i = P,S \\ H_1: & \mu_P \neq \mu_S \quad \text{o} \quad \exists \alpha_\text{i} \neq 0 \end{cases}$$

La variable respuesta y representa la edad estimada. En estas condiciones, el análisis de la varianza (ANOVA de los datos) detecta diferencias significativas, aceptándose la hipótesis alternativa (H_1) de la diferencia de medias. El cálculo del intervalo de confianza es otro modo de rechazar la hipótesis nula (H_0), sin necesidad de hacer el ANOVA de los datos.

Es necesario, empero, verificar la idoneidad del modelo. Para ello se calculan los residuales estandarizados y se determina cuántos son en valor absoluto mayores que $z_{0,95} = 1{,}96$. La suposición de normalidad en la población exige que haya un máximo de éstos valores (que se presumen *atípicos* u *outliers*). Si más del 5% del valor absoluto de los residuales estandarizados es mayor que 1,96, el modelo debe ser rechazado.

En nuestro caso el valor es aproximadamente 5%. Por lo tanto, el modelo queda en entredicho. Para validarlo, realizamos el diagrama de cajas y el gráfico de percentiles de la normal o qq-norm (figura 3). La simetría de la caja y el buen ajuste a la normal permiten entonces aceptarlo. Asimismo, se observan valores presumiblemente atípicos u outliers.

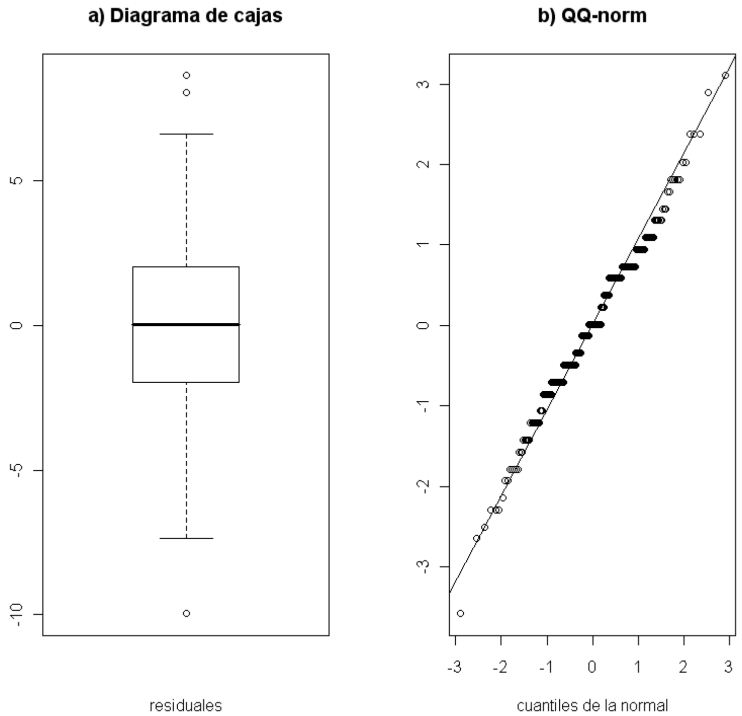

Figura 3. Validación del diseño completamente aleatorizado

Por último, el test de Levene permite aceptar la igualdad de varianzas con un nivel de significación de 0,95.

Criterios de idoneidad didáctica

Godino, Wilhelmi y Bencomo (2005) proponen tres dimensiones para analizar la idoneidad de un proceso de estudio matemático: *epistémica*, *cognitiva* e *instruccional.* Estas idoneidades deben ser integradas teniendo en cuenta los procesos que describen sus interacciones, lo cual requiere hablar de la *idoneidad didáctica* como criterio sistémico de pertinencia (adecuación al proyecto de enseñanza) de un proceso de estudio en una determinada institución. El principal indicador empírico de esta idoneidad puede ser la adaptación entre los significados personales logrados por los estudiantes y los significados institucionales pretendidos/implementados. Estos criterios orientan o "guían" la práctica educativa, pero no *aseguran* el logro de su idoneidad. A su vez, dichos criterios se pueden usar para valorar la idoneidad de procesos de estudio efectivamente implementados.

Cuestiones

- Valora la idoneidad didáctica de un proceso de estudio basado en el estudio matemático expuesto atendiendo a cada una de las dimensiones indicadas.
- Se desea trabajar con dicho ejemplo en el primer ciclo de educación secundaria, ¿es posible en las condiciones actuales de la institución de referencia realizar una transposición didáctica viable?

4. Resolución de Problemas: ilusiones y control

La modelización del saber mediante situaciones (Brousseau, 1998) es una forma de aportar un instrumento de control (*a priori*) de los procesos de estudio y de valoración (*a posteriori*) de las situaciones de enseñanza puestas en marcha. Este proceso de control y valoración es complejo. Se precisa de un instrumento que aporte indicadores precisos y objetivos. En esta sección mostramos un esquema para la descripción de situaciones de enseñanza que permite prever el funcionamiento de los conocimientos en los sistemas didácticos y valorar las interacciones de los niños con el saber y el medio. La construcción sistemática de estos esquemas permite:

- *Actuar*: *qué hacer en un aula*. Las prácticas de enseñanza requieren de pautas concretas de intervención en el sistema didáctico y de criterios de control sobre el funcionamiento de los objetos matemáticos en el medio y del sentido atribuido a los mismos.
- *Planificar: cómo diseñar una situación de enseñanza.* Las prácticas de planificación deben fundamentarse en presupuestos teóricos basados en otras investigaciones y en experiencias realizadas previamente. la *prueba de la contingencia* permite extraer datos que permiten prever tanto las intervenciones del profesor como los comportamientos esperados de los alumnos.
- *Analizar: cómo mejorar la práctica docente.* Es necesario contar con instrumentos que permitan el contraste del

funcionamiento de los sistemas didácticos según criterios preestablecidos, que sobrepasen una observación "naturalista" de los sistemas didácticos.

Las situaciones no son (y sería una pretensión inapropiada para la práctica docente) *situaciones fundamentales* ni necesariamente con carácter *adidáctico* (Brousseau, 1998). Son "situaciones potenciales de enseñanza". No tienen *a fortiori* que fundamentarse en desarrollos teóricos previos ni contrastada su pertinencia *a priori* con una prueba experimental. La importancia de la modelización de conocimientos mediante situaciones se centra en la posibilidad de disponer de un instrumento de valoración y análisis de la práctica docente. La confección de situaciones permite prever *a priori* el funcionamiento del sistema didáctico y planificar estrategias de control del mismo. La experimentación arroja consonancias y disonancias con las previsiones hechas. La comparación entre lo sucedido y lo previsto aporta conocimiento sobre las condiciones de difusión y de adquisición de saberes matemáticos.

La *ingeniería didáctica* (Artigue, 1989) se utiliza pues como fuente de referencia, sin la pretensión de elaboración de un corpus científico ni la obtención de funcionamientos estables del medio (*reproducibilidad de situaciones didácticas*). Con otras palabras, estas situaciones no pueden asegurar (aunque eventualmente podrían determinarlo) el progreso de la didáctica como instrumento técnico-práctico (*intervención crítica en los sistemas didácticos*) ni tampoco como disciplina científica (*prueba de la contingencia*).

Wilhelmi y Lacasta (2007) proponen un esquema de organización de situación de enseñanza: papel del profesor, actividad de los alumnos y función de las matemáticas en la actividad matemática. Este esquema es un instrumento objetivo para el contraste entre un análisis *a priori* y la realización efectiva. Está constituido por 4 bloques: 1) *Presentación*: descripción de la situación, sin abordar explícitamente el problema del alumno y del profesor. 2) *Desarrollo*: descripción pormenorizada de la situación según las tareas y responsabilidades de profesor y alumnos. 3) *Análisis didáctico*:

descripción de la situación según las dificultades previstas (necesidades de aprendizaje), los conocimientos matemáticos (tareas, técnicas, modos de justificación y significados de los objetos), los instrumentos de control del sistema didáctico previstos, etc. 4) *Contextualización*: justificación de la adecuación de la situación en la institución (referencias a leyes y disposiciones generales, su desarrollo y su aplicación o *propuestas de enseñanza*).

Cuestión. Construcción de una situación para la enseñanza del contraste de hipótesis introducido en la sección 3 para el primer ciclo de secundaria.

Reconocimiento: Este taller se realiza en el marco del proyecto SEJ2007-60110/EDUC.

Referencias

Artigue M. (1989) Ingénierie didactique. *Recherches en Didactique des Mathématiques*, 9(3), 282–307.

Brousseau G. (1998). Théorie des situations didactiques. La Pensée Sauvage : Grenobe.

Chevallard Y. (1991). *La transposición didáctica. Del saber sabio al saber enseñado*. Aique: Buenos Aires.

Chevallard Y., Bosch M., Gascón J. (1997). Estudiar matemáticas. El eslabón perdido entre enseñanza y aprendizaje. ICE-Horsori: Barcelona.

Godino J. D., Wilhelmi M. R., Bencomo D. (2005). Suitability criteria for a mathematical instruction. A teaching experience with the function notion. *Mediterranean journal for research in mathematics education*, 4(2), 1–26.

Livingstone R. (1941). *The Future in Education*. Cambridge University Press: London.

National Council of Teachers of Mathematics (NCTM) (2003). *Principios y estándares para la educación matemática*. Sevilla:

Thales. Traducción: M. Fernández, *Principles and Standards for School Mathematics* (NCTM, 2000).

Wilhelmi M. R. (2004). *Combinatoria y probabilidad*. GEEUG del Departamento de Didáctica de la Matemática (Universidad de Granada): Granada.

Wilhelmi M. R., Lacasta E. (2007). Un modelo docente para la formación en Geometría de maestros en Educación Infantil. En M. Camacho, P. Flores y P. Bolea (Eds.), *Investigación en educación matemática XI*. SEIEM-CajaCanarias: Tenerife.

ANEXO: órdenes con R para la obtención de gráficos y valores

```
############################################
########################Funciones adicionales
descriptiva<-function(variable){
    tamaño<- length(variable)
    rango<-range(variable)
    media<-mean(variable)
    varianza<-var(variable)
    destip<-sd(variable)
    print(list(tamaño=tamaño,rango=rango, media=media,
    varianza=varianza,desv.tip=destip))}
res1.std<- function(a, b = 3, d = 0.95) {
    c <- round((residuals(a) *
    sqrt(a$df.residual/deviance(a))), b);
    outlier <- as.character(ifelse(abs(c) > qnorm(d + (1 -
    d)/2), "***", " "))
    list(rs = c, o = outlier)}
############################################
########################################Datos
p<-
c(36,34,38,35,33,43,41,35,32,40,38,40,36,30,42,36,45,38,35,3
7,33,39,35,
    32,31,38,34,41,30,40,38,38,38,39,35,33,39,35,37,38,36,33
    ,34,35,40,38,29,
    38,34,35,35,35,34,39,35,36,39,36,31,35,35,38,35,33,35,40
    ,38,40,33,36,36,
    34,36,35,39,34,38,35,39,34,37,37,38,36,39,43,33,36,40,38
    ,42,43,38,38,36,
    38,35,34,36,38,33,34,38,35,38,38,38,39,33,35,37,30,34,36
    ,37,36,35,33,35,
    35,38,39,36,39,36,39,38,37,35,37,38,35,34,32)
s<-
c(38,40,38,40,38,44,41,39,43,40,40,44,36,30,43,42,48,43,37,4
3,33,42,40,
    35,36,42,45,41,40,40,40,38,42,42,40,36,41,37,40,39,42,41
    ,34,42,42,39,35,
    40,40,42,40,42,42,39,45,38,43,42,40,38,38,44,42,38,39,43
```

```
      ,42,42,37,42,38,
      39,40,36,39,35,42,40,41,39,41,43,40,38,39,43,36,38,41,39
      ,42,45,38,38,39,
      42,37,40,38,38,42,36,45,35,38,40,42,39,38,42,43,38,42,42
      ,40,38,40,38,35,
      41,40,41,40,43,41,39,40,40,40,40,45,38,41, 40)
############################################
##################### Descripción de los datos
descriptiva(p)
descriptiva(s)
edades<-c(p,s)
respuestas<-factor(c(rep("P",134),rep("S",134)))
datos<-data.frame(edades,respuestas)
par(mfrow=c(1,3))
      boxplot(split(edades,respuestas))
      title("a) Diagrama de cajas")
      plot.design(edades~respuestas, fun="mean", xlab=" ")
      title("b) Medias de las edades")
      plot.design(edades~respuestas, fun="median", xlab=" ")
      title("c) Medianas de las edades")
############################################
###################################ANOVA
edades.aov<-aov(edades~respuestas,data=datos)
summary(edades.aov)
############################################
###############################Int. confianza
t.test(edades[respuestas=="P"],edades[respuestas=="S"])$co
nf
############################################
###################################Outliers
cbind(datos,ajustado=fitted(edades.aov),
resid=round(residuals(edades.aov),1),resstd=res1.std(edades
.aov,2))
############################################
#######################Validación del diseño
par(mfrow=c(1,2))
      boxplot(resid(edades.aov), xlab="residuales")
      title("a) Diagrama de cajas")
```

```
    qqnorm(resid(edades.aov)/sqrt(7.72), xlab="cuantiles de
    la normal", ylab="",main="b) QQ-norm")
    qqline(resid(edades.aov)/sqrt(7.72))
############################################
########################Igualdad de varianzas
trat<-c(rep("P",134),rep("S",134))
x.mediana<-tapply(edades,trat,mean)
dife<-abs(c(edades[trat=="P"]-x.mediana["P"],
edades[trat=="S"]-x.mediana["S"]))
levene<-aov(dife~trat)
summary(levene)
```