

Perspectiva de la construcción hotelera en el mediterráneo europeo

Agustín ALONSO RODRÍGUEZ
Real Centro Universitario
«Escorial-María Cristina»
San Lorenzo del Escorial
aalonso@rcumariacristina.com

Resumen: Las nuevas tecnologías de la comunicación, y en concreto internet, están afectando profundamente a la industria turística. Ahora son los propios individuos los que planifican sus viajes, compran los billetes, hacen sus reservas de alojamiento, etc. Como consecuencia de estos cambios, han surgido múltiples formas de alojamiento, al margen de los tradicionales hoteles. Comprobar si esta nueva situación tiene efectos negativos inmediatos en la construcción de hoteles, en la zona turística del Mediterráneo europeo, es el objetivo del presente trabajo.

Abstract: The new communication technologies are having a serious impact in the tourist sector of the economy. Individuals are now the planners of their travels, they buy the tickets directly, they make reservations, and so on. As a consequence, new forms of lodging, leaving out the classic hotels, are strongly emerging. To see if this new situation is affecting, in the very short run, the construction of new hotels in the Mediterranean side of Europe, is the topic of this paper.

Palabras clave: Nuevas tecnologías de la comunicación, Internet, Turismo, Eurostat, Construcción de hoteles, NLME, Modelos de efectos mixtos, Previsión, R.

Key words: New communication technologies, Internet, Tourism, Eurostat, Hotels' construction, NLME, Mixed effects models, Forecasting, R.

Sumario:

- I. Introducción.**
- II. La industria turística.**
- III. Los datos.**
- IV. Modelos estadísticos.**
- V. Modelo lineal inicial.**
- VI. Modelo de efectos mixtos.**
- VII. Previsiones.**
- VIII. Conclusión.**
- IX. Bibliografía.**

I. INTRODUCCIÓN

En el presente trabajo se estudia la evolución de la construcción de hoteles en los países del área mediterránea europea, a saber: Grecia, España, Francia e Italia.

En el marco de referencia que las nuevas tecnologías y gustos suponen para la industria turística, se quiere vislumbrar si la construcción de hoteles es una realidad creciente o da señales de retroceso en el muy corto plazo.

II. LA INDUSTRIA TURÍSTICA

Viajar se ha convertido en nuestros días en una actividad con profundas repercusiones en todos los ámbitos de la vida humana, tanto a nivel privado como estatal.

Trabajo y placer parecen ser los principales motivos para viajar, y como consecuencia inmediata de los viajes, la hostelería está experimentando actualmente un notable desarrollo.

No es este el lugar para entrar en detalles sobre los factores que han impulsado el auge de la actividad turística, y en concreto el auge de la construcción de hoteles. El lector en español puede consultar la *Enciclopedia práctica profesional de turismo, hoteles y restaurantes*, editada por Océano/Centrum (1999) como un primer encuentro con el sector turístico. Aquí nos limitaremos a señalar algunos aspectos que han impulsado la presente situación.

En primer lugar, la construcción de hoteles lleva consigo grandes volúmenes de capital, lo que ha facilitado el nacimiento de grupos y cadenas hoteleros, así como el de grupos especializados de inversión.

Cabe también destacar la relación entre compañías aéreas y hoteles, una consecuencia de querer proveer a los viajeros de las líneas

aéreas de lugares de alojamiento dotados de grandes comodidades, y, a ser posible, en hoteles dirigidos por compatriotas, a fin de facilitar la comunicación lingüística y el atractivo de los lugares.

Por último, mencionar el tema referente a las ferias, exhibiciones y congresos, que ha obligado a los hoteles a ser algo más que lugares de alojamiento.

Frente a estas notables características de la industria turística, hoy, gracias a las nuevas tecnologías de la comunicación y en concreto gracias a internet, se nota un cambio en los gustos y aptitudes de los viajeros, que parecen poner en jaque algunos de los aspectos de la industria turística, hasta hoy desarrollada.

Así, por ejemplo, los viajeros se han convertido en sus propias agencias de viajes. Ellos gestionan la adquisición de los billetes, los lugares de alojamiento, las reservas y los tiempos de permanencia, y, en definitiva, la independencia personal, que busca objetivos no considerados suficientemente por la industria turística, como pueden ser alojamientos más baratos que permitan visitar más lugares, reduciendo los tiempos de permanencia, etc.

Ante esta nueva situación, ¿cabe pensar que la tendencia en la construcción de hoteles va a seguir creciendo en los países eminentemente turísticos del área mediterránea europea, o, por el contrario, se va a reducir en el futuro inmediato?

Este es el tema a considerar en el presente artículo, utilizando los datos de Eurostat para Grecia, España, Francia e Italia.

III. LOS DATOS

Conviene destacar la gran abundancia de datos disponibles en la actualidad, tanto a nivel mundial como nacional y regional.

La Organización Mundial del Turismo establecía para 2005 un total de 808 millones de turistas a nivel mundial, lo que representaba un incremento medio de más del 5,5 por 100 respecto al año anterior. Para Europa el crecimiento medio en el período 2000-2005 era estimado superior al 4,9 por 100, traducido en 444 millones de visitantes. Limitándonos al año 2005, último considerado por las estadísticas, Europa experimentó un crecimiento del 4 por 100, con un aumento en términos absolutos de 18 millones de llegadas nuevas. Y

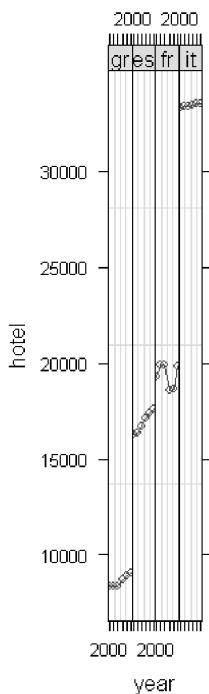
dentro de Europa, la Europa meridional y mediterránea experimentó un crecimiento del 6,2 por 100. El lector puede consultar el informe publicado por Exceltur, en su página web: *exceltur 2003*.

Eurostat recoge los datos de los países miembros de la Unión Europea, y los clasifica según numerosos criterios. En lo que a nosotros respecta, nos limitaremos a los hoteles y establecimientos semejantes, quedando fuera de nuestra consideración otros tipos de alojamientos. Con los datos de Eurostat, el número de hoteles actualmente existentes es el siguiente:

Años:	2000	2001	2002	2003	2004	2005
Grecia: gr	8.342	8.342	8.329	8.689	8.899	9.036
España: es	16.287	16.369	16.732	17.102	17.402	17.607
Francia: fr	19.315	19.928	19.889	18.598	18.689	19.811
Italia: it	33.361	33.421	33.411	33.480	33.518	33.527

Su representación gráfica viene recogida en la figura 1.

FIGURA 1



Saltan a la vista las diferencias existentes en el número de hoteles de cada país en el período 2000-2005.

IV. MODELOS ESTADÍSTICOS

En el establecimiento de modelos para datos experimentales u observados se trata de caracterizar la dependencia de la variable explicada o respuesta en su relación con una o más variables explicativas o covariables. Asimismo, se pretende caracterizar la variación no explicada de la variable respuesta.

Los modelos empíricos, es decir, los modelos derivados de los mismos datos y no de supuestos sobre el proceso generador de las observaciones, se prefieren lineales en los parámetros por ser más simples de entender e interpretar que los no lineales.

Algunas de las variables pueden ser repetidas en el sentido, al menos conceptual, de que es posible obtener nuevas observaciones con los mismos valores de las covariables. Un factor es repetible si el conjunto de sus posibles niveles es fijo, y cada nivel es, en sí, repetible.

El objetivo al modelizar covariables repetibles y no repetibles es diferente.

En el caso de covariables repetibles se busca caracterizar el cambio en la respuesta entre diferentes niveles, y para esto se utiliza el modelo de *efectos-fijos*, que representan la tasa de cambio de la respuesta con respecto al tiempo. Con variables no repetibles se quiere caracterizar la variación inducida en la respuesta por los diferentes niveles de las covariables, y para ello se utilizan *efectos-aleatorios*.

Un modelo que incorpore ambos efectos se denomina un modelo de *efectos mixtos*, Bates (2005), Littell et al. (2000), Pinheiro y Bates (2000), Verbeke y Molenberghs (2000). No es éste el único nombre para estos modelos. Así tenemos: modelos de panel: Frees (2004), Wooldridge (2002); modelos longitudinales: Hand y Crowder (1999), Frees (2004), Singer y Willet (2003); modelos jerárquicos: Bryk y Raudenbush (1992), Gelman y Hill (2007); modelos multinivel: Goldstein (2003), Gelman y Hill (2007) y modelos para datos repetidos: Davis (2003).

Hay que señalar la falta de terminología uniforme al caracterizar los efectos mencionados. Así se dice que los efectos son fijos si son interesantes en sí mismos, y son aleatorios si el interés se centra en la población subyacente. También se afirma que los efectos fijos son los coeficientes de las variables en un modelo de regresión que no son en sí modelizados; por ejemplo, un modelo de regresión con datos de 19 ciudades como regresores se considera un modelo de efectos fijos o de efectos fijos en relación a las 19 ciudades. También se consideran efectos fijos los coeficientes que no cambian en el grupo. Como la terminología no está asentada, Gelman y Hill (2007) prefieren denominar a los modelos de efectos mixtos como *modelos multinivel*, cf. o.c., p. 246.

Dado que tenemos varias observaciones por cada país, la técnica apropiada para nuestro estudio es la de los modelos de *efectos mixtos o multinivel*.

En una situación como la presente, tenemos un factor agrupante, *país*, al que se asocian cuatro niveles o naciones concretas: gr, es, fr e it. Cabe analizarla mediante un modelo de efectos-fijos y/o un modelo de efectos-aleatorios, según queramos hacer inferencia respecto a los cuatro niveles concretos del estudio o respecto a la población de la que consideramos los cuatro países son una muestra.

Entre los paquetes estadísticos disponible para estos modelos utilizaremos aquí el denominado *nlme*, de Pinheiro y Bates (2000), en su versión para R, y en su actualización de 2006. Entre las cualidades específicas de este paquete hay que mencionar sus representaciones gráficas, que permiten guiar la búsqueda del modelo más apropiado.

V. MODELO LINEAL INICIAL

Para resaltar la importancia de tener presente al factor de agrupamiento, olvidemos la estructura de grupo aquí presente y consideremos como una única muestra las 24 observaciones a nuestro alcance. Es decir, estimemos un modelo simple como el siguiente:

$$y_{ij} = \beta_1 + \beta_2 x_{ij} + u_{ij}; i = 1 \dots M; j = 1 \dots n_i \quad [1]$$

Siendo y_{ij} el número de hoteles contabilizados por observación j en el país i ; β_1 el número medio de hoteles en la población muestreada; x_{ij} la variable explicativa, en este caso el año (*year*), en el que se hizo la medición para cada país, y u_{ij} el término de error asociado con cada país y año, que supondremos con componentes independientes y distribución $N(0, \sigma^2)$.

Estimando el modelo por mínimos cuadrados ordinarios, tenemos:

```

modelo1 = lm(hotel ~ year, data = datos)
summary(modelo1)

Call:
lm(formula = hotel ~ year, data = datos)

Residuals:
    Min       1Q   Median       3Q      Max
-11207  -5000  -1642   3792  14028

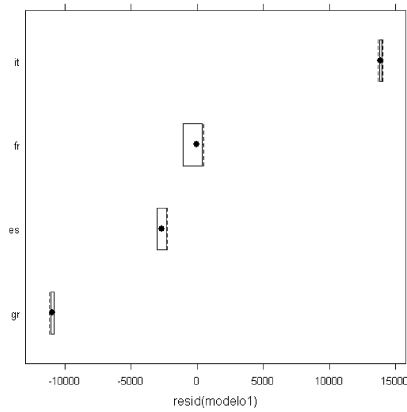
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -183981.6  2237841.6  -0.082   0.935
year           101.7     1117.5   0.091   0.928

Residual standard error: 9350 on 22 degrees of freedom
Multiple R-Squared: 0.000376,    Adjusted R-squared: -0.04506
F-statistic: 0.008275 on 1 and 22 DF,  p-value: 0.9283

```

Poco hay que decir de este primer modelo, salvo resaltar el alto valor estimado la varianza del término de error: $\hat{\sigma}^2 = 9350^2$. El gráfico de los residuos, figura 2, pone de relieve el problema de olvidar la estructura agrupada de los datos: residuos no centrados en cero y con muy diferentes varianzas. Es el efecto del factor agrupante *país*.

FIGURA 2



Cabe incorporar el efecto del factor *país* permitiendo que la ordenada en origen y la pendiente sean diferentes para cada país. Es el modelo de *efectos fijos*:

$$y_{ij} = \beta_{1i} + \beta_{2i}x_{2ij} + u_{ij}; i = 1 \dots M; j = 1 \dots n_i \quad [2]$$

Siendo ahora β_{1i} el número medio de hoteles para el país i , y β_{2i} la pendiente para cada país i , teniendo u_{ij} las mismas características que en [1].

Aplicando mínimos cuadrados ordinarios, obtenemos:

```

modelo2.lis = lmList(hotel.new)
> summary(modelo2.lis)
Call:
  lmerMod(<lm>, data = hotel.new, fixed = year | pais)

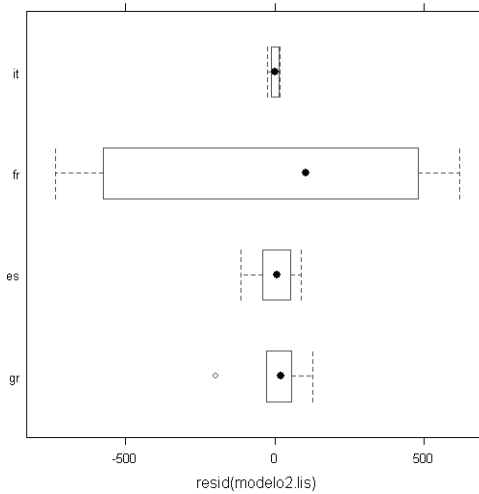
Coefficients:
  (Intercept)
  Estimate Std. Error t value Pr(>|t|)
  gr -306129.6  162251.9 -1.8867556 0.077475039
  es -559174.1  162251.9 -3.4463342 0.003319161
  fr  164009.4  162251.9  1.0108320 0.327140057
  it -34632.0  162251.9 -0.2134459 0.833675223
  year
  Estimate Std. Error t value Pr(>|t|)
  gr  157.17143  81.02462  1.9397984 0.070241713
  es  287.68571  81.02462  3.5505962 0.002663154
  fr -72.22857  81.02462 -0.8914398 0.385903085
  it  34.00000  81.02462  0.4196255 0.680339311

Residual standard error: 338.9503 on 16 degrees of freedom
    
```

Por el momento, nos limitamos a señalar que los valores medios representados en las ordenadas en el origen son muy diferentes, y lo mismo ocurre con las pendientes, si bien la varianza estimada del término de error: $\hat{\sigma}^2 = 338.95^2$ es ahora mucho menor.

Si analizamos y representamos los residuos, obtenemos la figura 3.

FIGURA 3



Ahora los residuos aparecen centrados en torno al cero, evidenciándose la gran variabilidad presente en los mismos.

VI. MODELO DE EFECTOS MIXTOS

Aunque el modelo de efectos fijos [2] tiene en cuenta el efecto *país*, no obstante, no da una visión completa de la situación. El problema fundamental es que sólo se modeliza la muestra concreta de estos cuatro países, mientras que también interesa la población de la que se ha obtenido la muestra. Es decir, el modelo [2] no brinda una estimación de la variabilidad entre países. Otro inconveniente que presenta el modelo [2] es que el número de parámetros aumenta con el número de países.

Un modelo que incorpore *efectos-aleatorios* obvia ambos inconvenientes, considerando el efecto *país* como variaciones aleatorias en torno a la media poblacional.

Para ello, reescribamos el modelo [2] sumando y restando $\bar{\beta}$:

$$y_{ij} = \bar{\beta} + (\beta_{1i} - \bar{\beta}) + \beta_{2i}x_{2ij} + u_{ij} \quad [3]$$

siendo, en nuestro caso, $\bar{\beta} = \frac{1}{24} \sum_{i=1}^{24} \beta_{1i}$ la media de las ordenadas en el origen.

El modelo de efectos aleatorios sustituye $\bar{\beta}$ por la media sobre la población, y reemplaza $(\beta_{1i} - \bar{\beta})$ por una variable aleatoria, b_i , cuya distribución hay que estimar.

Es decir, el modelo en [3] pasa a escribirse:

$$y_{ij} = \beta + b_i + \beta_{2i}x_{2ij} + u_{ij} \quad [4]$$

siendo:

β : ordenada media poblacional

b_i : variable aleatoria que representa las desviaciones respecto a la ordenada media poblacional del i -ésimo *país*; es el efecto aleatorio o variabilidad *between* países;

u_{ij} : variable aleatoria que representa la desviación para la observación j en el país i respecto a la ordenada media poblacional: es la variabilidad *within* o interna.

Ahora

$$b_i : i = 1 \dots M;$$

$$u_{ij} : i = 1 \dots M; j = 1 \dots n_i$$

Y comenzaremos suponiendo que ambas variables son independientes entre si, con media cero y varianza constante, es decir:

$$b_i \sim N(0, \sigma_b^2)$$

$$u_{ij} \sim N(0, \sigma^2)$$

supuestos a modificar si las circunstancias lo requieren.

Este modelo tiene dos fuentes de variabilidad, b_i y u_{ij} .

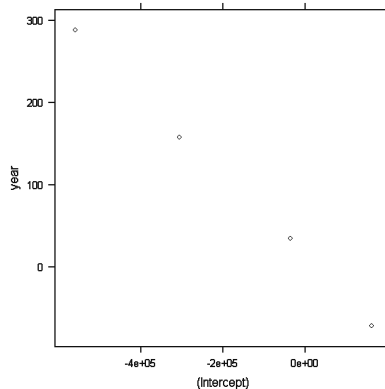
b_i se conoce como el efecto-aleatorio (*random effect*). El nombre se debe a su asociación con el experimento concreto en estudio, *países* en nuestro caso, que se considera una muestra tomada aleatoriamente de la población de interés. Se dicen efectos porque representan desviaciones respecto a la media global. En otras palabras, el efecto de elegir el país i consiste en desplazar el valor medio desde β a $\beta + b_i$. Dado que las observaciones corresponden al mismo país, tienen en común b_i , el efecto aleatorio, y estarán correlacionadas. La covarianza entre observaciones del mismo país es σ_b^2 y la correspondiente correlación viene dada por: $\sigma_b^2 / (\sigma_b^2 + \sigma^2)$, cf. Pinheiro y Bates (2000), p. 8; Gelman y Hill (2007), p. 258.

Los parámetros del nuevo modelo son β , σ_b^2 y σ^2 , es decir, tres, y siempre será así, a diferencia del modelo de *efectos fijos*, que veía aumentar el número de parámetros en función de los grupos.

Aunque los efectos aleatorios $b_i, i=1 \dots n_i$ puedan comportarse como parámetros, formalmente son niveles de variación en el modelo, y, por esto, no son estimados, sino predichos \hat{b}_i a partir de los valores de la muestra.

Con ayuda del paquete *nlme* en R, de Pinheiro y Bates, disponemos de una primera ayuda visual como guía para la determinación de la presencia de los *efectos aleatorios*, que es uno de los objetivos de los modelos *lm.List*, presentes en *nlme*. Se trata de la función *pairs*, que nos da una visión de la estructura de covarianzas en el modelo. Aplicada a nuestro modelo anterior, nos da como resultado el gráfico de la figura 4:

FIGURA 4



El gráfico muestra una correlación negativa entre el término constante y la pendiente del modelo. Esta alta correlación es debida a que los datos han sido recogidos a partir del año 2000, mientras que la ordenada en el origen conceptualmente se corresponde con el año 0. Cabe obviar esta situación centrando los datos, ajustándolos en torno al 2002.5. Al hacerlo, reestimamos el anterior modelo, y obtenemos el *modelo2bis.lis*:

```
modelo2bis.lis = update(modelo2.lis, hotel ~ I(year - 2002.5))
> summary(modelo2bis.lis)
```

Call:

```
Model: hotel ~ I(year - 2002.5) | pais
Data: hotel.new
```

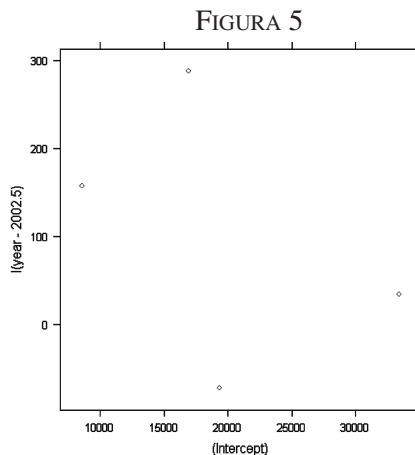
Coefficients:

```
(Intercept)
Estimate Std. Error t value Pr(>|t|)
gr 8606.167 138.3759 62.19412 0
es 16916.500 138.3759 122.25035 0
fr 19371.667 138.3759 139.99308 0
it 33453.000 138.3759 241.75455 0

I(year - 2002.5)
Estimate Std. Error t value Pr(>|t|)
gr 157.17143 81.02462 1.9397984 0.070241713
es 287.68571 81.02462 3.5505962 0.002663154
fr -72.22857 81.02462 -0.8914398 0.385903085
it 34.00000 81.02462 0.4196255 0.680339311
```

Residual standard error: 338.9503 on 16 degrees of freedom

Ahora, al aplicar la función *pairs*, aparece el gráfico de la figura 5:



Ha desaparecido la mencionada correlación.

Llegados aquí, podemos utilizar la función *intervals* como guía en la determinación de los posibles *efectos aleatorios* a tener presentes en el modelo. Al invocarla, tenemos como resultado:

```
intervals(modelo2bis.lis)
, , (Intercept)

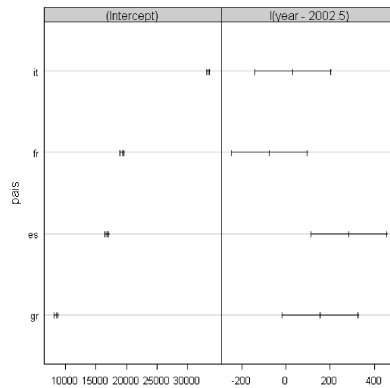
      lower      est.      upper
gr 8312.823 8606.167 8899.51
es 16623.156 16916.500 17209.84
fr 19078.323 19371.667 19665.01
it 33159.656 33453.000 33746.34

, , I(year - 2002.5)

      lower      est.      upper
gr -14.59310 157.17143 328.93595
es 115.92119 287.68571 459.45024
fr -243.99310 -72.22857 99.53595
it -137.76452 34.00000 205.76452
```

Para captar mejor el mensaje de estos intervalos, representémoslos; es la figura 6:

FIGURA 6



La falta de sobreposición o solapamiento en los intervalos que muestra esta figura es indicio de que es apropiado incluir *efectos aleatorios* en la constante y también en la pendiente del modelo.

El modelo de *efectos fijos* y *efectos aleatorios* tiene, en el paquete *nlme*, la formulación:

```
modelo3.lme = lme(hotel ~ I(year - 2002.5), datos, random = ~ I(year - 2002.5)|pais)
```

en esta formulación, la respuesta es *hotel*, hay unos efectos fijos en la ordenada en el origen o término constante del modelo y en la pendiente, y con `random = ~I(year-2002.5)|pais` se indica la presencia de efectos aleatorios en ambos términos, siendo *país* el factor aglutinante.

Al estimar el modelo, designado como *modelo3.lme*, obtenemos los siguientes resultados:

```
modelo3.lme = lme(hotel ~ I(year - 2002.5), data = hotel.new, random =
2002.5)|pais)

> summary(modelo3.lme)
Linear mixed-effects model fit by REML
Data: hotel.new
      AIC      BIC    logLik
367.4111 373.9573 -177.7055

Random effects:
Formula: ~I(year - 2002.5) | pais
Structure: General positive-definite, Log-Cholesky parametrization
      StdDev      Corr
(Intercept) 10327.3933 (Intr)
I(year - 2002.5) 132.6741 -0.488
Residual      338.9503

Fixed effects: hotel ~ I(year - 2002.5)
      Value Std.Error DF t-value p-value
(Intercept) 19586.833 5164.160 19 3.792840 0.0012
I(year - 2002.5) 101.657 77.729 19 1.307835 0.2065
Correlation:
      (Intr)
I(year - 2002.5) -0.417

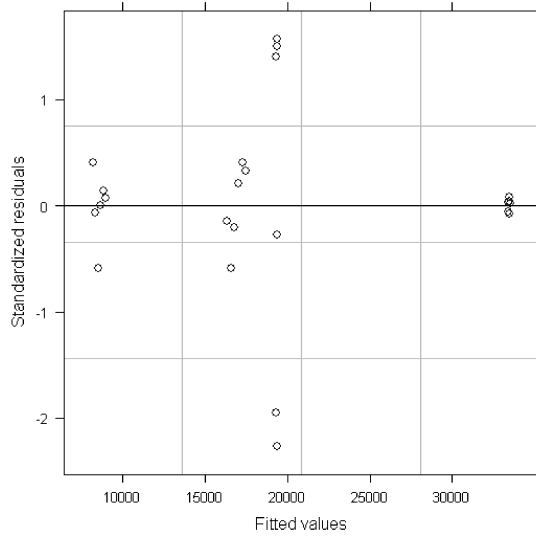
Standardized Within-Group Residuals:
      Min      Q1      Med      Q3      Max
-2.26418484 -0.15822808 0.02884873 0.23675618 1.57339446

Number of Observations: 24
Number of Groups: 4
```

Los residuos estandarizados de este modelo presentan el siguiente aspecto, figura 7:

```
plot(modelo3bis.lme)
```

FIGURA 7



que reflejan la heteroscedasticidad presente en los mismos, más claramente visible en las figuras 8 y 9:

FIGURA 8

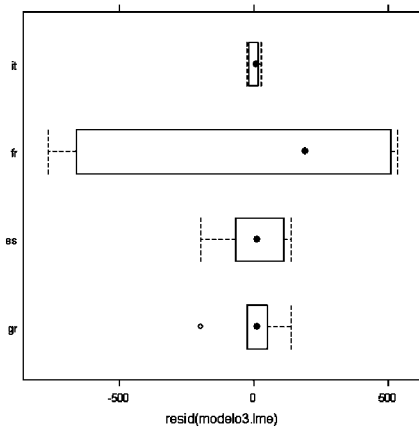
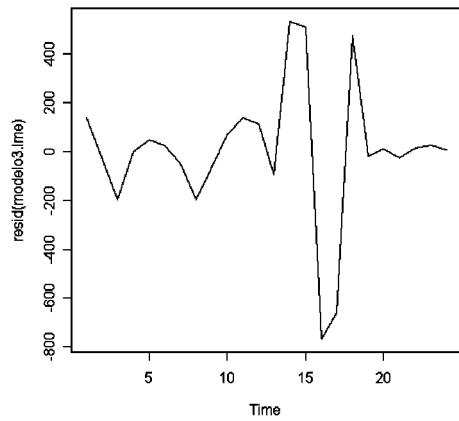


FIGURA 9



Para seguir adelante es necesario reestimar el modelo corrigiendo la heteroscedasticidad. Antes, sin embargo, indicar que el *modelo3.lme* ha sido estimado mediante RMLE (*Restricted Maximum Likelihood*), que tiende a obtener estimaciones más *conservadoras* de las varianzas de los componentes. Si se prefiere estimarlo por máxima verosimilitud, basta con precisar el *método*, obteniendo:

```

modelo3.ML = update(modelo3.lme, method = "ML")

> summary(modelo3.ML)
Linear mixed-effects model fit by maximum likelihood
Data: hotel.new
      AIC      BIC    logLik
396.4002 403.4685 -192.2001

Random effects:
Formula: ~I(year - 2002.5) | pais
Structure: General positive-definite, Log-Cholesky parametrisation
              StdDev   Corr
(Intercept)  8943.5311 (Intr)
I(year - 2002.5) 107.5201 -0.522
Residual      338.9502

Fixed effects: hotel ~ I(year - 2002.5)
              Value Std.Error DF t-value p-value
(Intercept)  19586.833  4671.166 19  4.193136  0.0005
I(year - 2002.5)  101.657    70.309 19  1.445866  0.1645
Correlation:
              (Intr)
I(year - 2002.5) -0.417

Standardized Within-Group Residuals:
              Min      Q1      Med      Q3      Max
-2.29358409 -0.16012042  0.03453114  0.27959469  1.65729301

Number of Observations: 24
Number of Groups: 4
>

```

Como era de esperar, las estimaciones de las desviaciones estándar de los efectos aleatorios son menores que las obtenidas por REML: 8943.5311 y 107.5201, frente a 10327.3933 y 132.6741, respectivamente. Continuamos con el *modelo3.lme*.

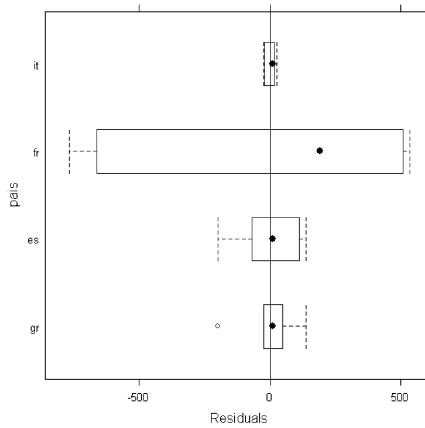
Validación del modelo3.lme

Comencemos por el análisis de los residuos. La instrucción

```
plot(modelo3.lme, pais ~ resid(.), abline =0)
```

nos devuelve la figura 10

FIGURA 10

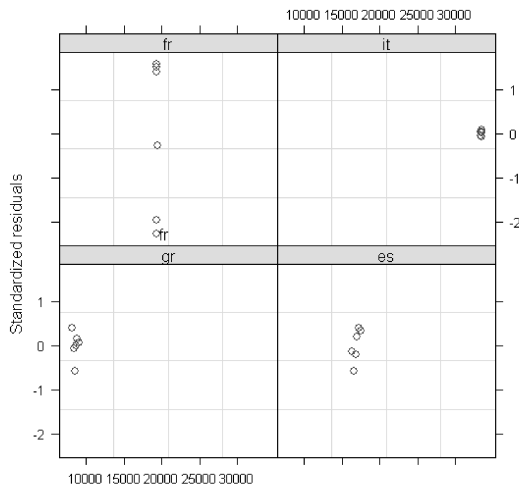


que claramente muestra la heteroscedasticidad presente en estos residuos. Podemos ver la situación, condicionada a cada *país* con la instrucción:

```
plot(modelo3.lme, resid(., type = "p") ~ fitted(.) |país, id=0.05, adj = -0.3)
```

que genera la representación de la figura 11:

FIGURA 11



Reestimemos el modelo corrigiendo la heteroscedasticidad. Es el nuevo modelo: **Nmodelo3b.lme**.

La instrucción:

```
Nmodelo3.lme = update(modelo3.lme, weights=varIdent(form = ~1 |pais))
```

nos da el resultado buscado:

```
summary(Nmodelo3.lme)

Linear mixed-effects model fit by REML
Data: hotel.new
      AIC      BIC    logLik
335.8618 345.6812 -158.9309

Random effects:
Formula: ~I(year - 2002.5) | pais
Structure: General positive-definite, Log-Cholesky parametrization
              StdDev      Corr
(Intercept)  10329.1606 (Intr)
I(year - 2002.5)  118.6388 -0.565
Residual      123.4188

Variance function:
Structure: Different standard deviations per stratum
Formula: ~1 | pais
Parameter estimates:
      gr      es      fr      it
1.0000000 0.6889113 5.5368868 0.1543462
Fixed effects: hotel ~ I(year - 2002.5)
              Value Std.Error DF  t-value p-value
(Intercept)  19587.69  5165.074 19  3.792334  0.0012
I(year - 2002.5)  140.94   64.633 19  2.180605  0.0420
Correlation:
              (Intr)
I(year - 2002.5) -0.519

Standardized Within-Group Residuals:
      Min      Q1      Med      Q3      Max
-1.5932572 -0.5270679  0.2145705  0.7305800  1.1332865

Number of Observations: 24
Number of Groups: 4
```

El aspecto global de los nuevos residuos viene recogido en la figura 12, y su función de autocorrelación en la figura 13.

```
ts.plot(resid(Nmodelo3.lme, type = "pearson"))
```

FIGURA 12

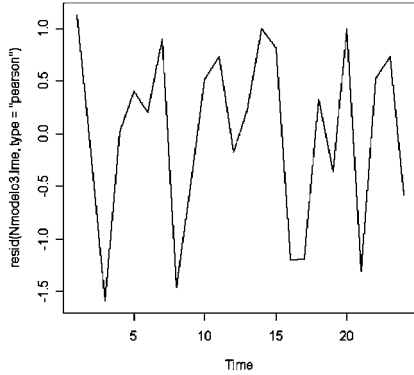
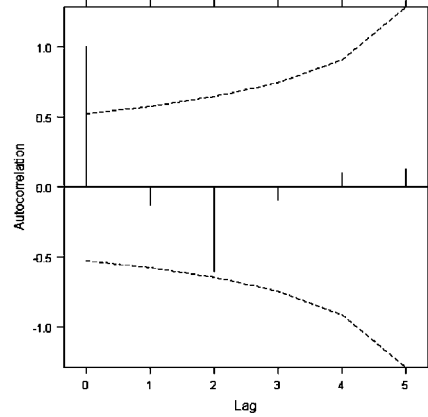
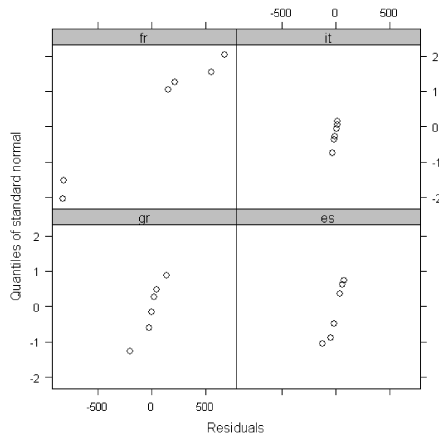


FIGURA 13



Los residuos son, pues, homoscedásticos e independientes entre sí. Y en lo que respecta a su distribución normal, los residuos presentan el aspecto que muestra la figura 14:

FIGURA 14



Al comparar ambos modelos

```
anova(modelo3.lme, Nmodelo3.lme)
      Model df      AIC      BIC    logLik    Test L.Ratio p-value
modelo3.lme    1    6 367.4111 373.9573 -177.7055
Nmodelo3.lme    2    9 335.8618 345.6812 -158.9309 1 vs 2 37.54931 <.0001
```

El bajo valor del *p-value* indica que el **Nmodelo3.lme** explica los datos mejor que el *modelo3.lme*.

Hemos llegado a un modelo estadísticamente aceptable, y podemos pasar a su análisis.

Comencemos por los *efectos fijos*. Sus estimaciones son:

```
Fixed effects: hotel ~ I(year - 2002.5)
              Value Std.Error DF  t-value p-value
(Intercept)  19587.69  5165.074  19  3.792334  0.0012
I(year - 2002.5)  140.94   64.633  19  2.180605  0.0420
```

coeficientes todos estadísticamente significativos, al 95 por 100 de confianza.

El modelo estimado de *efectos fijos*, el modelo promediado sobre los cuatro países, lo podemos escribir como:

$$\hat{y} = 19587.69 + 140.94x$$

$$(t=3.79) \quad (t=2.18)$$

representando \hat{y} el número de hoteles ajustado y x la variable años (*years*).

Las varianzas estimadas son: $\hat{\sigma}_b = 10329.1606$ y $\hat{\sigma}_y = \hat{\sigma}_u = 123.4188$.

Con la ayuda de la instrucción `coef(Nmodelo3.lme)`, obtenemos las estimaciones de las líneas de regresión de cada país:

```
coef(Nmodelo3.lme)
(Intercept) I(year - 2002.5)
gr      8606.515      161.75344
es     16916.429      282.33744
fr     19374.810       85.63218
it     33452.993       34.03663
```

y con la instrucción `random.effects(Nmodelo3.lme)` obtenemos los *efectos aleatorios*:

```
random.effects(Nmodelo3.lme)
(Intercept) I(year - 2002.5)
gr -10981.1719      20.81351
es -2671.2576      141.39752
fr -212.8768      -55.30774
it  13865.3063     -106.90329
```

que sumados a los *efectos fijos* anteriores nos permite establecer el nivel y pendiente de cada país, estimaciones que coinciden con las estimaciones dadas por la instrucción `coef`. Así, por ejemplo, para Grecia, *gr*, tendríamos:

$$\hat{y} = (19587.69 - 10981.1719) + (140.94 + 20.81351)x = 8606.515 + 161.75344 x$$

para España, *es*:

$$\hat{y} = (19587.69 - 2671.2576) + (140.94 + 141.39752)x = 16916.429 + 282.33744 x$$

y así sucesivamente.

VII. PREVISIONES

Dado que tenemos un modelo estadísticamente aceptable, pasemos a realizar la previsión o predicción con el mismo. Los valores futuros de la variable x , los años, para cada uno de los países, son: 2006, 2007 y 2008, y con la instrucción:

```
predict(Nmodelo3.lme, nuevos, level=0:1)
```

obtenemos las previsiones siguientes:

	pais	predict.fixed	predict.pais
1	gr	20080.98	9172.652
2	gr	20221.92	9334.405
3	gr	20362.86	9496.159
4	es	20080.98	17904.610
5	es	20221.92	18186.947
6	es	20362.86	18469.285
7	fr	20080.98	19674.522
8	fr	20221.92	19760.155
9	fr	20362.86	19845.787
10	it	20080.98	33572.121
11	it	20221.92	33606.158
12	it	20362.86	33640.194

En la columna *predict.fixed* se muestran las predicciones relativas al global, a la población, y en la columna *predict.pais*, las correspondientes a cada *pais* en el período 2006-2008.

VIII. CONCLUSIÓN

Llegados a este punto, procede comparar, para cada país, las tasas de incremento porcentual del número de hoteles durante el período 2000-2005, sobre la base de los datos de Eurostat, y para el período 2006-2008, basados en las anteriores previsiones.

Período 2000-2005:

Grecia: 8,3% de incremento.
España: 8,1% de incremento.
Francia: 2,6% de incremento.
Italia: 0,5% de incremento.

Período 2006-2008:

Grecia: 3,5% de incremento.
España: 3,2% de incremento.
Francia: 0,8% de incremento.
Italia: 0,2% de incremento.

Estas estimaciones tendenciales indican una desaceleración en las previsiones para la construcción de nuevos hoteles durante el período 2006-2008.

IX. BIBLIOGRAFÍA

- BATES, D. M., «Fitting linear mixed models in R», *R News*, 5 (mayo 2005) 27-30.
- BRYK, D., y RAUDENBUSH, S. W., *Hierarchical Linear Models: Applications and Data Analysis Methods*, Sage Publications, Newbery Park, California 1992.
- DAVIS, C. S., *Statistical Methods for the Analysis of Repeated Measurements*, Springer, New York 2003.
- EXCELTUR 2003, <http://www.exceltur.org>
- EUROSTAT, <http://epp.eurostat.ec.europa.eu> (consulta hecha el 03/12/2006).
- FREES, E. W., *Longitudinal and Panel Data, Analysis and Applications in the Social Sciences*, Cambridge University Press, New York 2004.
- GELMAN, A., y HILL, J., *Data Analysis Using Regression and Multilevel/Hierarchical Models*, Cambridge University Press, New York 2007.
- GOLDSTEIN, H., *Multilevel Statistical Models*, Arnold Publishers, London 2003.

- HAND, D., y CROWDER, M., *Practical Longitudinal Data Analysis*, Chapman & Hall/CRC, Boca Raton, Florida 1999.
- LITTELL, R. C.; MILLIKEN, G. A.; STROUP, W. W., y WOLFINGER, R. D., *SAS System for Mixed Models*, cuarta reimpression, SAS Institute Inc., Cary, NC 2000.
- PINHEIRO, J. C., y BATES, D. M., *Mixed-Effects Models in S and S-PLUS*, Springer, New York 2000.
- R, Development Core Team (2006).
- R: *A language and environment for statistical computing*.
R Foundation for Statistical Computing.
Viena, Austria, 2006.
- URL <http://www.R-project.org>
- SINGER, J. D., y WILLETT, J. B., *Applied Longitudinal Data Analysis*, Oxford University Press, New York 2003.
- VARIOS, *Enciclopedia Práctica Profesional de Turismo, Hoteles y Restaurantes*, Océano, Centrum, Barcelona 1999.
- VERBEKE, G., y MOLENBERGHS, G., *Linear Mixed Models for Longitudinal Data*, Springer, New York 2000.
- WOOLDRIDGE, J. M., *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge, Massachusetts 2002.