

ÁREA CIENTÍFICA: MARKETING, LOGÍSTICA E DISTRIBUIÇÃO

## **TURISMO NA REGIÃO NORTE DE PORTUGAL: APLICAÇÃO DO MODELO LINEAR GERAL**

Natália Sofia Matos dos Santos, [nspink@hotmail.com](mailto:nspink@hotmail.com), Escola Superior de Tecnologia e Gestão do Instituto Politécnico de Bragança

Paula Odete Fernandes, [pof@ipb.pt](mailto:pof@ipb.pt), Escola Superior de Tecnologia e Gestão do Instituto Politécnico de Bragança

### **RESUMO**

Pretendeu-se com o presente trabalho construir um modelo econométrico que melhor explicasse o comportamento da procura turística na Região Norte de Portugal. A série “Dormidas Mensais, nos estabelecimentos hoteleiros, na região Norte de Portugal”, foi a variável considerada para tal explicação. O período em referência para o estudo compreendeu a série construída desde 1990 até 2006. Para dar resposta ao principal objectivo construíram-se os modelos econométricos, assentes no modelo linear geral estocástico, estático e dinâmico. Confrontando os resultados de ambos os modelos chegou-se à conclusão que o melhor modelo para explicar o comportamento da procura turística na região Norte de Portugal foi o modelo dinâmico, pois produziu os melhores resultados, garantindo as hipóteses básicas do modelo linear geral estocástico e contendo os melhores estimadores obedecendo às propriedades do mesmo, sendo os melhores estimadores lineares não enviesados.

**PALAVRAS-CHAVE:** Turismo, Dormidas Mensais, Modelos Econométricos, Modelo Linear Geral.

### **ABSTRACT**

The main objective of this work was to build an econometric model that best explained the behaviour of tourism demand in northern of Portugal. Monthly Nights in hotel establishments in the north of Portugal, was the variable considered for such an explanation. The reference period for the study comprised a series built from 1990 to 2006. To meet the main objective were built econometric models, based on the general linear stochastic model, static and dynamic. Comparing the results of both models, static and dynamic, it was concluded that the best model to explain the behaviour of tourism demand in the North of Portugal was the dynamic model, it produced the best results ensuring the basic assumptions of general linear stochastic model and containing the best indicators complying with the same properties, and the best linear unbiased estimators.

**KEYWORDS:** Tourism, Nights Monthly, Econometric Models, General Linear Model.

## 1. INTRODUÇÃO

O turismo tem vindo a afirmar-se a nível mundial como um dos principais sectores económicos, que mais riqueza proporciona a uma nação.

Também em Portugal, a importância global do turismo na economia pode ser documentada através de diversos indicadores. O turismo para Portugal, sem dúvida, que é também um factor de equilíbrio das contas públicas, tendo as receitas turísticas um peso significativo na Balança Corrente. O turismo assume, deste modo, uma importância estratégica na economia portuguesa, gerando receitas significativas e revelando um conjunto de vantagens competitivas, em regra, não alcançáveis por outras actividades (Turismo de Portugal, 2009).

De referir que sobre o ponto de vista económico, considera-se que o turismo abrange todas as deslocações de pessoas, quaisquer que sejam as suas motivações, que obriguem ao pagamento de prestações e serviços durante a sua deslocação e permanência temporária fora da sua residência habitual superior ao rendimento que, eventualmente, auferam nos locais visitados (Cunha, 2003).

Relativamente à região Norte, pode dizer-se que é uma região que coloca à disposição de quem a visita uma grande variedade de opções no que diz respeito à oferta de produtos turísticos, que vai desde sol e mar, montanha, passando por estâncias termais até ao turismo rural. Este tipo de turismo oferece uma alternativa interessante e singular ao chamado turismo de massas que habitualmente encontra-se associado a impactos ambientais mais negativos (Fernandes, 2005).

Na área da modelação da procura turística, actualmente, encontra-se disponível uma grande multiplicidade de métodos que têm vindo a surgir para fazer face às mais variadas situações, apresentando características e metodologias diferentes. Seguindo esta linha de raciocínio, o principal objectivo deste estudo prende-se com a elaboração de um modelo econométrico que permita explicar o comportamento da procura turística na região Norte de Portugal. A variável dormidas mensais na região Norte de Portugal, para o período compreendido entre 1990 e 2006. No sentido de melhor explicar esta variável vão utilizar-se como variáveis independentes as seguintes: permanência média, variável *dummy* Páscoa, e as taxas de câmbio dos principais mercados emissores (Espanha, Reino Unido, Alemanha e França).

Em suma, o principal objectivo é obter um modelo que melhor explique o comportamento da variável dormidas na região Norte de Portugal, e verificar também se as variáveis escolhidas são as mais indicadas para explicar a variável dependente dormidas na região Norte de Portugal.

Para tal, o presente trabalho divide-se da seguinte forma. Na secção seguinte apresenta-se e explica-se o modelo econométrico utilizado e que serve de base a toda a análise efectuada, incorporando uma secção se apresentam resultados, e no último ponto apresenta-se uma conclusão.

## 2. A IMPORTÂNCIA DA MODELAÇÃO DA PROCURA TURÍSTICA

O principal objectivo da modelação de uma série económica, do sector do turismo, prende-se com a construção de um modelo que melhor se adequa à explicação e ao comportamento da série objecto de estudo, procurando-se também avançar com dados quantitativos para o futuro, ajudando assim os gestores e autoridades competentes a realizarem um planeamento turístico o mais rigoroso possível.

Nesta linha, nos últimos anos assistiu-se a um aumento do número de estudos científicos publicados tendo por base a modelação e previsão da procura turística (Witt & Witt, 1995; Lim, 1997; Thomakos & Guerard, 2004). O crescente interesse desta área de estudo tem-se relacionado com a rápida expansão e desenvolvimento do turismo Internacional e das economias de um país. Contudo existe um largo número de factores de índole sócio-culturais, económicos, políticos e tecnológicos, que podem influenciar a procura turística negativa ou positivamente. Planear debaixo destas circunstâncias torna-se singularmente difícil mas importante.

Vários são os autores que têm vindo a contribuir para o emergir de diferentes metodologias, robustas, de modelação e previsão, utilizando diferentes abordagens para solucionar diferentes problemas associados

aos modelos, desde os mais simples aos mais complexos (Makridakis & Hibon, 1997; Goh & Law, 2002; Thawornwong & Enke, 2004; Yu *et al.* 2006; Moutinho *et al.* 2008; Fernandes *et al.* 2008).

Assim, a modelação e previsão da procura turística, embora não sendo a única questão fulcral para os gestores e autoridades competentes, responsáveis pelo planeamento turístico de um local, região ou país, torna-se de fundamental importância para o planeamento turístico, recorrendo-se assim a diferentes métodos e instrumentos de modelação e previsão que permitam melhorar o rigor da previsão da procura turística (Preez & Witt, 2003).

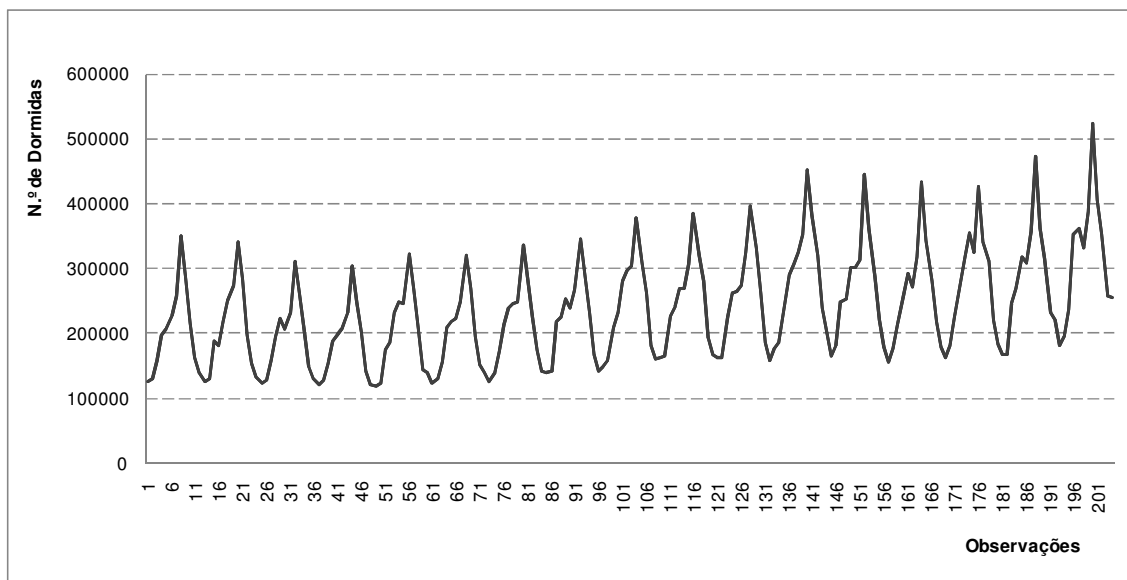
Com o mercado turístico cada vez mais competitivo, composto por um significativo número de produtos que tentam acompanhar uma procura heterogénea e cada vez mais crítica, por parte do turista, e concorrências agressivas entre destinos turísticos, é de primordial importância, para um eficiente planeamento turístico, obter uma rigorosa modelação e previsão da procura turística. Estas são apenas algumas das razões pelas quais o planeamento turístico apoiado pela modelação e previsão da procura turística se torna necessário. Conhecer os resultados fiáveis da procura turística permite as autoridades competentes oferecer condições favoráveis não só às pessoas que visitam um determinado local, bem como contribuir para o desenvolvimento económico da região e do país.

### 3. TURISMO NA REGIÃO NORTE: APLICAÇÃO DO MODELO LINEAR GERAL

#### 3.1. VARIÁVEIS DO MODELO

O modelo econométrico que irá ser analisado e estudado é constituído por uma variável dependente que são as dormidas registadas na região Norte de Portugal.

As dormidas na Região Norte de Portugal consistem na permanência de um indivíduo num estabelecimento que fornece alojamento por um período compreendido entre as 12 horas de um dia e as 12 horas do dia seguinte (INE, 2007), ou seja, é o número total de dormidas registadas na região em estudo. A variável dormidas da Região Norte de Portugal tem como unidade de medida o número de dormidas, o mesmo é dizer o n.º de noites que o hóspede passou na região e tem um comportamento sazonal, tal como se pode observar na figura seguinte:



**Figura 1:** Dormidas nas Unidades de Alojamento, no Norte de Portugal, no intervalo [Jan-90:Dez-06].

Na Figura 1, pode observar-se a série temporal “Dormidas registadas na Região Norte de Portugal”, cujos dados estão disponíveis, no INE (2007). Os dados recolhidos consideram o período compreendido entre Janeiro de 1990 e Dezembro de 2006, correspondendo a 204 dados mensais.

Optou-se por estudar esta série uma vez que tem sido considerada como significativa da actividade turística devido às suas características e contempla os visitantes que usufruíram das capacidades turísticas da Região Norte, nesse período. Analisando o seu comportamento verifica-se a existência de oscilações periódicas em torno de um valor médio crescente, sugerindo a presença de sazonalidade (valores máximos nos meses de Verão e mínimos nos meses de Inverno), isto é, os valores das dormidas vêm em função da altura do ano.

O grande objectivo deste trabalho é analisar, caracterizar e descrever o seu comportamento através das seguintes variáveis explicativas:

- a componente autónoma, também conhecida por constante do modelo, esta variável não é influenciada pelas variáveis explicativas e os seus valores são desconhecidos, ou seja, não se pode quantificar esta componente através das variáveis explicativas;

- a permanência média, consiste numa relação entre o número de dormidas na região Norte de Portugal e o número de hóspedes que deram motivo a essas dormidas. A unidade de medida desta variável é o número de dias que um determinado indivíduo permanece num estabelecimento hoteleiro;

- a variável Páscoa consiste numa variável fictícia ou *dummy*<sup>1</sup> e caracteriza a procura turística nesta época festiva, ou seja, esta variável pode assumir o valor 1 no mês que se festeja a Páscoa e o valor 0 nos restantes meses do ano;

- as restantes variáveis explicativas referem-se às taxas de câmbio dos quatro principais mercados emissores, ou seja, as taxas de câmbio de Espanha, Reino Unido, Alemanha e França. Este tipo de taxas pode definir-se como taxas de câmbio de termo incerto até 1999 e após 1999 a termo certo, com excepção da libra do Reino Unido. A unidade de medida destas variáveis é a unidade monetária, visto que, até 1999 os dados são em relação ao escudo e a partir deste ano até 2006 é uma proporção do euro;

- por fim, este modelo tem o termo de erro ou termo estocástico cujos dados não são observáveis, ou seja, não existem dados para poder representar no modelo.

### 3.2. APLICAÇÃO DO MODELO ESTÁTICO OU RESTRITO

Um modelo restrito, não é mais do que uma representação das relações entre as variáveis no mesmo momento de tempo, pelo que o modelo estático com todas as variáveis explicativas tem a seguinte forma:

$$\begin{aligned} Dormidas_t = a + b_0 Permanência Média + b_1 Pascoa Dummy + b_2 Tx Alemanha \\ + b_3 Tx Espanha + b_4 Tx França + b_5 Tx Reino Unido + u_t \end{aligned} \quad (1)$$

#### 3.2.1. MODELOS DE ESTIMAÇÃO

Pretende-se com a aplicação deste modelo explicar as variações das diversas variáveis, verificar se este é o modelo mais adequado para explicar as dormidas na região Norte de Portugal, pelo que para tal irá efectuar-se várias estimações OLS<sup>2</sup>. As primeiras estimações OLS dizem respeito à escolha das melhores variáveis para este modelo, ou seja, escolher somente as variáveis que mais influenciam o modelo, pois isso provoca com que o modelo seja o que mais se aproxima do comportamento real da série.

Em seguida, na Tabela 1, apresenta-se a estimação, dos diversos modelos para possibilitar a escolha do melhor modelo com as variáveis mais relevantes.

<sup>1</sup> *Dummy*: esta variável admite que pode assumir apenas um de dois valores possíveis, regra geral 0 e 1.

<sup>2</sup> Do inglês *Ordinary Least Squares*.

**Tabela 1:** Medidas de desempenho dos diferentes modelos.

	<b>Modelo 1</b> Contém todas as variáveis Explicativas	<b>Modelo 2</b> Exclusão da componente autónoma	<b>Modelo 3</b> Exclusão da variável Páscoa
R <sup>2</sup>	0,200032	0,906305	0,906131
R <sup>2</sup> ajustado	0,175668	0,903939	0,904244
Estatística F	F (6,197) = 8,20999 (p-value =0,00001)	F (6,198) = 319,205 (p-value =0,00001)	F (5,199) = 384,196 (p-value =0,00001)
Estatística <i>Durbin Watson</i>	0,482836	0,413061	0,419168
Estatística $\chi^2$	p-value=0,0403168	p-value=0,150826	p-value=0,0437851
Multicolinearidade	VIF<10	VIF<10	VIF<10
Variáveis com significância estatística	Existem 3 variáveis	Somente uma variável	Existem 3 variáveis

Através da análise à tabela anterior, pode concluir-se que o modelo 1 apresenta valores relativamente baixos em relação às medidas de precisão do ajustamento quando comparados com os modelos 2 e 3, podendo levar a pensar-se que não é o modelo mais adequado. Contudo não é verdade, pois a constante e a variável permanência média são estatisticamente significativas, a um nível de significância de 1%. A variável taxa de Espanha é estatisticamente significativa, a um nível de 5%, ou seja, 95% do seu valor é o mais correcto, tal situação é igual para o modelo 3 enquanto que no modelo 2 somente a variável permanência média tem significância estatística a um nível de significância de 1%.

Em relação à análise das variáveis em conjunto apesar de existir alguma insignificância em algumas variáveis individuais, através do teste de *Fisher* pode observar-se que em conjunto as variáveis incluídas no modelo, explicam bem as variações ocorridas na variável dependente, situação idêntica para ambos os modelos.

Observa-se ainda a não infracção à hipótese da multicolinearidade para os três modelos, ou seja, existe ausência de independência entre as variáveis explicativas. O teste de *Durbin Watson* apresenta o valor mais elevado para o modelo 1, quanto mais próximo estiver de 2 melhor, pois significa que se encontra na zona de independência dos erros e que não há infracção à hipótese da independência do termo de erro.

Pela análise pode observar-se que existe uma possível infracção de ambos os modelos à hipótese de independência dos erros, contudo esta análise será efectuada com maior cuidado no desenrolar do trabalho, e serão utilizados os testes adequados para ultrapassar este problema. Em relação ao teste de  $\chi^2$  os modelos 1 e 3 seguem uma distribuição normal a um nível de significância de 5%, a hipótese da normalidade não é violada. Em relação à homocedasticidade, o modelo 1 e 3 são homoscedásticos, não existe infracção a esta hipótese pois o valor de prova é superior a 10% e aceita-se a hipótese da homocedasticidade enquanto que o modelo 2 é heterocedástico.

Depois da análise aos três modelos verifica-se que o melhor modelo é o 3 em que se inclui variáveis explicativas, componente autónoma, permanência média e as taxas de câmbio dos quatro principais países emissores, com a excepção da variável Páscoa, apesar das medidas de ajustamento terem valores relativamente baixos.

As variáveis incluídas no modelo em conjunto explicam bem as variações ocorridas na variável dormidas na região Norte de Portugal. Apresenta o valor mais elevado comparando com os restantes modelos em relação ao teste de *Durbin Watson*, é homoscedástico, segue uma distribuição normal a um nível de significância de 5%, existe ausência de independência entre as variáveis explicativas e este modelo contém três variáveis com significância estatística.

Assim, tendo por base a análise anterior nos próximos pontos toda a análise recairá sobre os valores apresentados pelo modelo 3.

### 3.2.2. ESTIMAÇÃO OLS - INTERPRETAÇÃO DOS COEFICIENTES

O coeficiente de determinação não ajustado é igual a 0,906131 e significa que as variáveis permanência média, taxas de câmbio dos quatro principais mercados emissores explicam cerca de 91% das variações que ocorrem nas dormidas na região Norte de Portugal. Em relação à componente autónoma indica que 167.542 dormidas na região Norte de Portugal não são explicadas pelas restantes variáveis independentes. Esta variável é estatisticamente significativa, a um nível de significância de 1%, ou seja, 99% do valor da variável autónoma é um valor correcto. Se a variável permanência média variar um dia, a variável dormidas na região Norte de Portugal aumenta em 46.420 noites, existindo uma relação positiva entre estas duas variáveis. Esta variável é estatisticamente significativa, a um nível de significância de 1% ou seja, 99% do seu valor correcto. Se a taxa de câmbio da Alemanha variar em uma unidade monetária, a variável dormidas na região Norte de Portugal variará em cerca de 0.13 noites, existindo uma relação positiva entre estas variáveis. Esta variável não tem significância estatística. Relativamente à taxa de câmbio de Espanha, se esta variar em uma unidade monetária, a variável dormidas na região Norte de Portugal varia inversamente em cerca de 3 noites, existindo uma relação negativa entre estas duas variáveis, uma vez que, se a taxa de câmbio espanhola diminuir os espanhóis ficam mais limitados e não poderão passar tantas noites em Portugal. Esta variável é estatisticamente significativa, a um nível de significância de 5% ou seja, 95% do valor da variável é correcto. Se a taxa de câmbio de França variar em uma unidade monetária, a variável dormidas na região Norte de Portugal diminui em cerca de 0.25 noites, existe uma relação inversa entre estas variáveis. Esta variável não tem significância estatística. A taxa de câmbio do Reino Unido não tem significância estatística e se esta variar em uma unidade monetária, a variável dormidas na região Norte de Portugal variará em cerca de 0.009 noites, existe uma relação positiva entre estas variáveis.

**Tabela 2:** Estimativas OLS para o período de 1990:01 a 2006:12, Variável dependente: Dormidas\_RN.

<i>Variável</i>	<i>Coefficiente</i>	<i>Erro Padrão</i>	<i>Estatística de t</i>	<i>p-value</i>
Constante	167.542	30.095	5,5671	0,00001
Permanência	46.420	13.095,9	3,5446	0,00049
Tx_Alemanha	0,135868	0,747343	0,1818	0,85592
Tx_Espanha	-3,00383	1,47253	-2,0399	0,04269
Tx_França	-0,252355	0,68216	-0,3699	0,71183
Tx_Reino Unido	0,00939718	0,0788137	0,1192	0,90521

### 3.2.3. TESTES DE HIPÓTESE

O teste de significância individual que se utiliza é a estatística de *t*, já analisada no ponto anterior, quando se fez referência à significância estatística das variáveis de forma individual. Em relação aos testes de significância em conjunto, obteve-se através da estimação OLS a estatística do F (5,199) = 384,196 (valor de prova <0,00001). Como o valor de prova é inferior a 10% não se rejeita a hipótese de que existem variáveis que assumem valores diferentes de zero e como já foi referido, anteriormente, as variáveis incluídas no modelo em conjunto explicam de forma satisfatória as variações ocorridas na variável dormidas na região Norte de Portugal.

### 3.2.4. INFRACÇÃO ÀS HIPÓTESES BÁSICAS DO MLG

É necessário analisar se as hipóteses básicas são garantidas, pois o modelo só é válido caso se garantam as hipóteses básicas. Para tal, irá analisar-se com especial cuidado cada situação.

Os valores apresentados na tabela seguinte permitem analisar com maior precisão se existe ausência de independência entre as variáveis explicativas. Através da estimação OLS obtém-se dos valores do VIF<sup>3</sup> (*Variance Inflation Factors*) para as diversas variáveis. Através da análise à tabela seguinte verifica-se

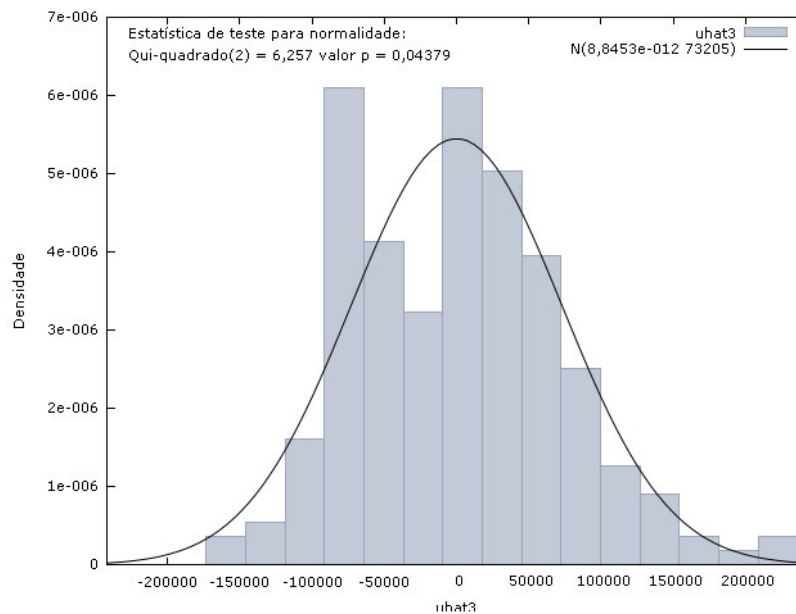
<sup>3</sup> *Variance inflation factors* serve para detectar a existência de relações exactas entre as variáveis independentes, e assim detectar multicolinearidade. Valores elevados sugerem que a variável explicativa está envolvida em multicolinearidade (Chaves, *et al.*; 2000).

que não existe infracção à hipótese básica da multicolinearidade, uma vez que, os valores do VIF para ambas as variáveis é inferior a 10 valores (caso fossem superiores a 10,0 podem indicar um problema de colinearidade). Pode concluir-se que existe ausência de independência das variáveis explicativas, ou seja, a variável permanência média, a variável taxa de câmbio de Espanha, Alemanha, França e Reino Unido não tem qualquer relação entre elas, com já tinha sido referido anteriormente e os estimadores permanecem BLUE<sup>4</sup>.

**Tabela 3:** *Variance Inflation Factors* do modelo estático.

Variáveis	VIF
Permanência média	1,321
Taxa da Alemanha	1,080
Taxa de Espanha	3,331
Taxa de França	3,674
Taxa do Reino Unido	3,958

Pela análise à Figura 1, pode observar-se que o teste da normalidade do resíduo feito através da estatística de teste  $\chi^2 = 6,25693$ , com valor de prova = 0,0437851, significa que este modelo segue uma distribuição normal a um nível de significância de 5%, logo esta hipótese não é violada. Verifica-se que a média apresenta um valor que é aproximadamente zero, pelo que então a hipótese da média zero também não é infringida. Teste de *White* para a heterocedasticidade através da estatística de teste  $TR^2 = 26,3038$  com valor de prova  $(\chi^2(20) > 26,3038) = 0,15601$ , como o valor de prova é superior a 10% conclui-se que não se rejeita a hipótese da homocedasticidade, de acordo com os resultados obtidos pode concluir-se que não existe infracção à homocedasticidade, isto é, a variância é constante de observação para observação. Não existe perda das características dos estimadores OLS, continuam a ser BLUE.



**Figura 2:** Distribuição normal do modelo estático.

<sup>4</sup> *Best Linear Unbiased Estimators*. As propriedades dos estimadores: (i) Não enviesados:  $E(\hat{b}) = b$ ; (ii) Eficientes: de entre os estimadores não enviesados têm variância mínima; (iii) Consistentes:  $P \lim_{x \rightarrow \infty} (\hat{b}) = b$ . (Gujarati, 1995).

O valor da estatística de *Durbin-Watson* de 0,419168, permite concluir que existe infracção à independência do termo de erro e que este modelo sofre de autocorrelação dos erros. Para tentar ultrapassar este problema, ou seja, tentar corrigir a infracção à hipótese da independência dos erros aplicou-se o teste de *Cochrane - Orcutt*, pelo que através da estimação obteve-se a seguinte estatística de *Durbin-Watson* = 1,28247; de facto melhorou mas não foi o suficiente pois continua a encontrar-se na zona de autocorrelação positiva dos erros. Um outro teste que se utiliza para tentar corrigir a infracção à hipótese básica da independência do termo de erro é o teste de *Hidreth - Lu*. Através da estimação obteve-se a seguinte estatística de *Durbin-Watson* = 1,28252. Melhorou em relação ao teste anterior mas continua a encontrar-se na zona de autocorrelação positiva dos erros. Este modelo sofre de autocorrelação dos erros, ou seja, os erros não são independentes entre si tendo como consequência que os estimadores dos mínimos quadrados não são os estimadores com variância mínima, isto é, não são eficientes embora permaneçam não enviesados. No entanto e no sentido de eliminar este problema podem ser aplicados testes mais avançados, sugerindo-se esta recomendação para futuros trabalhos.

### 3.3. APLICAÇÃO DO MODELO DINÂMICO

O modelo dinâmico representa as relações de uma determinada variável num determinado momento relacionado com variáveis relacionadas nos momentos anteriores.

O modelo dinâmico tem a seguinte apresentação:

$$\begin{aligned}
 Dormidas_t = & a + b_0 \text{Permanência Média} + b_1 \text{Tx Alemanha} + b_2 \text{Tx Espanha} \\
 & + b_3 \text{Tx França} + b_4 \text{Tx Reino Unido} + b_{5(t-1)} \text{Permanência Média} \\
 & + b_{6(t-1)} \text{Tx Alemanha} + b_{7(t-1)} \text{Tx Espanha} + b_{8(t-1)} \text{Tx França} \\
 & + b_{9(t-1)} \text{Tx Reino Unido} + b_{10(t-1)} \text{Dormidas RN} + u_{(t-1)}
 \end{aligned} \tag{2}$$

#### 3.3.1. ESTIMAÇÃO OLS – INTERPRETAÇÃO DOS COEFICIENTES

Uma vez que o modelo que produziu melhores resultados foi o 3, vai ser este que servirá de base para realizar a estimação do modelo dinâmico, embora com os períodos de tempo anteriores, uma vez que se pretende escolher o melhor modelo. Pois, tem lógica ambos os modelos terem as mesmas variáveis para que possam ser comparados.

**Tabela 4:** Estimativas OLS para o período de 1990:01 a 2006:12.

Variável	Coefficiente	Erro Padrão	Estatística de t	p-value
Constante	58913,3	21732,2	2,7109	0,00732
Permanência	60035,5	16591,7	3,6184	0,00038
Tx_Alemanha	-0,329066	0,483743	-0,6803	0,49717
Tx_Espanha	-0,344237	1,07501	-0,3202	0,74915
Tx_França	0,0527802	0,474397	0,1113	0,91153
Tx_Reino Unido	0,0264557	0,0592619	0,4464	0,65580
Permanência_1	-57423	16996,1	-3,3786	0,00088
Tx_Alemanha_1	-0,682499	0,484103	-1,4098	0,16022
Tx_Espanha_1	-0,371654	1,10837	-0,3353	0,73776
Tx_França_1	-0,358826	0,490661	-0,7313	0,46548
Tx_Reino_Un_1	-0,0140558	0,0562994	-0,2497	0,80312
Dormidas_RN_1	0,764745	0,0458907	16,6645	0,00001

A componente autónoma indica que 58913,3 dormidas na região Norte de Portugal não são explicadas pelas restantes variáveis independentes. Esta variável é estatisticamente significativa, a um nível de significância de 1% ou seja, 99% do valor da variável autónoma é um valor correcto. Se a variável



permanência média variar em um dia, a variável dormidas na região Norte de Portugal variará em 60.036 noites, existindo uma relação positiva entre estas duas variáveis. Esta variável é estatisticamente significativa, a um nível de significância de 1% ou seja, 99% do seu valor correcto.

Relativamente às variáveis das taxas de câmbio verifica-se o seguinte:

- se a taxa de câmbio da Alemanha variar em uma unidade monetária, a variável dormidas na região Norte de Portugal diminui em cerca de 0,32 noites, existe uma relação inversa entre as variáveis. Esta variável não tem significância estatística;

- se a taxa de câmbio de Espanha variar em uma unidade monetária, a variável dormidas na região Norte de Portugal diminui em cerca de 0,34 noites, existe uma relação inversa entre estas variáveis. Esta variável não tem significância estatística;

- se a taxa de câmbio de França variar em uma unidade monetária, a variável dormidas na região Norte de Portugal aumenta em cerca de 0.05 noites, existe uma relação positiva directa entre estas variáveis. Esta variável não tem significância estatística;

- se a taxa de câmbio do Reino Unido variar em uma unidade monetária, a variável dormidas na região Norte de Portugal aumenta em cerca de 0.09 noites, existe uma relação positiva entre estas variáveis. Esta variável não tem significância estatística;

- em relação às variáveis desfasadas se a variável permanência média variar um dia num determinado período de tempo, a variável dormidas na região Norte de Portugal varia inversamente em 57.423 noites, existindo uma relação negativa entre as variáveis. Esta variável é estatisticamente significativa, a um nível de significância de 1% ou seja, 99% do seu valor correcto. No entanto, se a taxa de câmbio da Alemanha variar em uma unidade monetária num determinado período de tempo, a variável dormidas na região Norte de Portugal varia inversamente em cerca de 0,68 noites, existe uma relação negativa entre as variáveis. Esta variável não tem significância estatística. A taxa de câmbio de Espanha se variar em uma unidade monetária num determinado período de tempo, a variável dormidas na região Norte de Portugal varia inversamente em cerca de 0,37 noites, existindo uma relação negativa entre as variáveis. Esta variável não tem significância estatística. Se a taxa de câmbio de França variar em uma unidade monetária num determinado período de tempo, a variável dormidas na região Norte de Portugal varia inversamente em cerca de 0,35 noites, existindo uma relação negativa entre as variáveis. Esta variável não tem significância estatística. A taxa de câmbio do Reino Unido se variar em uma unidade monetária num determinado período de tempo, a variável dormidas na região Norte de Portugal varia inversamente em cerca de 0,01 noites, existe uma relação negativa entre as variáveis. Esta variável não tem significância estatística. Ainda, as dormidas na região Norte de Portugal no período anterior ajudam a explicar cerca de 0.76 noites as variações ocorridas das dormidas na região Norte de Portugal no momento actual. Esta variável é estatisticamente significativa, a um nível de significância de 1% ou seja, 99% do valor da variável autónoma é um valor correcto.

**Tabela 5:** Medidas de desempenho dos diferentes modelos.

	<b>Modelo Dinâmico</b>
R <sup>2</sup>	0,680851
R <sup>2</sup> ajustado	0,66247
Estatística F	F (11,191) = 37,0423 (p-value =0,00001)
Estatística <i>Durbin Watson</i>	1,29352
Estatística $\chi^2$	p-value=0,0107027
Multicolinearidade	VIF<10
Variáveis com significância estatística	Existem 4 variáveis

Analisando a tabela anterior pode observar-se que o coeficiente de determinação não ajustado é igual a 0,680851 e significa que as variáveis permanência média, taxas de câmbio dos quatro principais mercados emissores explicam cerca de 68% das variações que ocorrem nas dormidas na região Norte de Portugal e que cerca de 32% das variações ocorridas nas dormidas da região Norte de Portugal não é explicado por qualquer factor conhecido. O coeficiente de determinação ajustado é igual a 0,66247 e significa que cerca de 66% das variações ocorridas nas dormidas da região Norte de Portugal foram explicadas pelas variações ocorridas nas variáveis independentes e que somente 34% das variações ocorridas nas dormidas não é explicada por nenhum factor conhecido.

### 3.3.2. TESTE DE HIPÓTESE

Os testes de significância individual, teste *t-student*, já foram analisados anteriormente, individualmente. Em relação aos testes de significância em conjunto, obteve-se uma estatística F (11,191) =37,0423 (valor de prova=0,00001), valor de prova inferior a 10%, aceita-se a hipótese que existem variáveis que assumem valores diferentes de zero. Apesar de existir alguma insignificância de algumas variáveis individuais, as variáveis no momento actual e no momento anterior incluídas no modelo, em conjunto explicam de forma satisfatória as variações ocorridas na variável dormidas na região Norte de Portugal.

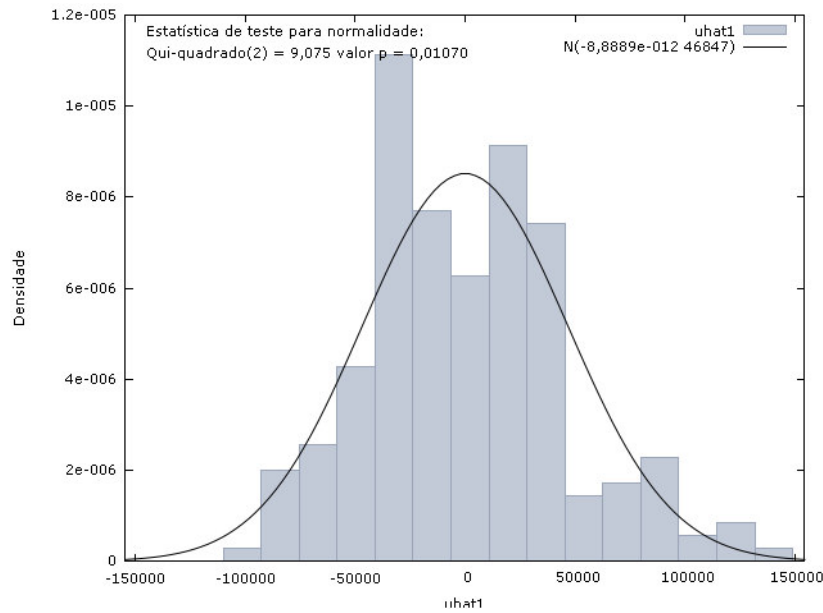
### 3.3.3. INFRACÇÃO ÀS HIPÓTESES BÁSICAS DO MLG

Através da análise à tabela seguinte verifica-se que não existe infracção à hipótese básica da multicolinearidade, uma vez que, os valores do VIF, para ambas as variáveis, no momento actual e no momento anterior é inferior a 10. Pode concluir-se que existe ausência de independência das variáveis explicativas, ou seja, a variável permanência média, a variável taxa de câmbio de Espanha, Alemanha, França e Reino Unido com as mesmas variáveis no período de tempo anterior não tem qualquer relação entre elas e os estimadores permanecem BLUE.

**Tabela 6:** *Variance Inflation Factors* (VIF) do modelo dinâmico.

<i>Variáveis</i>	<i>VIF</i>
Permanência Média	5,170
Taxa da Alemanha	1,105
Taxa de Espanha	4,306
Taxa de França	4,321
Taxa do Reino Unido	5,437
Permanencia_1	5,197
Tx_Alemanha_1	1,107
Tx_Espanha_1	4,587
Tx_Franca_1	4,623
Tx_Reino_Un_1	4,912
Dormidas_RN_1	1,273

Teste da normalidade do resíduo feito através da estatística de teste  $\chi^2(2) = 9,07451$ , com valor de prova=0,0107027, significando que este modelo segue uma distribuição normal a um nível de significância de 10%, logo esta hipótese não é infringida. Através da observação gráfica verifica-se o valor da média é aproximadamente zero então a hipótese da média zero também não é infringida. O teste de *White* para a heterocedasticidade através da estatística de teste  $TR^2 = 63,3619$  com valor de prova  $(\chi^2(55) > 63,3619) = 0,205263$ , como o valor de prova é superior a 10% conclui-se que não se rejeita a hipótese da homocedasticidade. De acordo com os resultados obtidos pode concluir-se que não existe infracção à homocedasticidade, isto é, a variância é constante de observação para observação. Não existe perda das características dos estimadores OLS, continuam a ser BLUE.



**Figura 3:** Distribuição normal do modelo dinâmico.

O valor da estatística de *Durbin-Watson* encontra-se na zona de autocorrelação positiva, pois o seu valor é de 1,29352, então pode concluir-se que existe infracção à independência do termo de erro e o modelo dinâmico sofre de autocorrelação dos erros. Para tentar ultrapassar este problema, ou seja, tentar corrigir a infracção à hipótese da independência dos erros irá efectuar-se o teste *Cochrane – Orcutt*. Através da estimação obteve-se a seguinte estatística de *Durbin-Watson*=1,88951, encontra-se na zona de independência dos erros e pode concluir-se que o modelo dinâmico não infringe a hipótese de independência dos erros, ou seja, os erros são independentes de observação para observação.

### 3.4. ESCOLHA DO MELHOR MODELO

De modo geral, pela interpretação dos resultados apresentada anteriormente o melhor modelo seria o dinâmico. Contudo para comprovar esta hipótese necessita-se de se efectuar o teste de significância das restrições lineares, contrastando as seguintes hipóteses:

$H_0$  : Modelo restrito

$H_1$  : Modelo dinâmico

Ou seja:

$$H_0 : b_7 = b_8 = b_9 = b_{10} = b_{11} = b_{12} = 0$$

$$H_1 : b_7, b_8, b_9, b_{10}, b_{11}, b_{12} \neq 0$$

**Tabela 7:** Conjunto de restrições.

1: b[Permanencia_1] = 0
2: b[Tx_Alemanha_1] = 0
3: b[Tx_Espanha_1] = 0
4: b[Tx_Franca_1] = 0
5: b[Tx_Reino_Un_1] = 0
6: b[Dormidas_RN_1] = 0

Pela estatística do teste do F ( $6,191$ ) = 48,2776, com valor de prova = 9,14621e-036, significa que existem evidências estatísticas suficientes para inferir que o modelo dinâmico é o que melhor se adequa à série em análise, uma vez que o valor de prova é inferior a 1%. As variáveis desfasadas contribuem para a explicação das dormidas na região Norte de Portugal.

#### 4. CONCLUSÃO

Este trabalho teve por objectivo construir um modelo para modelar a procura turística do norte de Portugal a partir de um modelo econométrico tendo por base o modelo linear geral estocástico, estático e dinâmico. Utilizaram-se dados mensais e foram recolhidos no site do Instituto Nacional de Portugal, cobrindo o período de 1990 a 2006. Para a construção do modelo utilizaram-se como variáveis explicativas a permanência média, taxa de câmbio e uma variável *dummy* Páscoa.

Assim, de entre os modelos encontrados e analisados, ao longo deste trabalho, o modelo OLS, na forma dinâmica, foi o modelo que melhores resultados produziu e se adequou à série original Dormidas na Região Norte de Portugal, comparando com os diferentes modelos estáticos analisados.

Concluiu-se que o modelo OLS estático necessita de corrigir a autocorrelação dos erros, com testes mais avançados, uma vez que autocorrelação apresentada e corrigida com os testes de *Hildreth Lu* e *Cochrane Orcutt* não foi suficiente para eliminar tal fenómeno, ou seja, este modelo apresentou dependências no termo de erro de observação para observação.

Uma vez que era também objectivo deste estudo trabalhar com um modelo dinâmico, construiu-se o modelo tendo-se concluído que o mesmo não violou as hipóteses básicas, apresentando um coeficiente de determinação e coeficiente de determinação ajustado de aproximadamente 68% e 66%, respectivamente, pelo que se entendeu ser um bom modelo, gerando estimadores BLUE (*Best Linear Unbiased Estimators*).

Deste modo concluiu-se, que não se deveria rejeitar o modelo dinâmico para prosseguir no estudo e assim ir de encontro ao objectivo deste trabalho, ou seja, foi o que apresentou qualidades estatísticas mais satisfatórias e o que melhor explicou o comportamento da variável Dormidas na Região Norte de Portugal, evidenciando ser adequado para a modelação e previsão da série de referência.

Como avaliação final, resta salientar que as autoras, no sentido de obter uma maior precisão do modelo, pretendem em investigações futuras introduzir no modelo variáveis explicativas e inerentes aos mercados emissores, tais como, preços de viagens, taxa de inflação, taxa de desemprego, entre outras.

#### REFERÊNCIAS BIBLIOGRÁFICAS

Chaves, Cristina; Maciel, Elisabete; Guimarães, Paulo e Ribeiro, José Carlos; (2000); “*Instrumentos estatísticos de apoio à economia: conceitos básicos*.” Lisboa: McGraw-Hill. ISBN 972-773-034-5.

Cunha, Licínio; (2003); “*Introdução ao turismo*.” Editorial Verbo, Lisboa/São Paulo.

Fernandes, Paula O.; (2005); “*Modelación, Predicción y Análisis del Comportamiento de la Demanda Turística en la Región Norte de Portugal*”; Dissertação de Doutoramento. Universidad de Valladolid, Espanha.

Fernandes, Paula O.; Teixeira, João P.; Ferreira, João M. and Azevedo, Susana G.; (2008); “Modelling Tourism Demand: A Comparative Study between Artificial Neural Networks and the Box-Jenkins Methodology”; *Romanian Journal of Economic Forecasting*; n.º 5(3), pp. 30/50.

Goh, Carey and Law, Rob; (2002); “Modeling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention”; *Tourism Management*; n.º 23, pp.499/510.

Gujarati, Damodar N.; (1995); “*Basic Econometrics*” (3ª edição). McGraw-Hill; ISBN 0-07-113964-8.

INE; (1990-2006); “*Anuários Estatísticos do Turismo de 1990-2006*”; Lisboa.

Liam, C.; (1997); “An econometric classification and review of international tourism demand models”; *Tourism Economics*, n.º 3, pp. 69/81.

MAKRIDAKIS, Spyros and HIBON, Michèle; (1997); “ARMA Models and the Box-Jenkins Methodology”; *Journal of Forecasting*; Vol. 16; pp.147/163.

Moutinho, Luiz; Huarng, K.H.; Yu, T.H.K. and Chen, C.Y; (2008); "Modeling and forecasting tourism demand: the case of flows from Mainland China to Taiwan"; *Service Business*; Vol. 2; n.º 3, pp. 219/232(14).

Preez, Johann and Witt, Stephen F.; (2003), "Univariate versus multivariate time series forecasting: an application to international tourism demand"; *International Journal of Forecasting*; n.º 19; pp. 435/451.

Thawornwong, S. and Enke, D.; (2004); "The adaptive selection of financial and economic variables for use with artificial neural networks"; *Neurocomputing*; n.º 6; pp.205/232.

Thomakos, Dimitrios D. and Guerard, John B.; (2004); "Naive, ARIMA, nonparametric, transfer function and VAR models: A comparison of forecasting performance"; *International Journal of Forecasting*; n.º 20; pp. 53/67.

Turismo de Portugal; (2009); "*Conta Satélite do Turismo - Turismo: Actividade Económica 2000 2007*"; <http://www.turismodeportugal.pt/>

Witt, Stephen F. and Witt, Christine A.; (1995); "Forecasting tourism demand: a review of empirical research"; *International Journal of Forecasting*; n.º 11; pp. 447/475.

Yu, Gongmei and Schwartz, Zvi; (2006); "Forecasting Short Time-Series Tourism Demand with Artificial Intelligence Models"; *Journal of Travel Research*; n.º 45, pp. 194/203.