



**GRADO EN ADMINISTRACION Y DIRECCION DE
EMPRESAS**

CURSO ACADEMICO
2018/2019

TRABAJO DE FIN DE GRADO

**ANALISIS DEL TURISMO CANTABRO. ENFOQUE
ECONOMETRICO.**

**ANALYSIS OF TOURISM IN CANTABRIA.
ECONOMETRIC APPROACH**

AUTOR: GONZALO GOMEZ RUIZ

**DIRECTORAS: CRISTINA RUIZ DEL RIO
PATRICIA MORENO MENCIA**

FECHA DE PRESENTACION: 3 DE JULIO DE 2019

INDICE

1	INTRODUCCION	4
2	APLICACION EMPIRICA.....	6
2.1	DATOS.....	6
2.2	VARIABLES	6
2.3	ANALISIS BIVARIANTE	7
2.4	CREACIÓN DEL MODELO.....	15
2.4.1	VALIDACION DEL MODELO (TUR).....	17
2.4.2	INTERPRETACION DE LOS PARAMETROS	22
3	PREDICCIONES	23
4	CONCLUSION	25
5	BIBLIOGRAFIA.....	26
6	ANEXO: FUENTE DE DATOS	27

RESUMEN

Con este trabajo, pretenderemos estudiar la demanda turística en Cantabria, tratando de entender mejor el comportamiento de los viajeros que deciden visitar la región. Para ello entenderemos la demanda como número de viajeros, y la intentaremos explicar a través de factores relevantes como el nivel de los precios en la región, la situación económica general del país y la oferta turística de la que ha dispuesto la región. Se estudiará tomando los datos de los últimos 14 años (2004-2018) de los factores que influyen en la demanda, que se explicarán más adelante. La variable dependiente será el número de viajeros total que recibe Cantabria a lo largo de un año, tanto en turismo hotelero, como extrahotelero. Tras elegir las variables que mejor expliquen la demanda turística en Cantabria, y calcular sus coeficientes a través del método MCO, se realizarán una serie de contrastes para ver la validez del modelo. El objetivo final es conseguir un modelo con un buen ajuste que permita predecir cual va a ser la demanda futura de turistas en la región, y, así, se puedan utilizar los recursos disponibles de la manera mas eficiente posible y sin realizar un exceso de gastos innecesario.

ABSTRACT

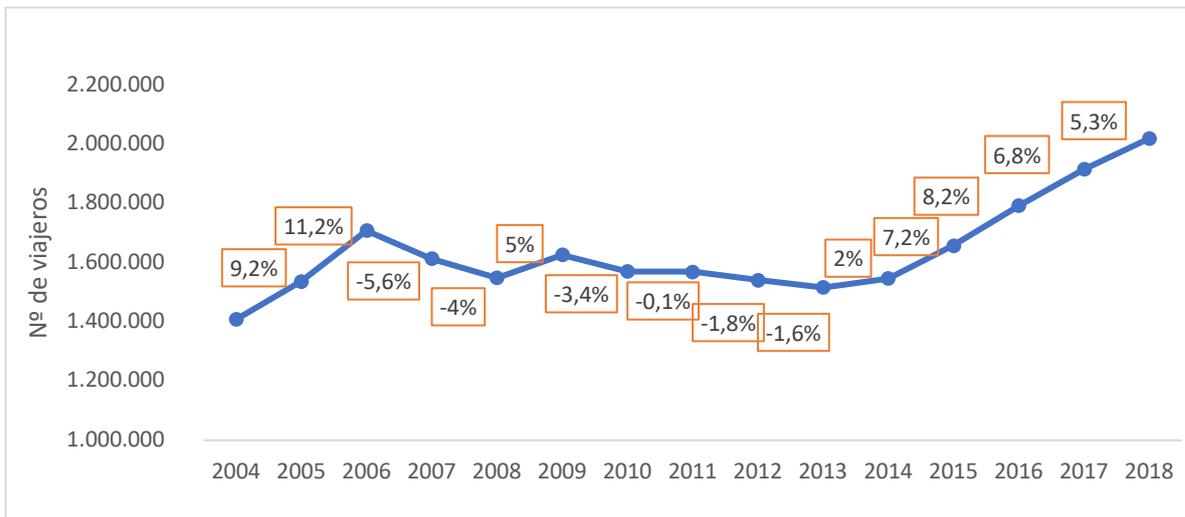
The aim of this project is to study the touristic demand in the community of Cantabria, by understanding the behaviour of the travellers that choose to visit this area. In order to do this, we will identify the demand with the number of travellers, and we will try to explain it through relevant factors such as the price levels of the area, the economic situation of the country and the touristic offer disposed by the region. These data of the factors that intervene in the demand will be analysed from the last 14 years (from 2004 to 2018). The dependent variable will be the total number of travellers that Cantabria receives throughout a year, in hotel tourism as well as extra hotel tourism. After having chosen the variables that explain best the touristic demand in Cantabria, and calculating their coefficients through the MCO method, a series of contrasts will be done in order to prove the validity of the model. The final objective is to reach a model with a good adjustment that enables us to predict what will be the future demand of tourists in the region, and, thus, the available resources will be used in the most efficient way and without making unnecessary and excessive expenses.

1 INTRODUCCION

El turismo es uno de los pilares de la economía en España, una de sus principales fuentes de ingresos, pues, según la Organización Mundial del Turismo (OTM), es el país que más turistas recibe al año, después de Francia y Estados Unidos. Por esto, este sector es uno de los motores para el crecimiento y desarrollo del país, ya que en 2018 aportó aproximadamente 178.000 millones de euros a la economía española, siendo un 14,6% del PIB español en ese año (EFE, 2019). Estos datos se deben a la amplia oferta turística que posee España, desde destinos de sol y playa, hasta una variedad de lugares con gran atractivo cultural, gastronómico, geográfico...

El objetivo de este estudio se centra en la región de Cantabria, que según el "Plan de Marketing turístico de Cantabria 2018-2019" (Oficina de turismo cantabro, 2019), en 2018 se consolidó como la región líder en el ámbito turístico de la zona norte de España, siendo un destino atractivo para todo tipo de turista, ofreciendo un turismo tanto de naturaleza, como cultural. La demanda de turismo en Cantabria creció en 2018, en un 5,3% en el número de viajeros y en un 5% en el número de pernoctaciones.

Ilustración 1.1 Crecimiento del número de viajeros en Cantabria



Fuente: Elaboración propia con datos del ICANE

Este crecimiento se logra, tanto en el mercado nacional, como en el internacional, pero este último logra un mayor dinamismo, pues en 2018, los turistas internacionales representaron un 20% de la demanda turística total, la mayoría procedentes de Francia y de Reino Unido. En total, en el último año, viajaron a Cantabria más de 2 millones de personas. Esto supuso para Cantabria un 10,9% del PIB de la región, por lo que su incidencia en la economía de la región es muy importante. Además, durante el año 2014, el turismo supuso un 11,4% del empleo regional (Exceltur, 2014).

Cabe destacar que en Cantabria el turismo está marcado por la estacionalidad, siendo julio y agosto los meses con más demanda turística, aunque en este estudio se tomarán datos anuales, por lo que no influirá en los resultados.

La mayoría de los turistas, entorno al 70%, eligen Cantabria como destino por temas de ocio y vacaciones (ICANE, 2017), pues cuenta con una amplia oferta de actividades y visitas culturales de gran interés, como el Parque Natural de Cabárceno, que fue visitado por casi 700.000 personas en 2017 (Cantur, 2017) o las Cuevas de Altamira, declaradas Patrimonio Mundial por la Unesco. También destaca el Año Jubilar Lebaniego, pues, en 2018, 1,2 millones de personas cruzaron la Puerta del Perdón del monasterio de Santo

Toribio y aportó a la economía española 132 millones de euros (Oficina de turismo cantabro, 2019).

A diferencia de otros sectores de la economía, la eficiencia y la rentabilidad del sector turístico no tiene tanto que ver con el aumento de la productividad o el mejor uso de las nuevas tecnologías, sino en intentar conocer la demanda futura para ofrecer una oferta que se ajuste a esa demanda, para no incurrir en gastos adicionales y repartir de una manera eficiente los recursos disponibles. Por ello, este trabajo se centra en intentar entender mejor esa demanda y conseguir crear un modelo econométrico que se ajuste a la realidad, y nos de la capacidad de poder hacer predicciones sobre el número de personas que visitan Cantabria, explicadas a través de una serie de variables que se consideran que influyen en la decisión de las personas para escoger Cantabria como destino turístico.

Antes de realizar el estudio, se han analizado otros trabajos acerca de la demanda turística (y de teoría econométrica), para entender como funciona y tener una idea de qué variables son las idóneas para crear un modelo que se ajuste a la realidad y sea capaz de realizar predicciones fiables sobre el número de turistas que se espera en la región.

Se han seguido los pasos de Wooldrige, J. en su libro "Introducción a la econometría. Un enfoque moderno", a la hora de realizar el análisis de los datos y la especificación del modelo a través de los contrastes. Este libro aplica problemas de econometría al mundo real para facilitar la comprensión de las técnicas econométricas existentes, que se han utilizado en el trabajo.

Uno de los primeros estudios econométricos que estudiaron la demanda española a través de técnicas econométricas fue "Introducción a un análisis econométrico del turismo" (Pulido San Roman, 1966), que tomó al turismo como un fenómeno económico y aplicó la econometría para crear el primer modelo capaz de estudiar y predecir el comportamiento del turismo en España.

Destaca también la tesis doctoral "Análisis de la demanda: aplicación a la actividad turística de las técnicas de predicción" (Esteban Talaya, 1987), que analizó los principales elementos y características de la demanda turística estableciendo una serie de hipótesis sobre su comportamiento, a través de factores socioeconómicos, demográficos, pictográficos y técnicos. Estudió aspectos, tanto cualitativos, como cuantitativos, para crear varios modelos capaces de predecir la demanda futura de turismo.

Otro trabajo, publicado por el Banco de España, fue "Un análisis econométrico de los ingresos por turismo en la economía española" (Espasa, et al., 1990), donde realizaron un acercamiento econométrico a partir de los ingresos nominales por motivos de turismo, donde se deflactaron esos ingresos utilizando un índice de precios de turismo. El modelo creado, reveló que "dichos ingresos dependen de los precios relativos de España respecto a los países de origen de los turistas, respecto de otros países oferentes de servicios turísticos, de la renta de los turistas, de una variable que recoja las mejoras en comercialización, infraestructura y calidad en la oferta turística española y de un elemento residual que recoja el efecto de "Otras Causas" no explicitadas en el modelo".

Alternativamente, el estudio "Un análisis econométrico del turismo hotelero y extrahotelero en las regiones y provincias españolas" (Guisan & Neira, 2001), presentó varios modelos econométricos para estudiar la evolución del turismo, tanto hotelero como extrahotelero a nivel regional y provincial en España, ya que consideraba que no había suficientes trabajos enfocados al turismo extrahotelero y lo consideraba igual de importante, o superior, en ciertas regiones de España. En estos modelos se utilizaron factores como el clima, la oferta del turismo hotelero o la demanda del turismo extrahotelero. Estos factores resultaron ser muy importantes, ya que encontró diferencias sustanciales en algunas regiones españolas condicionadas por el clima o la proximidad a núcleos urbanos de población grandes.

Otro trabajo que modelizó y analizó la evolución del turismo hotelero en las regiones y provincias españolas, utilizando las pernoctaciones hoteleras como indicador y sintetizando los factores que influyen en el desarrollo del turismo y la importancia que tiene este en el desarrollo económico a nivel regional, es "Evolución del turismo hotelero en las regiones y provincias españolas, 2001-2008" (Guisan & Aguayo, 2008).

2 APLICACION EMPIRICA

2.1 DATOS

Todos los datos utilizados para realizar el análisis han sido extraídos del Instituto Nacional de Estadística (INE), para los datos macroeconómicos que nos indican la situación económica de España (PIB y tasa de paro), y del Instituto Cántabro de Estadística (ICANE), para las variables referentes al nivel de precios en Cantabria, la oferta en el sector turístico de la región y el número de viajeros en Cantabria en los últimos años.

El análisis y procesamiento de los datos se realizarán con los programas Excel y Gretl.

2.2 VARIABLES

En cuanto a la variable endógena, he decidido escoger el número de viajeros, ya que puede ser un buen indicador de la demanda turística en Cantabria:

- *Número de viajeros (TUR)*: son el número total de personas que viajaron a Cantabria, independientemente de las pernoctaciones que realizaron.

Para calcular el total de viajeros se han sumado los datos de la coyuntura turística hotelera y la extrahotelera (campings, apartamentos turísticos y alojamientos en turismo rural) para así obtener el número total de viajeros.

El modelo se creará a partir de tres tipos de variables explicativas: una referente al nivel de los precios en Cantabria, una que refleje la situación económica del país al que pertenece y otra para relacionarlo con la oferta del sector turístico en la región (sumados los datos referentes a hoteles, campings, apartamentos turísticos y alojamientos en turismo rural):

- *Índice de precios al consumidor (IPC)*: es una medida estadística de la evolución de los precios de los bienes y servicios que consume la población residente en viviendas familiares en Cantabria.
- *Índice de precios del grupo hoteles y restaurantes (IPHR)*: es un índice de medida que se encuentra dentro del IPC, es uno de los numerosos grupos que influye en el valor de este.
- *Índice de precios hoteleros (IPH)*: es un índice que mide la evolución temporal de los precios recibidos por los empresarios hoteleros, de Cantabria, considerando todos sus clientes (hogares, empresas, touroperadores y agencias de viajes). Como se tiene en cuenta también los alojamientos extrahoteleros, se tomará este índice para ver su nivel de precios, ya que se supone que variará de una forma parecida y son similares independientemente del tipo de establecimiento.

- *Producto interior bruto (PIB)*: indicador económico que refleja el valor monetario de todos los bienes y servicios finales producidos por un país o región, en este caso, de España, en un determinado periodo de tiempo, normalmente un año.
- *Tasa de paro (PAR)*: mide el nivel de desocupación en relación con la población activa, en España.
- *Personal empleado (EMP)*: se define como el conjunto de personas, remuneradas y no remuneradas, que contribuyen mediante la aportación de su trabajo, a la producción de bienes y servicios en el establecimiento durante el mes que incluye el periodo de referencia de la encuesta, aunque trabajen fuera de los locales de este.
- *Nº de establecimientos abiertos (NEA)*: número de establecimientos hoteleros abiertos de temporada estimados por la encuesta.
- *Nº de plazas estimadas (NPE)*: número de plazas estimadas por la encuesta de los establecimientos hoteleros abiertos de temporada. El número de plazas equivale al número de camas fijas del establecimiento. No se incluyen, por tanto, las supletorias, y las de matrimonio dan lugar a dos plazas.

2.3 ANALISIS BIVARIANTE

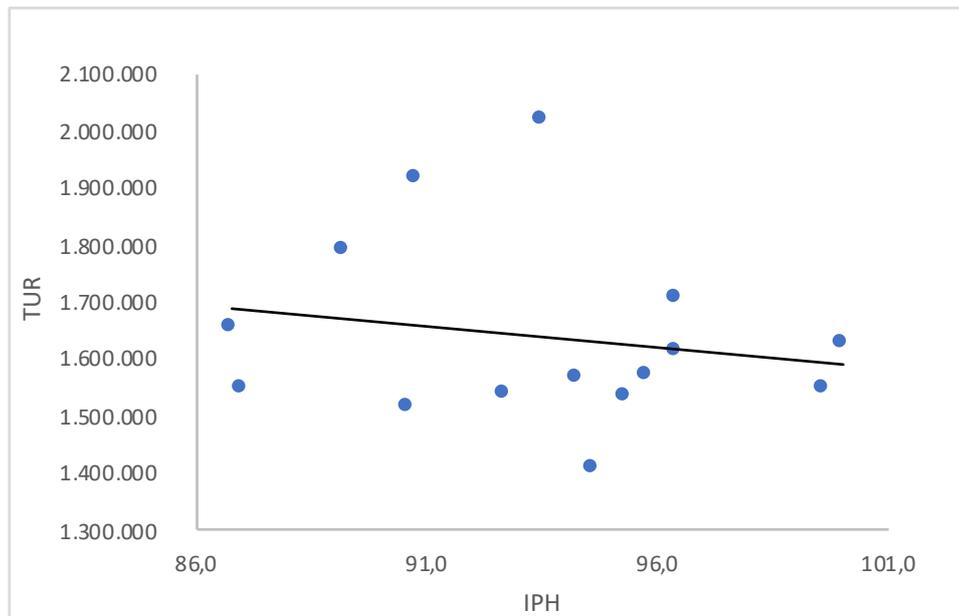
A continuación, se realizará un análisis de cada variable explicativa (X) con la variable dependiente (Y) de manera conjunta, para ver la relación de las variables endógenas con las variables explicativas de forma individual. Se hará a través del gráfico X-Y, donde las variables endógenas permanecerán en el eje de abscisas, mientras que el eje horizontal ira variando según la variable X que se este analizando. Luego se analizará la ecuación de la recta de la regresión, para ver el cambio que provoca cada variable exógena en la endógena. Por último, se creará un modelo lineal simple por el método de MCO, mediante gretl, para ver la bondad del ajuste y la significación de cada una de las variables exógenas con la variable endógena.

TUR vs IPH

$$\text{TUR} = 2314452,9 - 7226,51 \cdot \text{IPH}$$

$$R^2 = 0,0327$$

Ilustración 2.1 Gráfico X-Y para IPH



Fuente: Elaboración propia con datos del ICANE

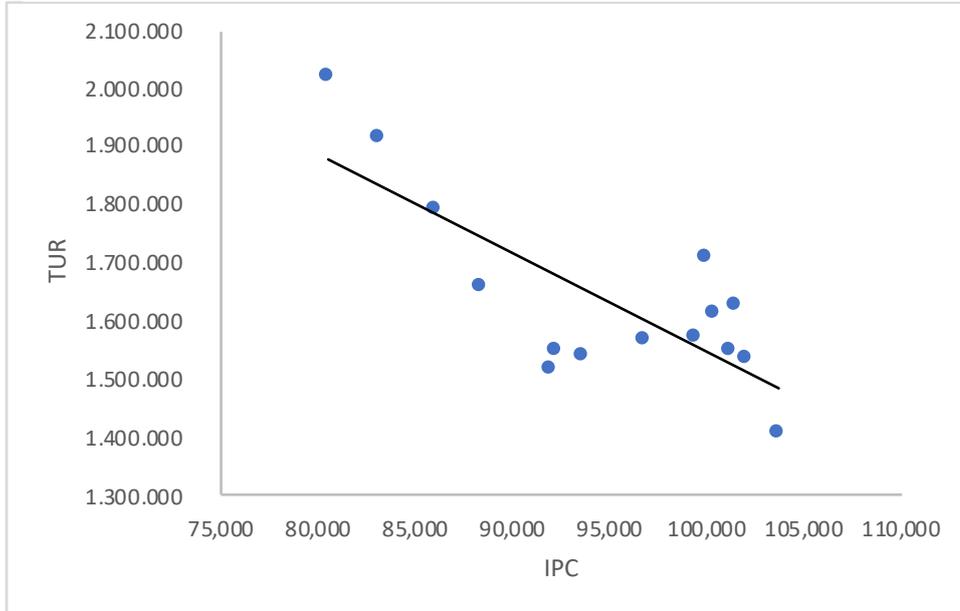
La ecuación de la recta nos muestra la variación de la variable dependiente, cuando la variable independiente varía en una unidad. En este caso, cuando hay un cambio de un 1% en el índice de precios hoteleros, el número de turistas disminuye en 7226,51. El coeficiente de determinación (R^2) nos indica la bondad del ajuste, que en este caso es de un 3,27%, por lo que esta variable solo explica un 3,27% de la variable TUR cuando se analizan individualmente. La relación entre TUR e IPH es inversa, como es lógico, ya que, a mayor índice de precios, el poder adquisitivo de las personas disminuye, y por tanto también el consumo de cualquier tipo de bien, en este caso del turismo. Al analizar la significación de esta variable en un modelo simple, se puede concluir que la variable IPH no es significativa respecto a TUR, ya que se obtiene un p-valor de 0,5193.

TUR vs IPC

$$\text{TUR} = 3245477,66 - 16963 \cdot \text{IPC}$$

$$R^2 = 0,6081$$

Ilustración 2.2 Gráfico X-Y para IPC



Fuente: Elaboración propia con datos del ICANE

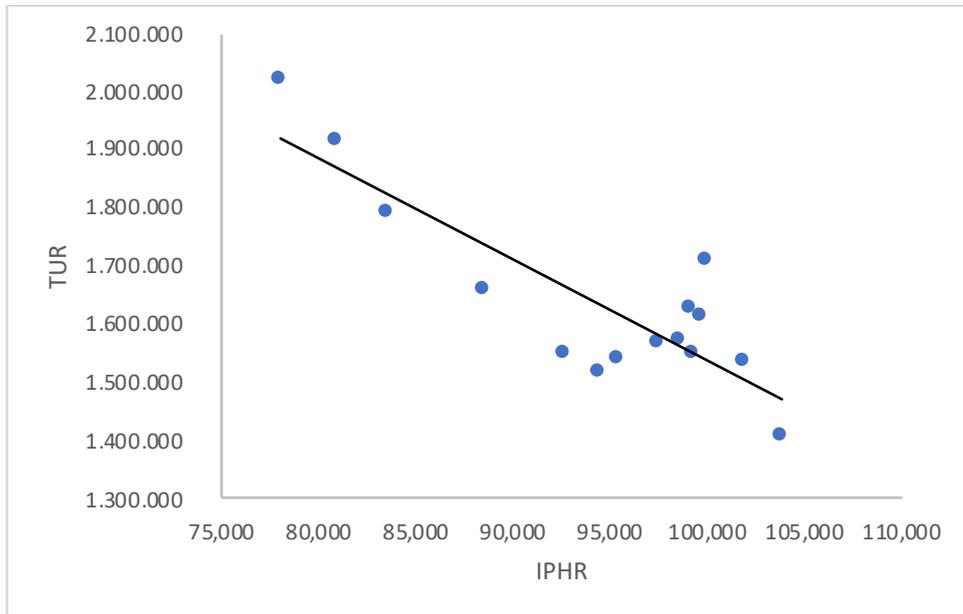
Para un cambio de un 1% en el IPC general, el número de turistas disminuye en 16963. El coeficiente de determinación indica que la bondad del ajuste es del 60,81%, un buen ajuste al tratarse de un modelo con una sola variable, que más adelante se hará conjunto a las otras variables. La relación de ambas variables es negativa, ya que, como en el caso anterior, un mayor índice de precios provoca una disminución del poder adquisitivo de la gente. La variable IPC tiene un p-valor de 0,0006, por lo que se puede afirmar que si es significativa en un modelo simple.

TUR vs IPHR

$$TUR = 3278694,97 - 17406,22 * IPHR$$

$$R^2 = 0,7303$$

Ilustración 2.3 Gráfico X-Y para IPHR



Fuente: Elaboración propia con datos del ICANE

La ecuación de la recta nos indica que a un aumento de 1% del nivel de precios de hoteles y restaurantes, el número de turistas disminuye en 17406,22. La bondad del ajuste es buena, la variable IPHR es capaz de explicar, en un modelo simple, un 73,03% de la variable dependiente. Como en los dos casos anteriores, la relación es inversa al tratarse de un índice de precios, que hace disminuir la capacidad adquisitiva de la población. Además, es significativa al obtenerse un p-valor de $4,96 \times 10^{-6}$.

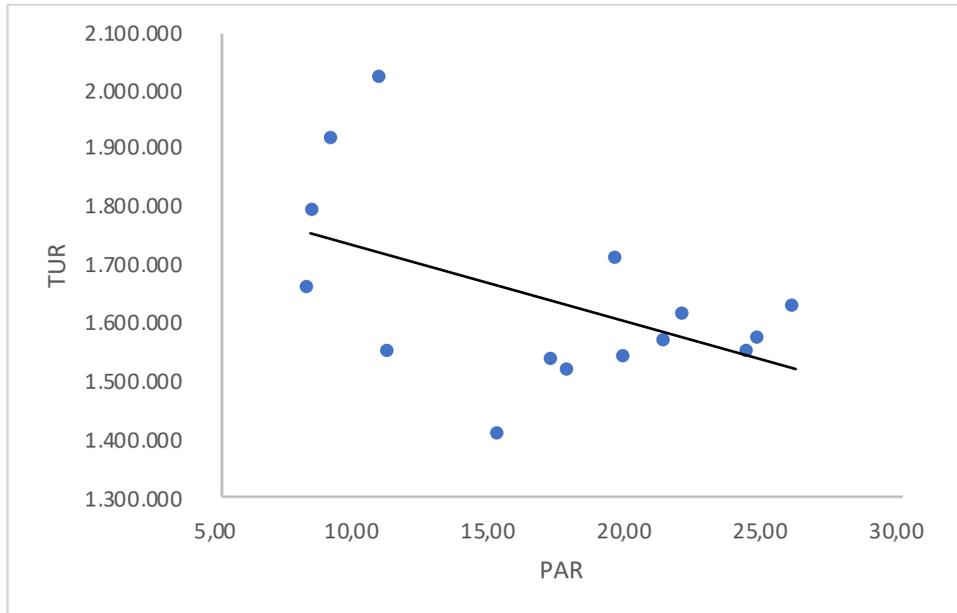
Después de observar los tres índices de precio, podemos afirmar que el que mayor coeficiente de determinación tiene, y, por lo tanto, mejor explica la variable TUR (de forma individual) es IPHR.

TUR vs PAR

$$\text{TUR} = 1858521,67 - 12857,39 \cdot \text{PAR}$$

$$R^2 = 0,2457$$

Ilustración 2.4 Gráfico X-Y para PAR



Fuente: Elaboración propia con datos del INE

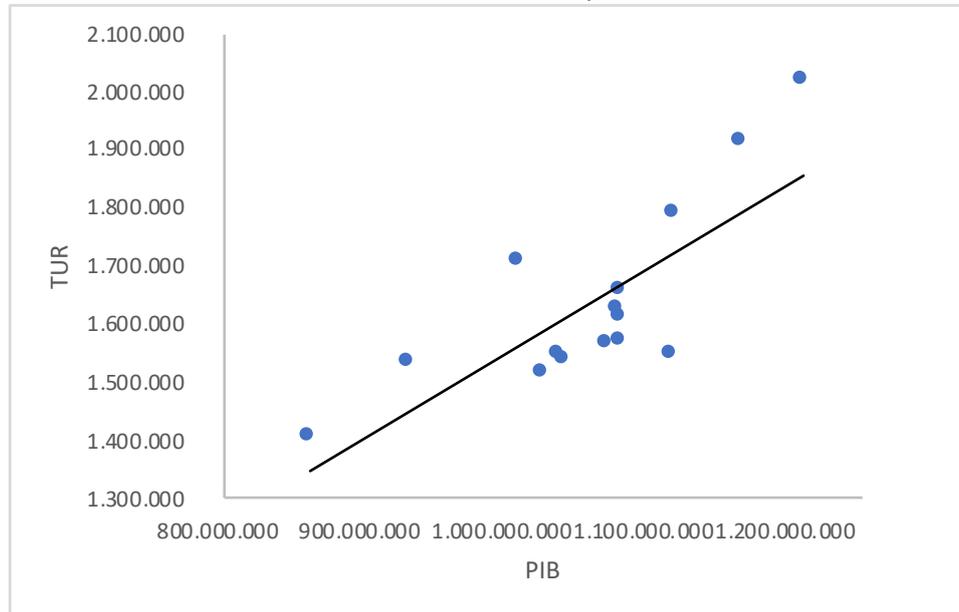
La recta de la regresión indica que a un aumento de un 1% en la tasa de paro de España, el turismo en Cantabria disminuye en 12857,39 turistas. La variable PAR es capaz de explicar en un 24,57% la variable dependiente. La relación entre ambas variables es negativa, pues una mayor tasa de paro querrá decir que la situación económica del país es peor, por lo que se realizarán menos viajes. Al crear un modelo simple con estas variables, se obtiene un p-valor de 0,06, por lo que la variable PAR no es significativa al 5% de significatividad.

TUR vs PIB

$$TUR = 79671 + 0,0015 * PIB$$

$$R^2 = 0,6068$$

Ilustración 2.5 Gráfico X-Y para PIB



Fuente: Elaboración propia con datos del INE

El número de turistas aumenta en 0,0015 cuando el PIB de España aumenta en un euro. Esta variación se ve muy pequeña, debido a lo grande que es el PIB en relación con el número de turistas. Si el PIB aumentase, por ejemplo, en 10 mil euros, el número de turistas aumenta en 15. Este aumento puede explicarse debido a que el PIB es un indicador económico de la situación económica actual, por lo que es normal que a mayor PIB se produzcan un mayor número de viajes entre la población. El PIB es capaz de explicar un 60,68% del número de turistas. Además, es significativa al obtener un p-valor de 0,0006 al crear un modelo.

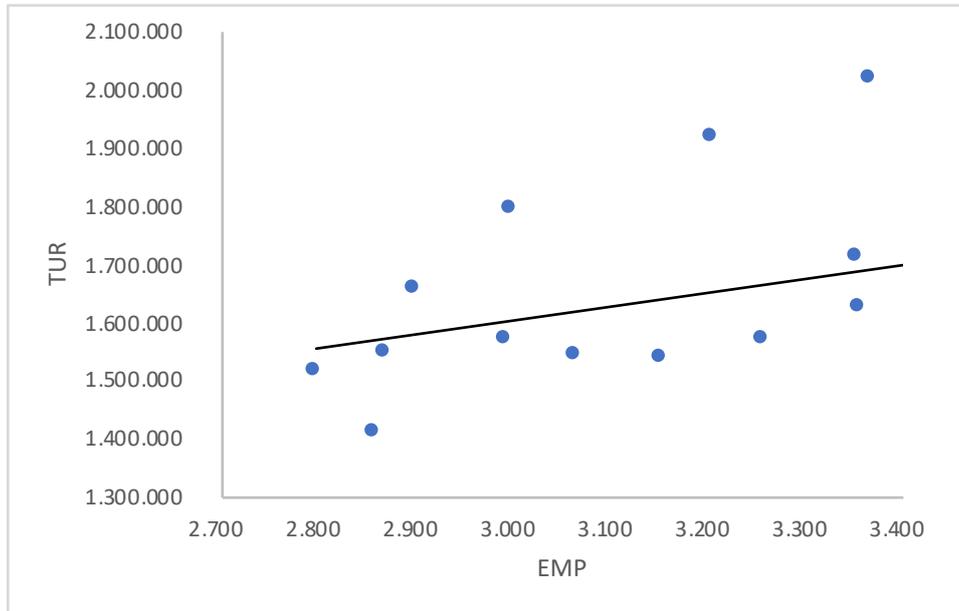
Entre las dos variables que representan la situación económica del país, la que mas explica la variable TUR, de forma individual, es PIB, con un coeficiente de determinación de 0,6068.

TUR vs EMP

$$TUR = 891122 + 237,66 * EMP$$

$$R^2 = 0,1239$$

Ilustración 2.6 Gráfico X-Y para EMP



Fuente: Elaboración propia con datos del ICANE

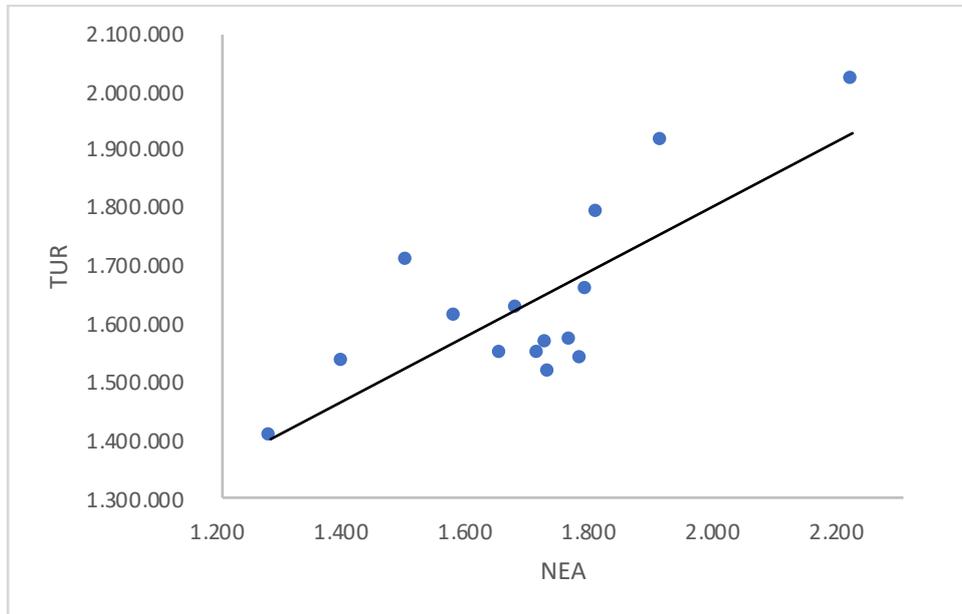
Ante cambios de un personal empleado adicional, el número de turistas aumenta en 237,66, pues la oferta turística sería mayor, y, por lo tanto, también la demanda. EMP es capaz de explicar en un 12,39% la variable endógena. En el modelo lineal simple con la variable TUR, nos encontramos con que no es significativa, al tener un p-valor (0,1991) mayor que 0,05.

TUR vs NEA

$$TUR = 685015,25 + 561,1 * NEA$$

$$R^2 = 0,579$$

Ilustración 2.7 Gráfico X-Y para NEA



Fuente: Elaboración propia con datos del ICANE

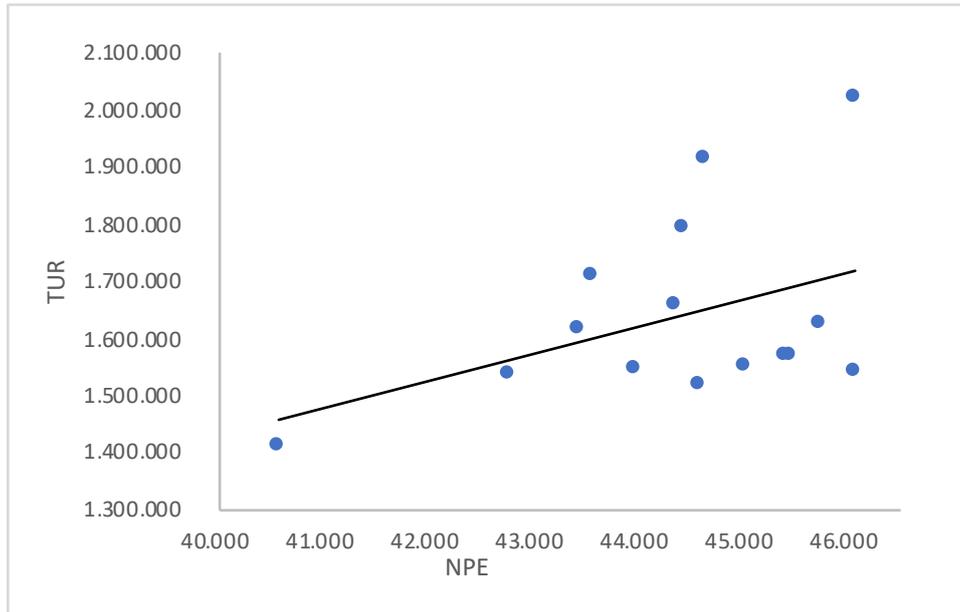
Un cambio de una unidad en la variable número de establecimientos abiertos provoca un aumento de 561,1 del número de turistas. Cuanto mayor sea la oferta de establecimientos abiertos, mayor será la demanda que puede ser satisfecha. La bondad del ajuste es del 57,9%. Esta variable si es significativa al 5% de significatividad, pues su p-valor es de 0,0010.

TUR vs NPE

$$\text{TUR} = 441699,45 + 46,86 * \text{NPE}$$

$$R^2 = 0,1751$$

Ilustración 2.8 Gráfico X-Y para NPE



Fuente: Elaboración propia con datos del ICANE

Por último, la relación entre las variables TUR y NPE es directa, a cambios de una unidad en el número de plazas estimadas, el número de turistas crece en 46,48, por los mismos motivos que las dos variables anteriores (a más oferta, más demanda). La variable NPE explica en un 17,51% la variable TUR, y no es significativa al realizar su modelo, pues su p-valor es igual a 0,1206.

Viendo los coeficientes de determinación de las variables que indican el nivel de oferta, la que mejor explica la variable endógena TUR es NEA (número de establecimientos abiertos), con un coeficiente de determinación de 0,579. Además, es la única que es significativa de forma individual.

2.4 CREACIÓN DEL MODELO

“El análisis de regresión múltiple es adecuado para realizar un análisis ceteris paribus debido a que permite controlar de manera explícita muchos otros factores que afectan en forma simultánea a la variable dependiente” (Wooldrige). Al aceptar más variables explicativas, que pueden estar correlacionadas, puede inferirse causalidad en casos en los que el modelo de regresión lineal simple podría no dar los resultados más óptimos. El modelo de regresión múltiple puede expresarse como:

$$Y = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \beta_3 * X_3 + \dots + \beta_k * X_k + \varepsilon$$

Donde β son los parámetros asociados a los términos independientes (X), β_0 es el término constante y la variable ε es el término de error o perturbación, que contiene los factores de los que depende Y que no han sido introducidos en el modelo. Los parámetros (β) se estiman a través del método de cuadrados ordinarios (MCO), que elige las estimaciones que minimizan la suma de los residuales cuadrados. Una vez

obtenidos estos parámetros se realizará una interpretación ceteris paribus de cada una de las estimaciones.

También se medirá la bondad del ajuste, que se define como:

$$R^2 = SEC/SCT$$

Donde SEC es la suma explicada de cuadrados y SCT es la suma total de cuadrados. Este término nos dirá cuanto son capaces de explicar las variables explicativas (X) a la variable dependiente (Y).

Una vez elegido el modelo y calculadas las estimaciones de los parámetros se pasará a validar el modelo, mirando si cumple algunos de los supuestos del MRLM:

- No hay colinealidad perfecta, es decir, ninguna de las variables independientes es constante y no hay ninguna relación lineal exacta entre las variables independientes
- Existe homocedasticidad, dado cualquier valor de las variables explicativas, el error tiene la misma varianza.
- Normalidad de los residuos. El error poblacional es independiente de las variables explicativas y esta distribuido normalmente, con media cero y varianza σ^2 .
- No existencia de autocorrelación, es decir, los residuos se generan de forma aleatoria y no de forma sistemática.

A continuación, se utilizará la matriz de correlaciones para ver qué relación tienen cada una de las variables que se pretenden introducir en el modelo. Nos fijaremos en las variables que mas se aproximen a 1, ya que serán las que mas correlacionadas estén con la variable endógena. Se tendrán en cuenta también las variables explicativas que estén muy correlacionadas entre sí, para que no haya problemas de multicolinealidad, y así buscar el modelo econométrico mas óptimo.

Ilustración 2.9 Matriz de correlaciones

	TUR	IPH	IPC	IPHR	PIB	PAR	EMP	NEA	NPE
TUR	1								
IPH	-0,1807	1							
IPC	-0,7798	0,6819	1						
IPHR	-0,8546	0,5823	0,9820	1					
PIB	0,7790	-0,0909	-0,6821	-0,7502	1				
PAR	-0,4957	0,8398	0,7694	0,7427	-0,1728	1			
EMP	0,3519	0,7044	0,1834	0,0701	0,4145	0,4478	1		
NEA	0,7609	-0,3264	-0,8472	-0,8595	0,8899	-0,3499	0,1003	1	
NPE	0,4185	0,0453	-0,4257	-0,4343	0,7717	0,1912	0,2683	0,7879	1

Fuente: Elaboración propia con datos del INE e ICANE

Observando la matriz de correlaciones (ilustración 3.9), se ve como la variable con mayor correlación con la endógena es IPHR (-0,8546), por lo que será la primera variable que se introducirá en el modelo. Se descartarán las otras variables que nos mostraban los índices de precios (IPC y IPH). Como se puede ver en la matriz, la variable IPHR tiene una fuerte correlación con las variables PIB, PAR y NEA. Al introducirla en el modelo sólo con la variable endógena, obtenemos un R^2 ajustado de 0,7303.

A continuación, pasaremos a escoger la variable indicativa de la situación económica del país. Al introducirlas en el modelo se obtiene un mayor R^2 con la variable PIB que con la variable PAR (0,7738 y 0,7734, respectivamente), por lo que escogemos la variable PIB para nuestro modelo. Aquí pueden aparecer problemas de multicolinealidad, ya que, tanto IPHR como PIB tienen una fuerte correlación con NEA. También la variable PIB tiene una fuerte correlación con NPE.

Por último, elegiremos la variable referente a la oferta turística en Cantabria. Se confirman los problemas de multicolinealidad mencionados anteriormente, ya que los coeficientes de determinación ajustado de modelos con las variables NEA (0,7948) y NPE (0,7981) son menores que el que conseguimos con la variable EMP (0,9148), a pesar de ser la que menos correlación tenga con la variable endógena. Por ello, escogemos la variable EMP para introducir en el modelo.

Finalmente, el modelo resultante a estudiar es:

$$TUR = \beta_0 + \beta_1 * IPHR + \beta_2 * PIB + \beta_3 * EMP + \varepsilon$$

Ilustración 2.10 Tabla coeficientes MRLM

Variable	Coefficiente (β)	Desv. típica	Estadístico t	P-valor
Constante	3093860	531964	5,816	0,0001
IPHR	- 22057,9	3503,71	-6,296	$5,87 \times 10^{-5}$
PIB	0,000479659	0,0003557	-1.348	0,2047
EMP	359,868	84,3437	4,267	0,0013

R-cuadrado	0,9148
F	39,372
Durbin-Watson	1,9892

Fuente: Elaboración propia a través de gretl

Las tablas (Ilustración 3.10) muestran los resultados obtenidos al introducir el modelo en gretl y estimar los coeficientes del modelo econométrico mediante el método MCO. También se muestran algunos estadísticos que nos serán de utilidad para el trabajo. Finalmente se muestra el modelo de regresión múltiple con los coeficientes de las distintas β estimados:

$$TUR = 3093860 - 22057,9 * IPHR - 0,000479659 * PIB + 359,868 * EMP$$

Como indica el R-cuadrado las variables X elegidas explican la variable Y en un 91,48%, por lo que el ajuste de el modelo es elevado.

2.4.1 VALIDACION DEL MODELO (TUR)

Una vez obtenido el primer modelo, se analizará para ver si cumple los supuestos clásicos, antes de interpretar el coeficiente estimado de los parámetros.

2.4.1.1 *Signos de los parámetros*

El modelo resultante indica que el IPHR tiene una relación negativa respecto a la entrada de turistas, como era de esperar, ya que se cumple que, un mayor nivel de precios hace que la demanda disminuya, ya que se reduce el poder adquisitivo de las personas.

En cuanto a la variable PIB, vemos como también influye de forma negativa en la variable dependiente, sin embargo, el signo esperado de esta variable era positivo, y de forma individual obteníamos esta positividad, ya que una mejor situación del país debería indicar una mayor entrada de turistas en la región, pues supone un mayor nivel de vida y una mayor renta per cápita. Esto se puede deber a la correlación con los otros regresores, ya que la variable no es significativa de forma individual, pero si cuando la variable dependiente se encuentra explicada por todos los regresores en conjunto.

La variable EMP tiene un coeficiente positivo, por lo que se ajusta también al signo esperado para esta variable, pues, ya que representa la oferta, al aumentar, crecerá el número de turistas debido a que puede satisfacerse una cantidad mayor de demanda.

2.4.1.2 *Contraste de significación individual*

Primero, se analizará cada variable explicativa individualmente para ver si, una vez realizado el modelo y estimados los coeficientes, son significativas en nuestro modelo. Se calculará el p-valor y se comparará con el nivel de significación para ver si cada de una de las variables es significativa (Wooldrige, p. 133). Este contraste se hará a un nivel de significación (α) del 5%:

Constante

$$\left\{ \begin{array}{ll} H_0: \beta_0=0 & \text{p-valor}=0,0001 \\ H_1: \beta_0 \neq 0 & \text{p-valor} < \alpha; 0,001 < 0,05 \end{array} \right.$$

Como el p-valor es menor que 0,05, rechazamos la hipótesis nula, por lo que podemos afirmar que la constante es significativa a un nivel de significación del 5%.

IPHR

$$\left\{ \begin{array}{ll} H_0: \beta_1=0 & \text{p-valor}=5,87 \cdot 10^{-5} \\ H_1: \beta_1 \neq 0 & \text{p-valor} < \alpha; 5,87 \cdot 10^{-5} < 0,05 \end{array} \right.$$

En este caso también se rechaza la hipótesis nula (p-valor < 0,05), así que la variable IPHR es significativa al 5% de significatividad.

PIB

$$\left\{ \begin{array}{l} H_0: \beta_2=0 \\ H_1: \beta_2 \neq 0 \end{array} \right. \quad \begin{array}{l} \text{p-valor}=0,2047 \\ \text{p-valor}>\alpha; 0,2047>0,05 \end{array}$$

Como el p-valor es mayor que 0,05, debemos aceptar la hipótesis nula, por lo que la variable PIB no es significativa en nuestro modelo. Mas adelante, en la prueba de significación global, veremos si son significativas en el modelo, aunque no lo sea de manera individual, ya que al el R^2 del modelo aumenta al introducir esta variable. En el análisis bivalente pudimos ver también como esta variable era significativa si se analizaba solo con TUR.

EMP

$$\left\{ \begin{array}{l} H_0: \beta_2=0 \\ H_1: \beta_2 \neq 0 \end{array} \right. \quad \begin{array}{l} \text{p-valor}=0,0013 \\ \text{p-valor}<\alpha; 0,0013<0,05 \end{array}$$

Rechazamos la hipótesis nula (p-valor<0,05), entonces la variable EMP es significativa al 5% de significatividad.

2.4.1.3 *Contraste de significación global*

Para este contraste usaremos el estadístico F, que se obtiene con la siguiente fórmula:

$$F = \frac{\frac{R^2}{k-1}}{\frac{1-R^2}{n-k}}$$

Para nuestro modelo tenemos una F de 39,37, que compararemos con el valor crítico tabulado en $F_{3;11;0,05}$, que es igual a 3,58743.

$$\left\{ \begin{array}{l} H_0: \beta_0=\beta_1=\beta_2=\beta_3=0 \\ H_1: \beta_0 \neq \beta_1 \neq \beta_2 \neq \beta_3 \neq 0 \end{array} \right.$$

La F es mayor que $F_{3;11;0,05}$ ($39,37>3,58743$), por lo que rechazamos la hipótesis nula de que todas las betas son 0 y podemos asegurar que el modelo esta bien explicado por

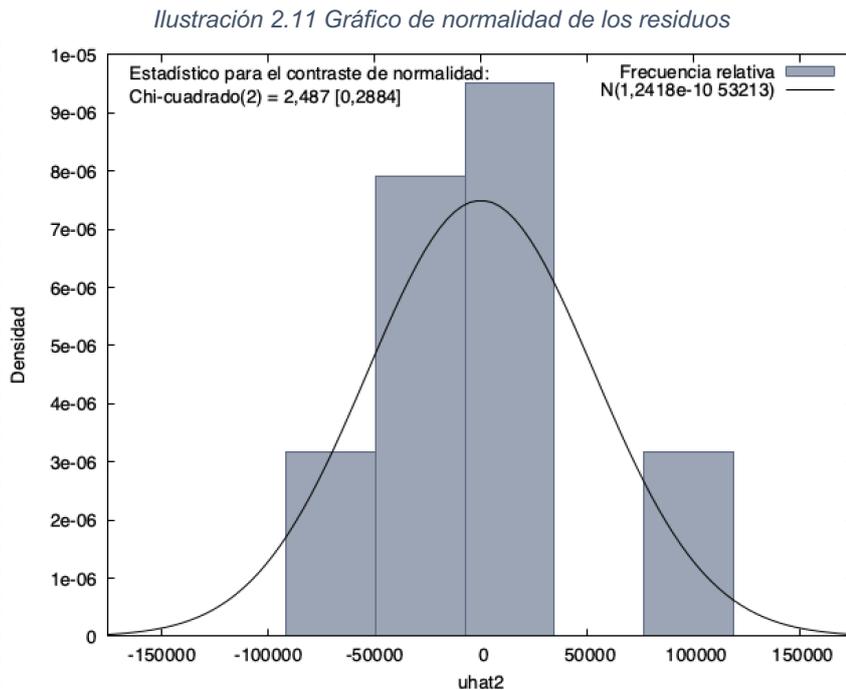
las variables independientes, transmiten información al modelo. Como se vió anteriormente, la variable PIB no era significativa de forma individual, pero al tener una significatividad conjunta de todos los parámetros, se sigue incluyendo en el modelo.

2.4.1.4 *Contraste normalidad de los residuos*

- H_0 : El error tiene una distribución normal
- H_1 : No existe normalidad en el error

Este contraste se realizará a través del p-valor, que se consigue mediante gretl. Se obtiene un p-valor de 0,288443, por lo que aceptamos la hipótesis nula ($0,288443 > 0,05$) y existe normalidad en los residuos.

Como se puede observar en el grafico (Ilustración 3.11) de distribución de las perturbaciones se puede ver este supuesto de normalidad, pues los residuos se distribuyen cerca de la línea y se ajustan bastante bien a la campana de Gauss.



2.4.1.5 *Contraste de heterocedasticidad de White*

- H_0 : Hay homocedasticidad
- H_1 : Hay heterocedasticidad

Realizando el contraste de White, se obtiene un p-valor de 0,4291, aceptando la hipótesis nula, entonces hay homocedasticidad, por lo que las perturbaciones no sufren variaciones sistemáticas cuando varían los valores de las variables.

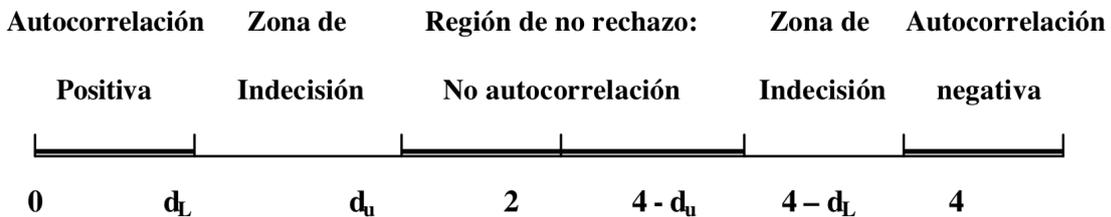
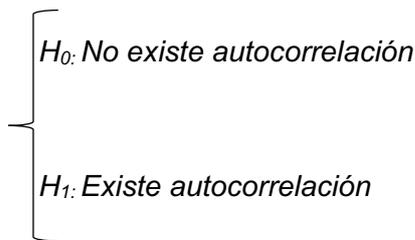
2.4.1.6 *Contraste de autocorrelación*

Este contraste de hipótesis es especialmente importante en estudios realizados con datos temporales. Con él se pretende verificar que los errores de las observaciones estimadas no están correlacionados.

Para verificar la ausencia de autocorrelación en el modelo, utilizaremos el test de Durbin-Watson, que propone el estadístico dado por:

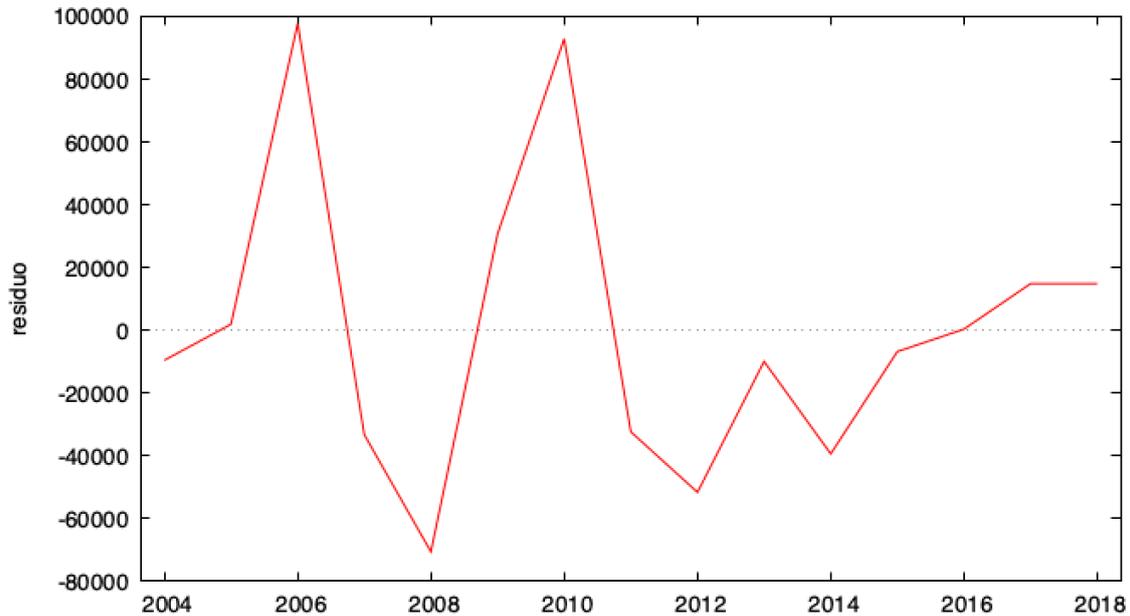
$$d = \frac{\sum_{i=2}^{i=n} (e_i - e_{i-1})^2}{\sum_{i=1}^{i=n} e_i^2}$$

Este estadístico se comparará con los límites superior e inferior de la tabla Savin y White. Posteriormente se mirarán esos límites dentro de las siguientes regiones de contraste para ver si rechazamos o no la hipótesis nula.



Con gretl se han calculado los límites superior (1,75) e inferior (0,814). Nuestra d era igual a 1,989, por lo que se encuentra dentro de la región de no rechazar la hipótesis nula. Podemos afirmar que no existe autocorrelación en nuestro modelo. A demás, fijándose en el p-valor que aporta gretl (0,9989) al realizar el contraste, vemos como es mayor que 0,05, por lo que se confirma que no existe autocorrelación.

Ilustración 2.12 Grafico de residuos contra el tiempo



Fuente: Elaboración propia a través de gretl

El gráfico (Ilustración 3.12) muestra la distribución de los residuos de la estimación realizada por MCO. Se puede ver como los residuos se generan de forma aleatoria, sin seguir ningún patrón, viendo gráficamente la no existencia de autocorrelación.

2.4.2 INTERPRETACION DE LOS PARAMETROS

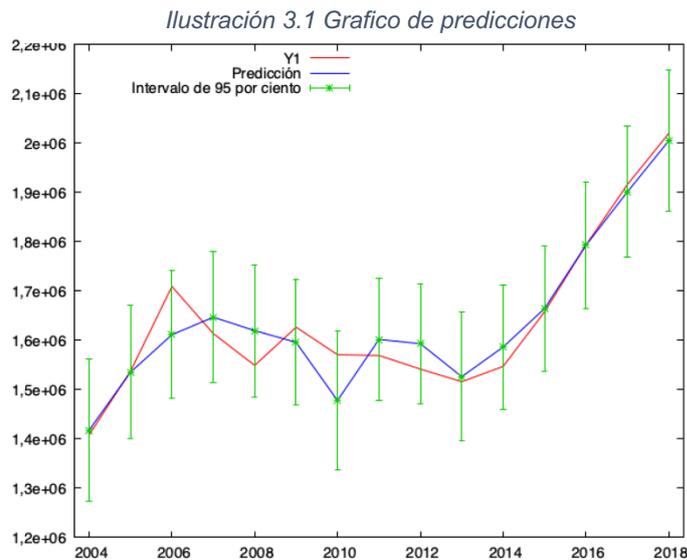
Vemos como el modelo con TUR cumple todas las hipótesis clásicas, por lo que se pasará a analizar los coeficientes obtenidos a través de MCO:

- β_0 : este es el valor de la constante, que es igual a 3093860, por lo que es el valor que toma la variable dependiente TUR cuando todas las variables independientes son iguales a 0.
- β_1 : obtenemos un valor de -22057,9. Esto nos indica que a variaciones de un 1% del IPHR, el número de turistas disminuye en esa cantidad, "ceteris paribus".
- β_2 : para la variable PIB, se puede ver como un cambio de una unidad en el PIB de España, hace disminuir la entrada de turistas en -0,000479659 unidades, "ceteris paribus".
- β_3 : tiene un coeficiente de 359,868, por lo que este es el valor que aumenta la variable TUR cuando aumenta en una unidad el empleo ofertado en el sector turístico, "ceteris paribus".

3 PREDICCIONES

Una vez estimado y estudiado el modelo, se realizarán unas predicciones para ver si las variables independientes explican de forma correcta la variable dependiente y saber si el modelo definitivo es "bueno".

Se verá si los datos reales con los que se ha trabajado se acercan a las predicciones de número de turistas que obtendríamos a través del modelo. Las predicciones se conseguirán a través de gretl. Como se puede observar en la tabla (ilustración 4.2), gretl calcula los intervalos de confianza de los valores que puede tomar la variable dependiente. Todos los valores predichos se encuentran dentro de los intervalos, por lo que el modelo tiene una buena capacidad de predicción.



Fuente: Elaboración a través de gretl

En el gráfico (ilustración 4.2) se observa como las predicciones se ajustan de una forma correcta a la realidad, pues, tanto la predicción, como los valores reales de TUR, siguen la misma tendencia y todas las predicciones se encuentran dentro de los intervalos de confianza.

Ilustración 3.2 Tabla predicciones del modelo

Año	TUR	Predicción	Desv. típica	Intervalo de 95%
2004	1408420	1417930,00	65279,10	(1,27425e+06, 1,56161e+06)
2005	1537420	1535470,00	61409,40	(1,40031e+06, 1,67063e+06)
2006	1709340	1611590,00	58771,10	(1,48223e+06, 1,74094e+06)
2007	1613870	1647120,00	60012,30	(1,51503e+06, 1,77920e+06)
2008	1549260	1619810,00	60832,90	(1,48592e+06, 1,75370e+06)
2009	1626820	1596110,00	57654,80	(1,46921e+06, 1,72301e+06)
2010	1570950	1478080,00	63587,60	(1,33813e+06, 1,61804e+06)
2011	1569490	1601860,00	56030,10	(1,47854e+06, 1,72518e+06)
2012	1541760	1593370,00	55232,60	(1,47181e+06, 1,71494e+06)
2013	1516360	1526310,00	59398,80	(1,39557e+06, 1,65704e+06)
2014	1547220	1586570,00	57466,50	(1,46008e+06, 1,71305e+06)
2015	1657990	1664770,00	57838,30	(1,53747e+06, 1,79207e+06)
2016	1793210	1792840,00	58725,00	(1,66358e+06, 1,92209e+06)
2017	1916000	1901130,00	60359,80	(1,76828e+06, 2,03398e+06)
2018	2019900	2005060,00	64707,10	(1,86264e+06, 2,14748e+06)

Fuente: Elaboración propia a través de gretl

A continuación, se realizará una predicción con datos reales de abril de 2019, por lo que se suponen los siguientes datos:

- IPHR: 105,74
- PIB: 307.640.000
- EMP: 3.602

Año	Predicción	Desv. típica	Intervalo de 95%
Abril 2019	1.910.208,85	268283,403	(1.319.721,06 , 2.500.696,64)

Se obtiene una predicción de número de viajeros de 1910208,85, mientras que la predicción por intervalo con un nivel de confianza del 95% para esta observación es 1319721,06 - 2500696,64. En muestreos repetidos o en el largo plazo, en 95 de cada 100 intervalos como el anterior estará incluido el verdadero valor individual del número de turistas.

4 CONCLUSION

El objetivo de este trabajo era conseguir realizar un modelo econométrico que explicase la demanda turística (medida a través del número de viajeros) en Cantabria, con el fin de poder prever la demanda de la región y así poder ajustar la oferta turística, a esa entrada de viajeros, de una forma más eficaz.

Finalmente, se ha conseguido un modelo donde la variable número de turistas (TUR) se ha explicado a través de las variables: índice de precios de hoteles y restaurantes (IPHR), producto interior bruto de España (PIB) y cantidad de personal empleado en el sector turístico en Cantabria. El ajuste del modelo ha sido bueno (91,48%) y las predicciones realizadas han sido bastante precisas. El modelo, también, ha cumplido las hipótesis clásicas del modelo de regresión lineal múltiple, por lo que la especificación está bien. Sin embargo, se han encontrado una serie de limitaciones a la hora de realizar el trabajo, como la dificultad de obtener algunos datos que podían ser significativos para realizar el modelo, pero al no tener acceso a ellos se ha decidido omitirlos para realizar el análisis. La variable PIB ha resultado no ser significativa de forma individual en el modelo, pero se ha decidido mantener, ya que la idea principal era representar la demanda turística a través de un índice de precios, una magnitud que indique la situación económica del país y una variable indicativa de la oferta turística en Cantabria. Algunos de los resultados obtenidos no son los que se esperaban en un principio:

El PIB ha resultado ser no significativo de forma individual, y en el modelo tiene un coeficiente negativo. El resultado esperado era que la situación económica del país, representado por el PIB, fuese significativo en la entrada del número de turistas, y, además, que un aumento del PIB produjese un incremento del número de turistas al poseer la población una renta mayor. Sin embargo, al realizar el análisis bivalente con el PIB, explicaba bastante bien la variable dependiente, y en este modelo simple sí que era significativa. Esto se puede deber también, a que las variables indicativas del nivel de precios o de la oferta están directamente relacionadas con el turismo, mientras que, la situación económica de un país debería influir en la demanda turística, pero no de forma tan directa.

Otra limitación ha podido ser la representación de los datos de las variables, que se han medido en unidades. Es posible que, si se hubiesen analizado las variables en tasas de crecimiento, se hubiera obtenido un mejor resultado.

También, los datos se recogieron de forma anual, por lo que se habría obtenido una mayor precisión en el modelo si se hubieran tomado de forma trimestral o mensual, para así tener una mayor muestra de los datos.

5 BIBLIOGRAFIA

Cantur, 2017. *Plan de marketing turístico para la campaña de promoción de Cantabria 2018-2019.*

EFE, 2019. *EFE.COM.* [En línea]

Available at: <https://www.efecom.com/efe/espana/economia/el-turismo-aporto-178-000-millones-a-la-economia-espanola-en-2018-2-4-mas/10003-3916283>

Espasa, A., Gomez-Churrua, R. & Jareño, j., 1990. *Un análisis econométrico de los ingresos por turismo en la economía española.*

Esteban Talaya, A., 1987. *Análisis de la demanda: aplicación a la actividad turística de las técnicas de predicción.*

Exceltur, 2014. *Estudio del impacto económico del turismo sobre la economía y el empleo de Cantabria.*

Guisan, M. C. & Aguayo, E., 2008. *Evolución del turismo hotelero en las regiones españolas.*

Guisan, M. C. & Neira, I., 2001. *Un análisis econométrico del turismo hotelero y extrahotelero en las regiones y provincias españolas.*

ICANE, 2017. *Perfil del turista en Cantabria.*

Oficina de turismo cantabro, 2019. *Plan de Marketing turístico de Cantabria 2018-2019*

Pulido San Roman, A., 1966. *Introducción a un análisis econométrico del turismo.*

Wooldrige, J. M. *Introducción a la econometría. Un enfoque moderno.*

6 ANEXO: FUENTE DE DATOS

Encuesta de ocupación hotelera
Periodicidad: Mensual
Resumen: La Encuesta de ocupación hotelera ofrece información sobre viajeros, pernoctaciones y estancia media, distribuidos por país de residencia para los viajeros extranjeros o comunidad autónoma de procedencia para los viajeros españoles, así como la categoría de los establecimientos que ocupan. También proporciona estimaciones del número de establecimientos abiertos, plazas, grado de ocupación y empleo en el sector, según categoría del establecimiento.
Ficha técnica: Tipo de encuesta: continua de periodicidad mensual Ámbito poblacional: todos los establecimientos hoteleros Ámbito geográfico: todo el territorio nacional Periodo de referencia de los resultados: el mes Periodo de referencia de la información: siete días seguidos de cada mes, elegidos aleatoriamente de tal manera que entre todos los establecimientos cubran el mes completo Diseño muestral: muestreo estratificado por provincia y categoría del establecimiento, siendo exhaustivos aquellos estratos de categorías 4 y 5 estrellas de oro Tamaño muestral: aproximadamente 9.250 establecimientos en invierno y 11.200 en verano. Método de recogida: cuestionario cumplimentado directamente por el establecimiento hotelero
URL: http://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736177015&menu=resultados&idp=1254735576863
Variables analizadas: De esta encuesta se tomarán las variables: "nº de viajeros", "nº de establecimientos abiertos", "nº de plazas estimadas" y "personal empleado".

Índice de precios hoteleros
Periodicidad: Mensual
Resumen: El Índice de precios hoteleros es una medida estadística de la evolución del conjunto de precios aplicados por los empresarios a los distintos clientes que se alojan en los hoteles de España. Mide, por tanto, la evolución de los precios del sector desde la óptica de la oferta.
Ficha técnica: Tipo de encuesta: continua de periodicidad mensual Ámbito poblacional: todos los establecimientos hoteleros Ámbito geográfico: todo el territorio nacional Periodo de referencia de los resultados: el mes Periodo de referencia de la información: siete días seguidos de cada mes, elegidos aleatoriamente de tal manera que entre todos los establecimientos cubran el mes completo Diseño muestral: muestreo estratificado por provincia y categoría del establecimiento, siendo exhaustivos aquellos estratos de categorías 4 y 5 estrellas de oro Tamaño muestral: aproximadamente 9.250 establecimientos en invierno y 11.200 en verano. Método de recogida: cuestionario cumplimentado directamente por el establecimiento hotelero
URL: http://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736177015&menu=resultados&idp=1254735576863
Variables analizadas: De esta encuesta se tomarán las variables: "Índice de precios hoteleros"

Índice de precios de consumo
Periodicidad: Mensual
<p>Resumen: El índice de precios de consumo (IPC) es una medida estadística de la evolución de los precios de los bienes y servicios que consume la población residente en viviendas familiares en España.</p> <p>El conjunto de bienes y servicios, que conforman la cesta de la compra, se obtiene básicamente del consumo de las familias y la importancia de cada uno de ellos en el cálculo del IPC está determinada por dicho consumo.</p> <p>A partir de enero de 2002 la metodología del IPC se renovó completamente. Los cambios metodológicos introducidos en este Sistema han hecho del IPC un indicador más dinámico, que se adapta mejor a la evolución del mercado, ya que se pueden actualizar las ponderaciones con más frecuencia. Además, se pueden incluir nuevos productos en la cesta de la compra en el momento en que su consumo comience a ser significativo.</p>
<p>Ficha técnica:</p> <p>Tipo de encuesta: continua de periodicidad mensual</p> <p>Período base: 2016</p> <p>Periodo de referencia de las ponderaciones: año anterior al corriente</p> <p>Muestra de municipios: 177</p> <p>Número de artículos: 479</p> <p>Número de observaciones: aproximadamente 220.000 precios mensuales</p> <p>Clasificación funcional: 12 grupos, 43 subgrupos, 101 clases y 219 subclases; 57 rúbricas y 29 grupos especiales</p> <p>Método general de cálculo: Laspeyres encadenado</p> <p>Método de recogida: agentes entrevistadores en establecimientos y recogida centralizada para artículos especiales</p>
<p>URL:</p> <p>http://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176802&menu=metodologia&idp=1254735976607</p>
<p>Variables analizadas: De esta encuesta se tomarán las variables: "Índice de precios de consumo" e "Índice de precios de hoteles y restaurantes"</p>

Contabilidad nacional trimestral de España: principales agregados (CNTR)
Periodicidad: Trimestral
<p>Resumen: La Contabilidad nacional trimestral de España: principales agregados (CNTR) es una estadística de síntesis, de carácter coyuntural, cuyo objetivo es proporcionar una descripción cuantitativa de la evolución reciente de la economía nacional en el marco de la Contabilidad Nacional de España. Se ofrecen estimaciones del PIB y sus componentes en sus tres enfoques (oferta, demanda y rentas), tanto a precios corrientes como en términos de volumen (en el caso del enfoque de renta, sólo a precios corrientes), así como de empleo (personas, puestos de trabajo, puestos de trabajo equivalentes a tiempo completo y horas trabajadas) y renta nacional. Además, todas las estimaciones se presentan tanto no ajustadas como ajustadas de estacionalidad y calendario.</p>
<p>Ficha técnica:</p> <p>Tipo de operación: continua de periodicidad trimestral.</p> <p>Período base: 2010.</p> <p>Ámbito geográfico: todo el territorio nacional.</p> <p>Periodo de referencia de los resultados: trimestre.</p> <p>Método de recogida: estadística de síntesis.</p>
<p>URL:</p> <p>http://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736164439&menu=ultiDatos&idp=1254735576581</p>
<p>Variables analizadas: De esta encuesta se tomarán las variables: "Producto interior bruto"</p>

Encuesta de población activa
Periodicidad: Trimestral
<p>Resumen: La Encuesta de población activa (EPA) se realiza desde 1964, siendo la metodología vigente en la actualidad la de 2005.</p> <p>Se trata de una investigación continua y de periodicidad trimestral dirigida a las familias, cuya finalidad principal es obtener datos de la fuerza de trabajo y de sus diversas categorías (ocupados, parados), así como de la población ajena al mercado laboral (inactivos).</p> <p>La muestra inicial es de unas 65.000 familias al trimestre que equivalen aproximadamente a 160.000 personas.</p>
<p>Ficha técnica:</p> <p>Tipo de encuesta: continua de periodicidad trimestral</p> <p>Ámbito poblacional: población que reside en viviendas familiares</p> <p>Ámbito geográfico: todo el territorio nacional</p> <p>Período de referencia de los resultados: el trimestre</p> <p>Período de referencia de la información: la semana anterior a la entrevista</p> <p>Tamaño muestral: 3.822 secciones censales, alrededor de 65.000 viviendas y 160.000 personas</p> <p>Tipo de muestreo: bietápico con estratificación en las unidades de primera etapa. Las unidades de primera etapa son las secciones censales y las de segunda etapa son las viviendas familiares habitadas</p> <p>Método de recogida: entrevista personal y telefónica</p>
<p>URL:</p> <p>http://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176918&menu=metodologia&idp=1254735976595</p>
Variables analizadas: De esta encuesta se tomarán las variables: "Tasa de paro"