

ESTUDIO DE LA RELACIÓN EXISTENTE ENTRE EL NÚMERO DE MEDICOS EN LOS HOSPITALES DE LA COMUNIDAD VALENCIANA Y UNA SERIE DE VARIABLES DE ÁMBITO HOSPITALARIO

**TRABAJO FIN DE GRADO
CURSO 2014-2015**



**Autor: Román Carpio Silvente
Tutor: Juan Aparicio Baeza
Fecha de entrega: 20 de Mayo de 2015
Grado de Administración y Dirección de Empresas
Facultad de Ciencias Sociales y Jurídicas de Orihuela
Universidad Miguel Hernández**

ÍNDICE GENERAL

1. Resumen.....	4
2. Introducción.....	5
2.1 La importancia de cuidar la salud.....	5
2.2 El Sistema Sanitario Español.....	6
2.3 Situación actual del Sistema Nacional de Salud.....	7
2.4 Hospitales y camas hospitalarias.....	9
2.5 Situación actual de la sanidad en la Comunidad Valenciana.....	10
3. Objetivos.....	10
4. Metodología.....	11
5. Estudio econométrico.....	15
5.1 Estudio econométrico completo.....	15
1- Análisis preliminar.....	16
2- Ajuste del modelo.....	18
3- Bondad del ajuste.....	20
4- Diagnóstico del modelo.....	22
Procedimiento secuencial de selección de variables.....	27
5.2 Estudio econométrico simplificado.....	28
1- Análisis preliminar.....	28
2- Ajuste del modelo.....	30
3- Bondad del ajuste.....	32
4- Diagnóstico del modelo.....	33
5.3 Estudio econométrico simplificado 2.....	37
1- Análisis preliminar.....	38
2- Ajuste del modelo.....	40
3- Bondad del ajuste.....	41
4- Diagnóstico del modelo.....	42
6- Conclusiones.....	47
7-Bibliografía.....	49

1. RESUMEN

Mediante la elaboración de este estudio se ha tratado de averiguar si existe algún tipo de relación estadística entre el número de médicos en los hospitales de la Comunidad Valenciana y una serie de variables relacionadas con el mundo hospitalario, y al mismo tiempo crear un modelo matemático lineal que explique este tipo de relación.

Los datos se han obtenido de una base de datos de 2009, formada por 26 hospitales de la Comunidad Valencia cuya dependencia funcional es la Agencia Valenciana de Salud y con dependencia patrimonial de la Seguridad Social en algunos casos y en otros de la Comunidad Autónoma Valenciana.

Nuestro desarrollo econométrico consta de tres estudios que se han realizado de manera secuencial, tratando de encontrar el modelo que mejor explique qué variables influyen más en el número de médicos de la Comunidad Valenciana.

En este trabajo hemos sido capaces de determinar un modelo de regresión lineal múltiple con un alto poder explicativo ($R^2=96\%$), con el objetivo de relacionar el número de médicos en las plantillas hospitalarias y otra serie de variables. En nuestro modelo final, las variables estadísticamente significativas resultaron ser camas, que es una variable proxy que nos habla del tamaño del hospital, y número de urgencias atendidas. Según nuestras conclusiones, *ceteris paribus*, si aumentásemos, por ejemplo, en 100 unidades el número de camas de un hospital, la plantilla de médicos debería aumentar en unos 76 nuevos profesionales. Por otro lado, si el número de urgencias aumentara, por ejemplo, en 1000 atenciones/año, el número de médicos aumentaría aproximadamente en dos unidades.



2. INTRODUCCIÓN

2.1 La importancia de cuidar la salud

Según la definición elaborada por la Organización Mundial de la Salud (OMS), "la salud es un estado completo de bienestar físico, mental y social, y no solamente la ausencia de afecciones o enfermedades". Así, queda claro que más allá de patologías, la salud abarca diversos factores que intervienen en la calidad de vida de cada uno. Una persona sana, por tanto, sería aquella que disfruta de una vida plena y feliz a todos los niveles.

Por lo tanto podemos determinar que estar sano es algo más que la ausencia de la enfermedad. Hábitos diarios, como hacer ejercicio, llevar una dieta sana o vivir con optimismo, pueden influir positivamente, sin embargo, situaciones como puede ser la actual crisis económica y los recortes están propiciando un aumento de algunas enfermedades sobre todo mentales.

Según un estudio del Observatorio Europeo de Sistemas y Políticas de Salud (Grupo respaldado por la OMS), la crisis económica actual acompañadas de políticas de recortes presupuestarios en sanidad, que en definitiva, dificultan el acceso a la atención y servicios médicos ha influido negativamente sobre la salud de los ciudadanos, provocando un repunte significativo de enfermedades mentales tales como la depresión o la ansiedad. No podemos olvidar tampoco, que como consecuencia del empeoramiento de las condiciones de vida y la caída de la calidad del sistema público de salud, se está favoreciendo el desarrollo de enfermedades infecciosas que en algunos casos ya se encontraban prácticamente erradicadas.

Centros hospitalarios

Si nos detenemos a definir el concepto de hospital, comprobaremos que ha variado en los últimas décadas, tradicionalmente se consideraba un hospital como la institución sanitaria que tenía como finalidad la prestación fundamental de una prestación sanitaria en régimen de internado.

En la década de los años 90, se pasó de la estructura tradicional de hospitales y ambulatorios, a una forma de organización unificada, que integraba hospitales y centros de asistencia especializada, y en algunos casos de atención primaria del Área Sanitaria. Esta forma de organización, es la que se encuentra más o menos extendida por la geografía nacional, aun así, depende en gran medida de particularidades de cada Comunidad Autónoma, esto es debido en gran medida al hecho de que en el año 2002 finalizó el proceso de transferencias de las funciones y servicios sanitarios a las Comunidades Autónomas, creándose además en el Real Decreto 840/2002 un nuevo organismo denominado Instituto de Gestión Sanitaria al que le corresponde la gestión de los derechos y obligaciones del INSALUD y las prestaciones sanitarias en el ámbito de las ciudades de Ceuta y Melilla.

Por otra parte, un hospital puede estar constituido por un único centro hospitalario o por dos o más que se organizan e integran en el complejo hospitalario. En estos casos es la unidad de dirección y gestión la que sirve para su identificación. De esta forma, un complejo hospitalario puede estar constituido por dos o más hospitales, incluso distantes entre sí y uno o varios centros de especialidades.

2.2 El Sistema Sanitario Español

Fue en 1986, mediante la ley General de Sanidad¹ (Ley 14/1986) cuando fue creado, el denominado Sistema Nacional de Salud (SNS) con la función principal de articular las prestaciones sanitarias públicas. La promulgación de esta ley supuso un paso muy importante para el desarrollo y consolidación del actual Sistema Nacional de Salud.

Antecedentes a la Ley General de Sanidad (Ley 14/1986)

La preocupación de los estados por los problemas de salud y la prevención de enfermedades ha sido una constante desde la edad media cuando se intentaba controlar las epidemias cerrando murallas o prohibiendo la comunicación con las ciudades afectadas.

Más recientemente, se han aprobado diferentes leyes antecesoras de la actual Ley General de Sanidad, destacando las siguientes:

La Ley de Coordinación Sanitaria del 11 de Julio de 1934.

La Ley de Bases de 1944 tras la Guerra Civil Española.

Constitución del Seguro Obligatorio de Enfermedad (SOE), el 14 de Diciembre de 1942.

Ley General de la Seguridad Social de 1974.

Leyes reguladoras de la sanidad española

Ley 14/1986 General de Sanidad.

El 25 de Abril de 1986 se promulgaba la actual Ley General de Sanidad y se creaban las consejerías de sanidad dependientes del Ministerio de Sanidad. La ley nacía para dar respuesta a la Constitución de 1978 en materia de salud pública, que establece en sus artículos 43 y 49 el derecho de todos los ciudadanos a la protección de la salud. Todos los ciudadanos y extranjeros residentes en España, tienen derecho a obtener prestaciones del sistema sanitario español. Por otro lado, prevé que las competencias a nivel sanitario pueden ser transferidas a las Comunidades Autónomas. Conforme se han realizado transferencias en materias de sanidad, se han ido creando los servicios de Salud de las Comunidades Autónomas.

Ley 16/2003 de cohesión y calidad del Sistema Nacional de Salud

En 2003, una vez que todas las comunidades han asumido competencias a nivel sanitario y existe un modelo de financiación más o menos estable, se produce una modificación de la ley para dar respuesta a las nuevas necesidades generadas en la sociedad.

Al igual que la sociedad española cambia sus hábitos y costumbres de manera veloz, el Sistema Nacional de Salud debe adaptarse a la manera de vivir y enfermar de la sociedad actual.

¹ Ley 14/1986, de 25 de abril, General de Sanidad. Publicado en BOE núm. 102 de 29 de Abril de 1986

Por ello la Ley crea diferentes organismos, tales como la Agencia Española de Medicamentos y Productos Sanitarios, la Agencia de Evaluación de las Tecnologías, la Comisión de Recursos Humanos, el Instituto de Salud Carlos III, el Instituto de Información Sanitaria, la Agencia de Calidad del sistema Nacional de Salud, la Comisión Asesora de Investigación en Salud o el Observatorio del Sistema Nacional de Salud, que pretenden potenciar la participación y la cohesión a nivel sanitario, siendo el Consejo Interterritorial del Sistema Nacional de Salud el más importante a la hora de tomar decisiones de forma rápida y vinculando a las partes a la asunción de las mismas.

La Ley clasifica en función del ámbito de colaboración en: Prestaciones del Sistema Nacional de Salud, la Farmacia, Profesionales Sanitarios, la Investigación, el Sistema de Información Sanitaria y la Calidad del Sistema Sanitario.

El Sistema Sanitario en la Comunidad Valenciana

Como ya hemos comentado anteriormente, las diferentes Comunidades Autónomas, incluida la Valenciana, mediante el proceso de descentralización, han ido obteniendo las competencias en materia de sanitaria. La Comunidad Valenciana, mediante el Real Decreto Constitutivo del Servicio de Salud Autonómico 1612/1987, de 27 de noviembre constituyó la Agencia Valenciana de la Salud.²

2.3 Situación actual del Sistema Nacional de Salud

El gasto sanitario público en el año 2012 en España supuso 64.150 millones de euros, lo que representa el 6,2 por ciento del producto interior bruto (PIB). Como se observa en la tabla siguiente, podemos ver una reducción de más de 6.429 millones de euros respecto al gasto en 2009. Si comparamos los datos de 2009 con los datos provisionales del año 2013, la reducción se aproxima a las 8.000 millones de euros.

El gasto per cápita en 2012 ascendió a 1.357 euros por habitante, que respecto a valores del año 2009, supone una reducción de 153 euros por habitante.

² Con la publicación en el Boletín Oficial del Estado el 27 de noviembre de 1987 (con fecha de efectividad 1 de enero de 1988), se traspasaban las funciones del Instituto Nacional de la Salud a la Comunidad Valenciana, así como los correspondientes servicios e instituciones y medios personales, materiales y presupuestarios precisos para el ejercicio de aquéllas. BOE» núm. 312, de 30 de diciembre de 1987, páginas 38171 a 38176 (6 págs.)

	2008	2009	2010	2011	2012	2013
Millones de euros	66.918	70.579	69.357	68.059	64.150	63.006
Porcentaje sobre el PIB	6,2%	6,7%	6.6%	6,5%	6.2%	-
Euros por Habitante	1.450	1.510	1475	1.442	1.357	-

Gasto sanitario público consolidado. Millones de euros, porcentajes sobre el PIB y euros por habitante. España, 2008-2012 (ESTADÍSTICA DE GASTO SANITARIO PÚBLICO 2012: Principales resultados. (Edición abril 2014). Los datos de 2013 son provisionales y proporcionados por el Ministerio de Hacienda.

Según el informe de la Organización para la Cooperación y el Desarrollo Económico (OCDE) elaborado en 2011, España es junto a Irlanda, Reino Unido y Grecia los únicos países de la OCDE que han reducido el gasto sanitario varios años consecutivos. Entre los años 2000 a 2009, el gasto en sanidad de los países de la OCDE, creció a ritmo del 4´8% mientras que España lo hizo al 5´6%, sin embargo, la progresión se ralentizó en 2010 como consecuencia de la crisis en España y también en los países de la OCDE, siendo del 0´5%. En 2011, el gasto sanitario en el grupo de países considerados desarrollados creció 0´6% y en España por culpa del agravamiento de la crisis, paso a decrecer a un ritmo del -2´8%.

Pese a la crisis económica, y a los duros recortes que ha sufrido la sanidad española, según estudios independientes hechos por la prestigiosa agencia Bloomberg (2013), que pone en común datos del Banco Mundial, el Fondo Monetario Internacional y la Organización Mundial de la Salud, España se seguía situando entre los países más eficientes del mundo, más concretamente el estudio, la situaba en el año 2013 en el puesto 5 en eficiencia del gasto; solo superada por Israel, Japón, Singapur y Hong Kong.

En Europa, después de España, se situaba Italia, después nos encontrábamos otros países como Suiza o Suecia. A nivel mundial, Estados Unidos, aun siendo aún una súper potencia mundial, se encontraba en las últimas posiciones de la lista. Brasil, es dentro de los países de la OCED, el que peores datos registra, situándose en la cola.

Sin embargo, la disminución del gasto público en sanidad ha hecho mella en el ranking que ha vuelto a publicar la agencia Bloomberg en 2014. En este caso, España se desplomaba a la decimocuarta posición. España se queda lejos de los cinco primeros puestos y por el contrario, otros países del entorno como Italia (tercera en el ranking), Francia (octava), Reino Unido (decimo) o Noruega (undécimo) escalan posiciones y se sitúan en este último estudio por encima de España. A nivel mundial, Estados Unidos se sitúa en la puesto 44 y el las dos últimas posiciones de un ranking de 51 países encontramos a Brasil y Rusia.

A la hora de llevar a cabo los estudios, tuvieron en cuenta la esperanza de vida, el coste relativo de los cuidados de la salud per cápita, y el coste per cápita de los cuidados de salud en términos absolutos. La ponderación del primer término se valoró en un 60%, el segundo criterio en un 30% y el tercero en un 10%

Otros datos que nos ayudan a entender la situación de la sanidad española respecto del resto de países OCDE es que España se sitúa por encima de la media respecto al número

de médicos por cada 1000 habitantes, 4´1 frente al 3´2 y se sitúa muy por debajo de la media, respecto al número de auxiliares de enfermería, 5´5 frente al 8´7. Otro dato negativo para los ciudadanos españoles es que el número de camas hospitalarias por cada 1000 habitantes también se encuentra por debajo de la media de la OCED que es de 4´8 y en España de 3´2.

2.4 Hospitales y camas hospitalarias.

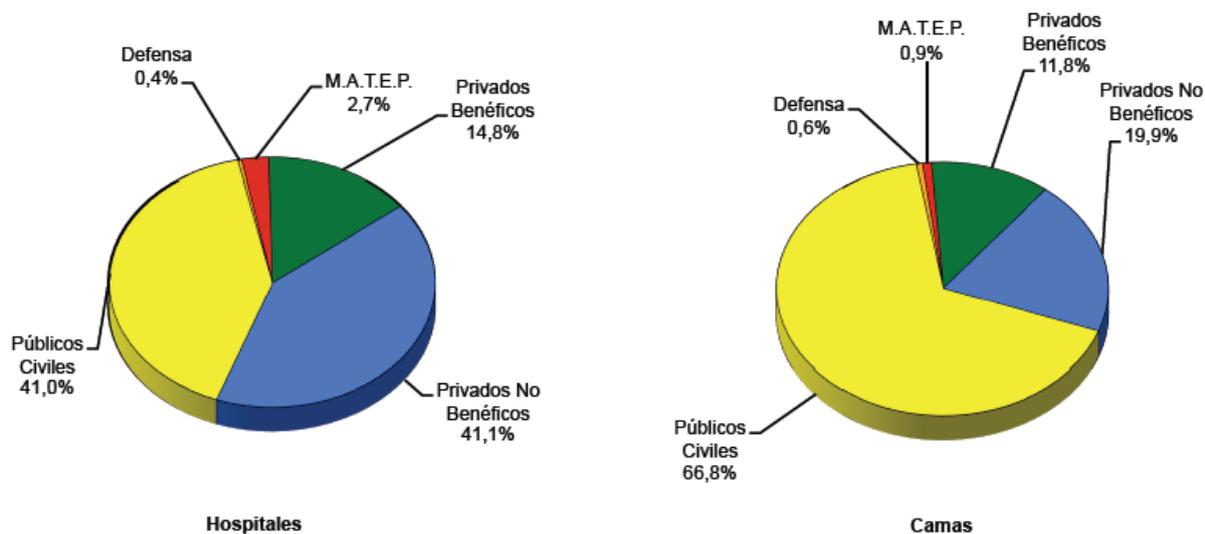
Si nos centramos en el presupuesto que la administración destinó a sostener la actividad de hospitales y centros de salud, ésta, cayó un 16,3% durante el periodo de mayores recortes, el que va de 2009 a 2013. Como consecuencia de esa reducción en el presupuesto, los estudios y los barómetros sanitarios realizados por el Ministerio de Sanidad y el CIS, avalan listas de espera cada vez más extensas y una percepción de los ciudadanos cada vez más negativa. A mediados de 2009 los pacientes que aguardaban más de seis meses para operarse eran el 5% del total; en 2013 ya suponían el 13,7%.

Según Juan Oliva, economista y expresidente de la Asociación de Economía de la Salud, a los investigadores les gustaría ir más allá y poder comprobar si la reducción en sanidad en los presupuestos generales tiene efectos concretos sobre indicadores asistenciales como las muertes evitables o las hospitalizaciones potencialmente evitables relacionadas con la reagudización de enfermedades crónicas (diabetes, EPOC, insuficiencia cardíaca). El problema es que los datos se suelen publicar con bastante retraso y este hecho se ha acrecentado por la crisis económica que además ha reducido el número de estudios. Oliver pone como ejemplo el efecto del crecimiento actual de las tasas de pobreza, que entre otras cosas incidirá en problemas de salud en el futuro.

Ante la necesidad realizar recortes en el Sistema Nacional de Salud (SNS) para reducir el déficit presupuestario, desde las Comunidades Autónomas se ha optado por la opción más sencilla, cebarse con los recursos humanos. Las Comunidades han metido la tijera dejando de cubrir bajas médicas y no sustituyendo al personal médico de enfermeros que se jubilan.

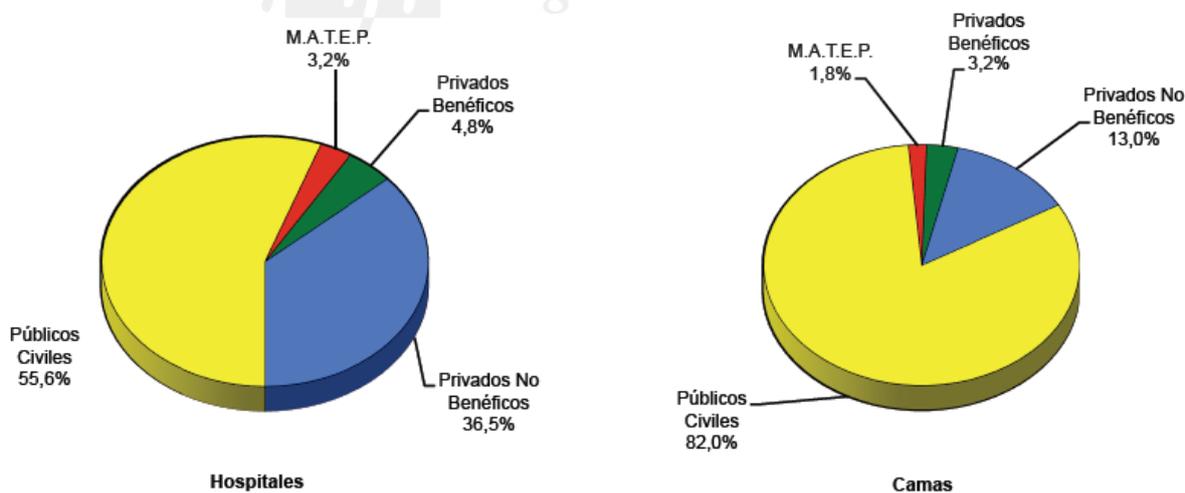
En España, el SNS ha perdido casi 5000 trabajadores (médicos y expertos) entre el 2010 y el 2013, según revela el Informe Anual del Sistema Nacional de Salud de 2013. El SNS contó en 2013 con 3.004 centros de salud y 10.112 consultorios de atención primaria. El colectivo profesional más numeroso es el de la enfermería: 165.000 personas. Algo más de 115.000 médicos atendieron a los ciudadanos en el sistema público. Además de los recursos humanos, una de las formas de medir la caída del gasto es comprobar el número de camas que en los hospitales se mantienen operativas (las llamadas "*funcionantes*", que son menos que las instaladas). En 2010, según el informe, había 115.418 camas "*funcionantes*" en el sistema. Tres años después cayeron a 109.484. Es decir, casi 6.000 camas operativas menos, un 5,1% de descenso.

**DISTRIBUCIÓN DE HOSPITALES Y CAMAS
 TOTAL NACIONAL**



(Según dependencia funcional)

**DISTRIBUCIÓN DE HOSPITALES Y CAMAS
 COMUNIDAD VALENCIANA**



(Según dependencia funcional)

Catálogo Nacional de Hospitales 2013, actualizado a 31 de diciembre de 2012. Ministerio de Sanidad, Servicios Sociales e Igualdad y las Consejerías de Sanidad de las Comunidades Autónomas, el Ministerio de Defensa, los órganos competentes de las Ciudades Autónomas de Ceuta y Melilla.

2.5 Situación actual de la sanidad en la Comunidad Valenciana.

A nivel autonómico, Valencia, Alicante y Castellón disponen de un total de 64 hospitales de los que 35 son totalmente públicos, de los 29 hospitales privados restantes, (12 con concierto y 17 sin concierto con la sanidad pública)

La importancia de la Sanidad Privada en la Comunidad Valenciana

Según el Barómetro de la Sanidad Privada impulsado por el instituto IDIS, el sector hospitalario privado representa el 43% de los hospitales y el 18% de las camas en Valencia. Aproximadamente el 15% de los valencianos tiene sanidad privada y la Comunidad Valenciana, se consolida como la cuarta con mayor número de camas privadas del conjunto nacional tras Cataluña, Madrid y Andalucía.

La estimación de gasto sanitario en el año 2012 supuso en la Comunidad Valencia 8.498 millones de euros, de los que el 3.282 millones de euros, lo que supone el 28% del gasto total en sanidad³.

3. OBJETIVO

El objetivo fundamental de este estudio es determinar si existe algún tipo de relación estadística, y analizar si está es modelizable linealmente, entre la variable "Número de médicos en la Comunidad Valenciana y una serie de indicadores relacionados con la calidad y el servicio ofrecido de los hospitales valencianos tales como son: número de camas, número de ingresos hospitalarios, número de pacientes que han sido atendidos en Urgencias y número de pacientes que han acudido a las consultas de atención primaria.

En el caso de que sea posible, crearemos un modelo matemático (de carácter lineal) que permita explicar o predecir el número de Médicos en la Comunidad Valenciana a través de los valores de las restantes variables.

³ Fuente:

- Gasto sanitario público y gasto sanitario público (conciertos): Ministerio de Sanidad, Servicios Sociales e Igualdad. Cuenta satélite del gasto público, Serie 2002-2012. MSSSI, 2013.
- Gasto sanitario privado: estimado a partir de los datos del INE, Encuesta presupuestos familiares, COICOP grupo 6 correspondiente a Salud y grupo 12.3 de Protección social; e ICEA. El seguro de salud, Estadística año 2012. Informe n 1280, marzo 2013.

4. METODOLOGÍA

Para la determinación de relaciones lineales entre variables haremos uso de gráficos de dispersión y cálculo de correlaciones (simples y parciales), junto al test de la t de Student. Para el establecimiento del mejor modelo lineal (regresión lineal múltiple), utilizaremos el estimador de máxima verosimilitud de los parámetros del modelo, junto con el método Stepwise ligado al criterio AIK (Akaike Information Criteria). Para analizar la bondad del ajuste haremos uso del error estándar residual, la tabla de ANOVA (análisis de la Varianza) y el coeficiente de determinación. Finalmente, las hipótesis del modelo de regresión lineal múltiple (linealidad, normalidad, homocedasticidad, media cero e incorrelación) serán analizadas mediante el test de Shapiro –Wilk, el test de Breusch-Pagan, el test de Durbin-Watson, el histograma de los residuos, el gráfico qq-plot y otra serie de gráficos habituales. Para todos nuestros análisis asumiremos un error no superior al 5%.

La guía de referencia utilizada para la elaboración de este análisis estadístico es el manual, Modelos Lineales Aplicados en R (Aparicio, Martínez Mayoral & Morales, Modelos Lineales Aplicados en R (2004)), así como los video-tutoriales impartidos en la asignatura de Econometría por el profesor D. Juan Aparicio.

El programa utilizado para todos y cada uno de los pasos seguidos en este estudio estadístico es R (www.r-project.org). R es un software libre (GNU), cuyo uso se ha generalizado en los últimos años entre estadísticos y usuarios de estas técnicas en todo el mundo.

Definición de las variables

A continuación mostramos la definición de las variables utilizadas en este estudio:

Médicos: número total de médicos en plantilla en los Hospitales de la Comunidad Valenciana en el año 2009. El término médico se refiere al profesional que, tras cursar los estudios necesarios y obtener el título correspondiente, cuenta con autorización legal para ejercer la medicina.

Camas: número de camas totales que se encuentran instaladas y dispuestas las 24 horas del día para el uso regular de pacientes hospitalizados de las que se dispone en los hospitales de la Comunidad Valenciana durante el año 2009. Se cuentan como camas de hospital, las camas de adultos y niños (cunas e incubadoras usadas para el tratamiento de

niños hospitalizados). No son camas de hospital las que se usan sólo temporalmente con fines de diagnóstico o tratamiento (rayos X, camas de parto, camas de recuperación,...) o las camillas ubicadas en salas de emergencia

Ingresos: número de pacientes que han sido ingresados en hospitales de la Comunidad Valencia en el año 2009. El ingreso hospitalario es la aceptación formal de un paciente por el hospital para su atención médica, observación, tratamiento y recuperación. Todo ingreso al hospitalario involucra la ocupación de una cama hospitalaria y la mantención de una historia clínica para el registro de todas las atenciones otorgadas. No deben considerarse ingresos los bebés nacidos vivos sanos o los nacidos muertos en el establecimiento hospitalario, las personas fallecidas mientras son trasladadas al hospital y las personas que fallecen en la sala de espera de la Unidad de Urgencias. No es necesario que el paciente pernocte para que se considere ingreso.

Urgencias: cifra de personas que ha acudido a la Unidad de Urgencias de hospitales de la Comunidad Valencia durante el año 2009. La unidad de Urgencias es un departamento en el que se atiende a heridos y enfermos que necesitan cuidados médicos inmediatos.

Primconsult: número de ciudadanos que acudieron a consultas de atención primaria en los hospitales de la Comunidad Valencia durante el año 2009. Estos servicios atienden las necesidades básicas de salud de la población; la atención primaria incluye las visitas médicas generales, la prescripción de medicamentos, la atención preventiva y la detección de enfermedades. Los médicos de atención primaria son los encargados de desviar las consultas hacia los especialistas.

Muestra de datos

Hemos utilizado una base de datos de 2009 formada por 26 hospitales de la Comunidad Valencia cuya dependencia funcional es la Agencia Valenciana de Salud y con dependencia patrimonial de la Seguridad Social en algunos casos y en otros de la Comunidad Autónoma Valenciana.

HOSPITALES	Médicos	camas	ingresos	Urgencias	primconsult
Hospital Arnau de Vilanova	273	287	13264	64958	34225
Hospital Clínico de Valencia	652	574	24158	152875	64208
Hospital de Denia (Marina Salud)	308	206	9689	41352	54095
Hospital de Gandía	228	236	13224	66091	81105
Hospital de General de Elche	431	473	22415	118680	57170
Hospital de la Marina Baixa (La Vila)	204	262	12664	59360	28417
Hospital de La Plana	183	242	12090	55598	46866
Hospital de la Ribera	337	301	21385	92047	124260
Hospital de Manises	163	173	5197	36216	48966
Hospital de Onteniente	55	47	2658	27930	0
Hospital de Orihuela	281	348	15151	61596	34241
Hospital de Requena	96	85	3665	24887	40123
Hospital de Sagunto	210	242	11504	57232	31385
Hospital de Torrevieja	201	210	15918	58132	112814
Hospital de Vinaroz	129	129	6089	32325	32016
Hospital Dr Peset	622	515	24319	133600	52562
Hospital General de Alicante	771	813	33136	143355	72852
Hospital General de Castellón	476	563	20261	86214	58700
Hospital General de Elda	299	339	18304	74960	55862
Hospital General de Valencia	818	565	25123	144089	84140
Hospital La Fe	1353	1267	46784	200442	103023
Hospital Lluís Alcanyis (Xàtiva)	191	238	12027	56945	29437
Hospital Malvarrosa	48	45	1585	0	14833
Hospital Provincial de Castellón	186	253	4401	11017	40084
Hospital San Juan	391	370	17461	76848	53116
Hospital Virgen de los Lirios	180	276	11440	42848	37859

Tabla 1. Base datos con las variables objeto de estudio

5. ESTUDIO ECONOMÉTRICO

5.1 Estudio Econométrico Completo

En nuestro estudio, la variable respuesta es el número de médicos por hospital en la Comunidad Valenciana, siendo las variables explicativas el número de camas, la cifra de ingresos hospitalarios, el número de pacientes atendidos en servicio de urgencias de cada hospital y el número de pacientes atendidos en consultas de atención primaria.

Con nuestro estudio vamos a tratar de describir la relación entre la variable respuesta y las variables explicativas a través de un hiperplano lineal en las variables explicativas, lo más próxima posible a los valores observados de la respuesta. Como es bien sabido, la idea básica de la regresión múltiple es explicar todo lo posible sobre la variable respuesta con una variable explicativa, y a continuación utilizar información adicional que aporta otra variable explicativa y así sucesivamente con todas las variables explicativas para completar la predicción sobre la variable respuesta.

La formulación del modelo para nuestro análisis es:

$$\text{Médicos} = \beta_0 + \beta_1 * \text{camas} + \beta_2 * \text{ingresos} + \beta_3 * \text{Urgencias} + \beta_4 * \text{primconsult}$$

Definición de los parámetros:

β_0 : término de interceptación o intercepto.

$\beta_1, \beta_2, \dots, \beta_4$ son los coeficientes de regresión, son las pendientes de sus respectivas variables, por lo tanto el valor β_1 me indica que “ceteris paribus”, si aumento en una unidad la variable camas, la variable Médicos aumentara en el valor β_1 . El valor del coeficiente β_1 cuantifica el peso que tiene la variable camas explicando la respuesta media de Médicos. Cabría decir lo mismo para el resto de parámetros.

El modelo de regresión múltiple se va a analizar en cuatro fases:

- 1.-Análisis preliminar
- 2.-Ajuste del modelo
- 3.-Bondad del ajuste
- 4.-Diagnostico del modelo

FASE 1: ANÁLISIS PRELIMINAR

El primer paso a dar cuando queremos modelizar un banco de datos es llevar a cabo la inspección gráfica de los mismos, con el fin de descubrir de qué tipo son las relaciones entre las variables, si las hay. Así, estudiaremos el gráfico de dispersión y éste nos dará una primera aproximación a las relaciones existentes entre las variables.

En esta fase también realizaremos un análisis numérico de las correlaciones simples y parciales que nos dará una visión más objetiva sobre si la asociación se puede catalogar como lineal o no y nos permitirá cuantificar el grado de asociación lineal entre las variables explicativas.

Por otra parte, el coeficiente de correlación parcial nos cuantificará el grado de relación lineal existente entre la variable respuesta y cada una de las variables explicativas, considerando siempre la información de las restantes variables a la hora de calcular el coeficiente.

Entrando ya en el análisis, la Figura 1 nos muestra todos los gráficos de dispersión posibles. Así se deduce que existe cierta relación lineal “positiva” (creciente) entre el número de médicos en plantilla en los hospitales de la Comunidad Valenciana y número de camas y número de ingresos hospitalarios y el número de entradas por urgencias. No se aprecia tendencia lineal clara entre el número de médicos y el número de consultas en atención primaria.

No obstante, calcularemos a continuación los coeficientes de correlación pertinentes para asegurarnos numéricamente de si realmente existen relaciones lineales entre la variable respuesta y las variables explicativas o no.

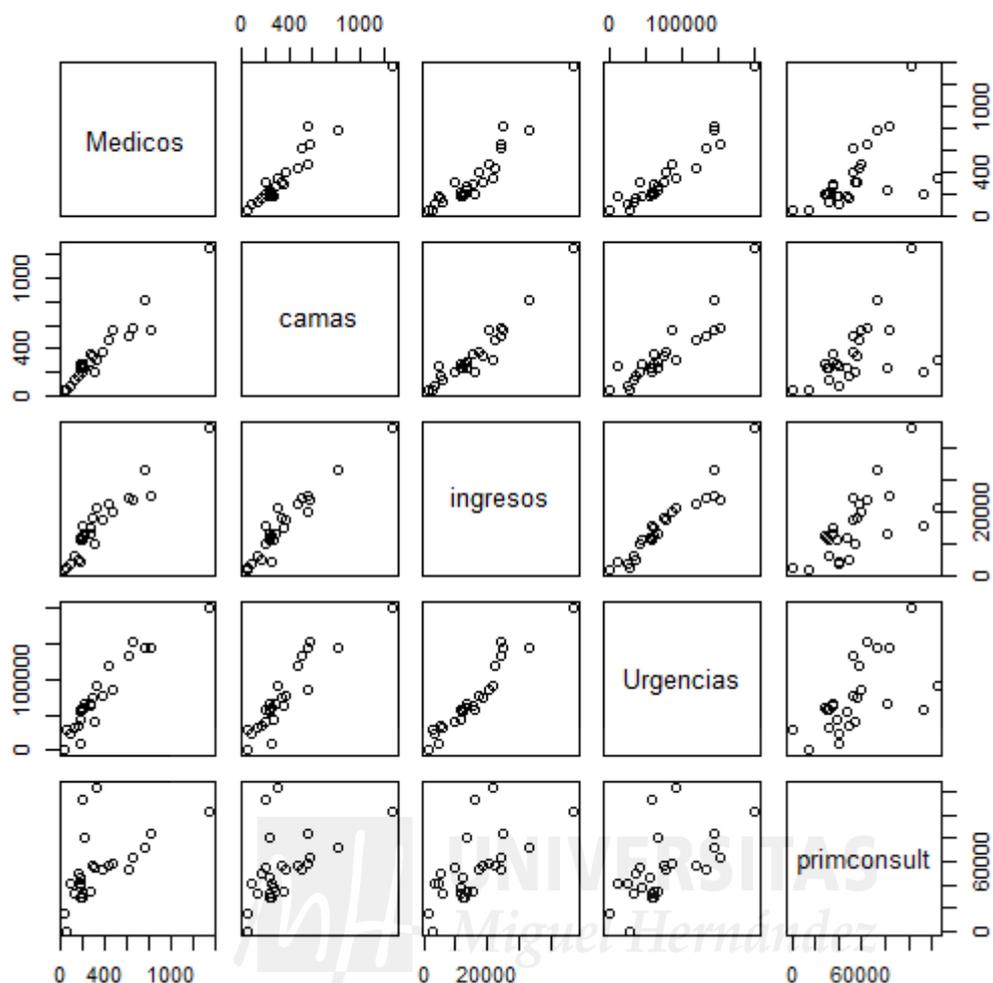


Figura 1.- Relación entre las variable y Médicos.

Analizando los resultados de la Tabla 2, se confirma unas correlaciones simples altísimas entre número de médicos y Camas e Ingresos y Urgencias, con unos valores muy próximos a 1. Esto nos indica que tienen relación positiva, es decir, al aumentar las variables explicativas el número de médicos también aumenta.

Las variables por nivel de importancia son Camas (0.9686), Ingresos (0.9416), Urgencias (0.9363) y por último, primconsult (0.5602).

VALORES DE CORRELACIÓN SIMPLE					
	Médicos	Camas	Ingresos	Urgencias	primconsult
Médicos	1	0.9686856	0.9416815	0.9363900	0.5602617

Tabla 2.- Valores correlación simple, Modelo Completo

En cuanto a las correlaciones parciales, la Tabla 3, nos muestra los valores estimados. En este sentido, observamos que la variable más importante, es el número de Camas en hospitales de la Comunidad Valenciana con un valor de 0,7964 ya que tiene un mayor grado de relación lineal con la variable respuesta Médicos, siguiendo en importancia: Urgencias (0.6201), Ingresos (0.4420) y Primconsult (0.3644). En general hemos obtenido unas correlaciones parciales poco elevadas, indicando que las relaciones lineales entre las variables explicativas y la variable respuesta no son demasiado evidentes

Cabe destacar que Camas, Ingresos y Urgencias habían obtenido valores cercanos a 1 en las correlaciones simples, y nos hacía pensar que podrían llegar a ser importantes para el modelo, y sin embargo al obtener las correlaciones parciales, en presencia del resto de covariables parece indicar que no son tan relevantes.

VALORES DE CORRELACIÓN PARCIAL					
	Médicos	Camas	Ingresos	Urgencias	primconsult
Médicos	0.0000	0.79647	-0.44205	0.62010	0.36440

Tabla 3.- Valores de Correlación Parcial, Modelo Completo

FASE 2: AJUSTE DEL MODELO

En esta segunda fase, vamos a ver los valores que estima el modelo para β_0 , β_1 , β_2 , β_3 y β_4 , los intervalos de confianza al 95% para cada uno de los coeficientes y el contraste de hipótesis asociado a la cuestión de si cada coeficiente es igual o no a cero. Así, y en concreto, si se rechazase la hipótesis de coeficiente cero, sabríamos que esa variable es útil, es decir, contiene suficiente información para explicar la variable respuesta Médicos, en presencia del resto de covariables.

Los valores de los parámetros, por ejemplo β_1 , representan la estimación del incremento que experimenta la variable respuesta cuando la variable explicativa asociada a dicho coeficiente aumenta su valor en una unidad y las demás variables se mantienen constantes.

En la Tabla 4, mostramos las estimaciones de máximo verosímiles de los parámetros, así como el p-valor asociado a los contrastes anteriormente explicados. Además, en la tabla aparecen otros valores que más tarde analizaremos y que se encuentran relacionados con

la bondad del ajuste. En definitiva, el modelo ajustado se correspondería con la siguiente expresión:

$$\text{Medicos} = -7.225e+01 + (1.080e+00) \cdot \text{camas} + (-1.650e-02) \cdot \text{ingresos} + (3.254e-03) \cdot \text{Urgencias} + (1.149e-03) \cdot \text{primconsult}$$

RESUMEN DEL MODELO					
	β_0	β_1	β_2	β_3	β_4
Parámetros	-7.225e+01	1.080e+00	-1.650e-02	3.254e-03	1.149e-03
p-valor	0.0110	5.45e-06	0.0347	0.0016	0.0874
Valor Residual Estándar					59.52
Múltiple R-Squared					0.9649
Ajuste R-Squared					0.9582
P-Value					5.906e-15

Tabla 4.-Resumen del modelo, Modelo Completo



Por otra parte, en la Tabla 5 se muestran los intervalos de confianza al 95% para cada uno de los coeficientes del modelo lineal.

INTERVALOS DE CONFIANZA		
	Extremo Inferior	Extremo Superior
Intercepto	-1.261251e+02	-18.366654454
Beta de Camas	7.078411e-01	1.451945130
Beta de Ingresos	-3.169432e-02	-0.001305834
Beta de Urgencias	1.385699e-03	0.005121982
Beta de Primconsult	-1.835610e-04	0.002482045

Tabla 5.-Intervalos confianza, Modelo Completo

IC95% $\beta_1 = (7.078411e-01, 1.451945130)$

IC95% $\beta_2 = (-3.169432e-02, -0.001305834)$

IC95% $\beta_3 = (1.385699e-03, 0.005121982)$

IC95% $\beta_4 = (-1.835610e-04, 0.002482045)$

A continuación, estudiaremos si cada coeficiente puede ser o no igual a cero. En la Tabla 4 aparecen los p-valores asociados a cada par de contraste-coeficiente. Por ejemplo, para el caso del intercepto, el contraste de hipótesis a resolver sería el siguiente:

$H_0: \beta_0 = 0$

$H_1: \beta_0 \neq 0$

En este caso el p-valor es igual a 0.0110, como no es superior a 0.05, podemos rechazar H_0 . Por tanto, nos indica que en presencia de las demás variables ésta es relevante y no asumiremos que el parámetro es nulo.

p-valor $\beta_1 = 5.45e-06 < 0,05$ Rechazamos H_0 a favor de H_1

p-valor $\beta_2 = 0.0347 < 0,05$ Rechazamos H_0 a favor de H_1 .

p-valor $\beta_3 = 0.0016 < 0,05$ Rechazamos H_0 a favor de H_1 .

p-valor $\beta_4 = 0.0874 > 0,05$ No rechazamos H_0 , asumiremos que el parámetro es nulo.

En beta 4, el p-valor es superior a 0.05, por lo que no se rechaza H_0 y es posible que dicho betas sea cero, lo que significa que la variables asociada (primconsult) no es lo suficientemente útil, es decir, la información contenida en esta variable no explica la variable Médicos y es probable que la información contenida en ella se solape.

FASE 3: BONDAD DEL AJUSTE

Cuando realizamos el ajuste en un modelo de regresión lineal debemos verificar si dicho modelo proporciona un buen ajuste a la hora de predecir la variable respuesta, en nuestro

caso, Médicos. Para ello contamos con varias medidas para cuantificar la bondad del ajuste de diversos modos, en concreto tres:

1.- Error estándar residual: es una medida de bondad del ajuste relativa a la escala de medida utilizada. Se prefieren modelos con menor error residual estimado. No obstante una vez estimado el error estándar residual del modelo es habitual determinar si éste es grande o pequeño a través del cálculo del coeficiente de variación: $CV = (\text{error estándar residual}/\text{la media de Médicos}) \%$. Así, si el valor del coeficiente de variación es inferior al 10% nos indica que estamos ante un buen modelo, según este criterio.

En nuestro caso el error estándar residual tiene un valor 59.52, y el cv de 17.03192% (> 10% recomendado por el instituto nacional de estadística). No aprueba el primer criterio del error estándar residual, aun así es “sólo” relativamente grande.

2.- Tabla de Anova: en este criterio cuantificamos cuánta de la variabilidad contenida en los datos ha conseguido ser explicada por dicho modelo. Un modelo es bueno si la variabilidad explicada es abundante, o lo que es lo mismo, si las diferencias entre los datos y las predicciones según el modelo son pequeñas. Construir la tabla de Anova consiste en descomponer la variabilidad de los datos en la parte que es explicada por el modelo y la parte que se deja sin explicar, es decir, la variabilidad de los residuos, y compararlas y valorarlas estadísticamente para ver si la variabilidad explicada por el modelo es suficientemente grande. En particular, damos respuesta al contraste siguiente:

$$H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$$

H1: lo contrario

A este respecto, si obtenemos un p-valor inferior a 0.05 rechazamos H_0 y el modelo será bueno, mientras que por el contrario, si el p-valor es superior al 0.05 el modelo será malo; contrastamos si las variables explicativas explican suficientemente bien la variable respuesta a través del modelo lineal propuesto.

El p-valor asociado al contraste de la tabla de ANOVA de nuestro modelo es 5.906e-15 por lo tanto inferior a 0.005, y rechazamos H_0 . En definitiva el modelo es bueno según este criterio.

3.- Coeficiente de determinación R_2 : es otro estadístico útil para comprobar la bondad del ajuste, se define como la proporción de la varianza que es explicada por la recta de regresión. Obtenemos un número que nos indicará el valor en tanto por ciento de la variabilidad total de los datos. Cuando el valor sea cercano a 1 (100%) implicará que buena parte de la varianza es explicada por la recta de regresión y cuando R_2 tenga un valor cercano a cero significará que prácticamente toda la variabilidad de los datos queda sin

explicar por el modelo de regresión. Comparemos, además el valor del coeficiente R_2 con el valor de coeficiente R_2 ajustado para ver si son parecidos, ya que cuando hay varias variables explicativas, el mero hecho de tener muchas variables puede distorsionar el R_2 original, por eso lo comparamos con el ajustado. Si el ajustado fuera muy diferente, tomaríamos este como bueno y no el coeficiente de R_2 .

Para nuestros datos, observamos un valor del coeficiente $R^2 = 0.9649$ (>65%), por lo que el modelo es muy bueno, explica el 96.49% de la variabilidad total de la variable Médicos.

Por otra parte, el valor de coeficiente R^2 ajustado = 0.9582, es prácticamente igual al coeficiente de R^2 . Por lo tanto, damos por válido el valor del coeficiente R_2 original, que como hemos indicado es muy bueno.

FASE 4. DIAGNÓSTICO DEL MODELO:

Después de ajustado un modelo y habiendo superado las pruebas de bondad de ajuste (salvo Error estándar residual), ahora en la cuarta fase vamos a verificar si el modelo satisface las hipótesis básicas del modelo de regresión lineal para los errores:

- Normalidad
- Varianza constante (Homocedasticidad)
- Incorrelación
- Media Cero

El análisis de los residuos nos permitirá detectar deficiencias en la verificación de estas hipótesis, así como descubrir observaciones anómalas o especialmente influyentes en el ajuste. Una vez encontradas las deficiencias, si existen, debemos analizar si replantearnos el modelo empleando alguna transformación sobre las variables.

La fase del diagnóstico del modelo suele ser gráfica, aunque también existen varios test estadísticos que mostraremos y resolveremos. Los residuos de un modelo lineal los definimos como habitualmente se hace en la literatura, como las desviaciones entre las observaciones y los valores ajustados.

1-Normalidad: construiremos, en primer lugar, un histograma para comparar nuestro modelo con la campana de Gauss; a este respecto, cuanto más se desvíe, menos normal será el modelo. Además, analizaremos el gráfico qq-plot, donde se representan los residuos ordenados versus los cuantiles correspondientes de una normal; si la normalidad de los residuos es cierta, los puntos han de estar alineados con la diagonal. Desviaciones de la

diagonal más o menos severas en las colas o en el centro de la distribución implican desviaciones con respecto a la hipótesis de normalidad.

En nuestro caso particular, la Figura 2, el histograma no se asemeja a la campana de Gauss, no hay simetría y se muestra más peso en la cola de la izquierda de la distribución que en la de la derecha.

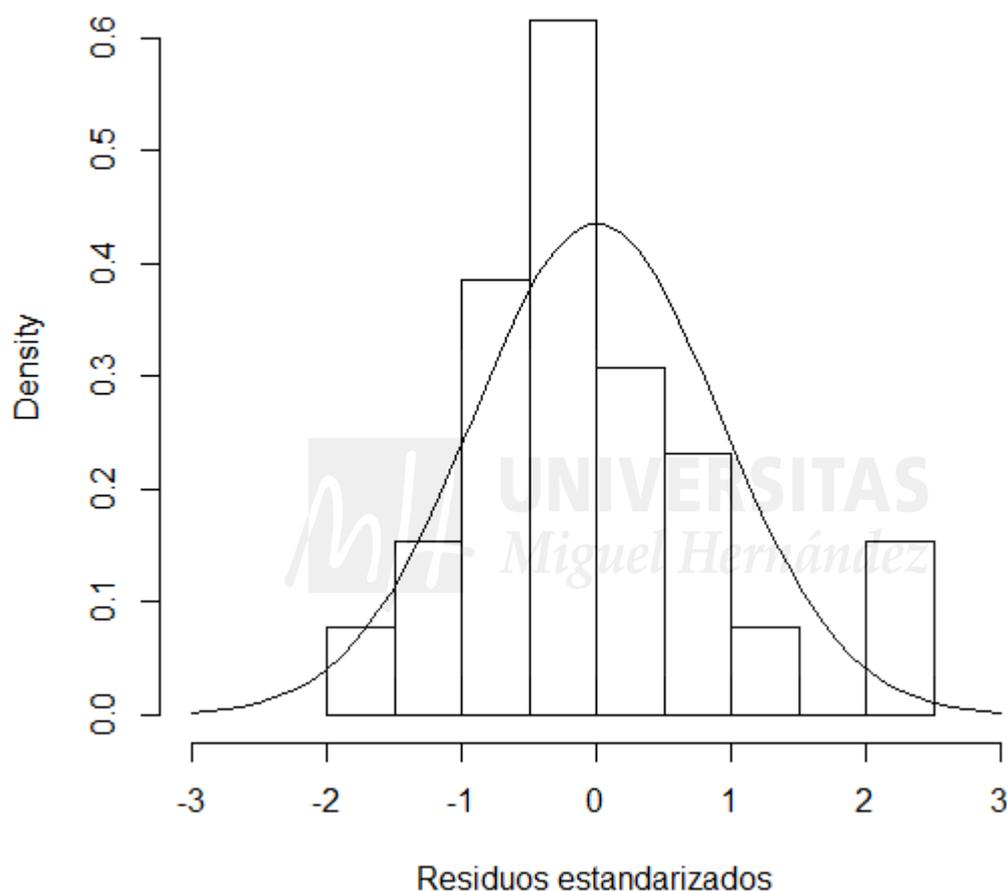


Figura 2.-Histograma. Modelo Completo

Por otra parte, en el gráfico qqplot, la mayoría de los puntos están sobre la línea recta, lo que indica que si hubiera problemas con la hipótesis de normalidad, estos no son muy graves.

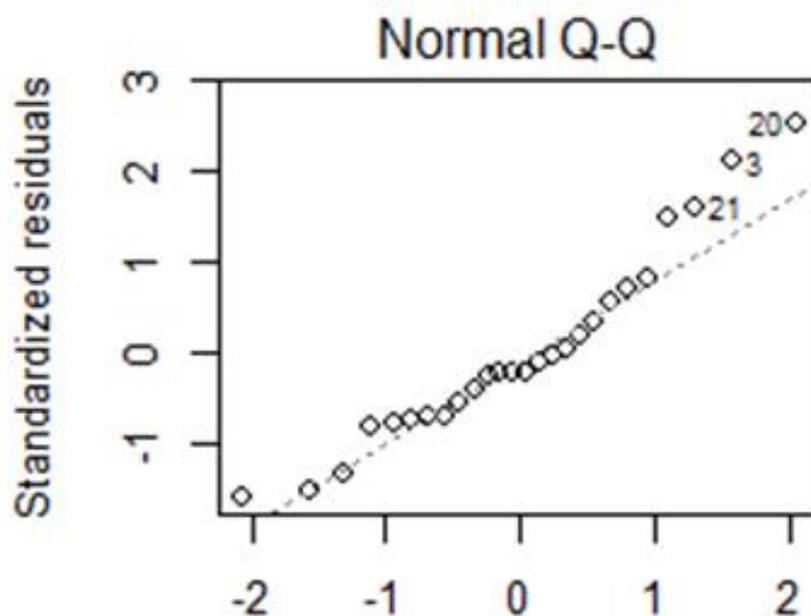


Figura 3.-Gráfico QQPLOT. Modelo Completo



La normalidad también será contrastada a través del contraste de hipótesis de Shapiro-Wilk:

H_0 : $e \sim \text{Normal}$

H_1 : lo contrario

En este caso, obtenemos un p-valor de $0.173 > 0,05$, por lo que no podemos rechazar la hipótesis de normalidad y con lo cual asumiré que los errores son normales. En definitiva, esta hipótesis podemos asumir que se verifica.

2- Homocedasticidad: (hipótesis de varianza constante) Construiremos un gráfico de residuos versus valores predichos de la variable respuesta Médicos y numéricamente resolveremos el contraste de hipótesis de Breusch-Pagan.

En el gráfico lo deseable es que los residuos aparezcan representados en una banda horizontal sin tendencia alrededor del cero.

En el gráfico, lo deseable es que los residuos aparezcan representados en una banda horizontal sin tendencias alrededor del cero.

Resolvemos a continuación el contraste de hipótesis debido a Breusch y Pagan:

H_0 : Homocedasticidad

H_1 : Heterocedasticidad

Obtenemos en concreto un p-valor de $0.2915 > 0.05$. No podemos rechazar H_0 . Por lo tanto, asumiremos la Homocedasticidad, es decir, que la varianza es constante. Esta hipótesis se verifica.

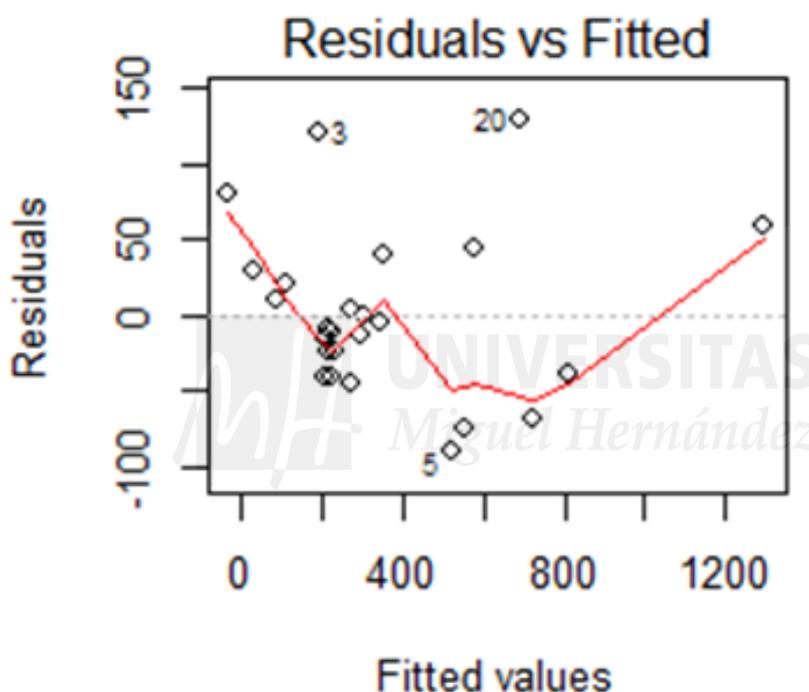


Figura 4.-Residuos frente a valores predichos. Modelo Completo

3.- Incorrelación: la Incorrelación significa que las observaciones de la variable respuesta no están correladas entre sí, es decir, los valores de unas no afectan a los de otras. Para comprobar esta hipótesis construiremos, en primera instancia, el gráfico del residuo i frente al anterior $i-1$.

Gráficamente no se ve ninguna tendencia clara, ni creciente ni decreciente.

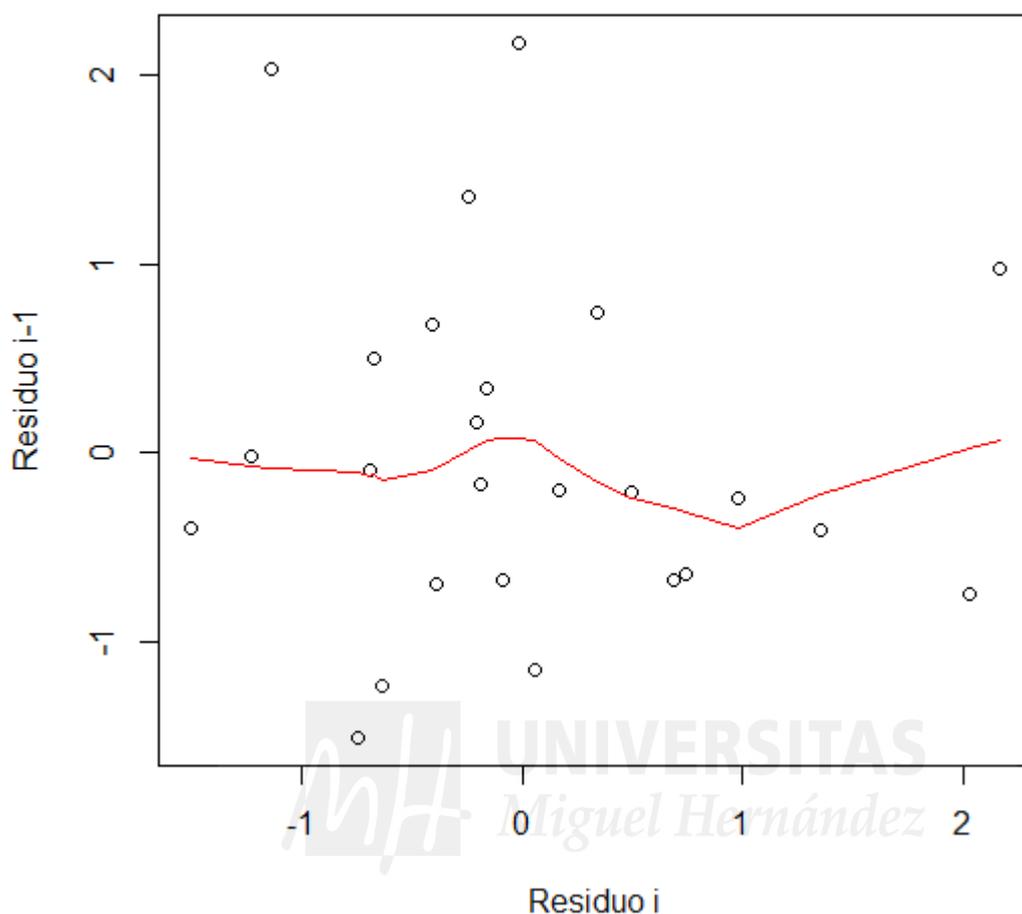


Figura 5.-Gráfico Incorrelación. Modelo Completo

A continuación intentaremos comprobar la hipótesis de forma numérica a través de un p-valor.

En este sentido, el test que debemos de usar se debe a Durbin y Watson:

$H_0: \rho = 0$ (incorrelación)

$H_1: \rho \neq 0$

En este caso, obtendremos un $p\text{-valor} = 0.7317 > 0.05$ No podemos rechazar H_0 , lo que indica que no se puede rechazar la hipótesis de Incorrelación el modelo es bueno para este criterio: hay Incorrelación.

4.- Media Cero: Esta hipótesis siempre se cumple porque usamos el método de máxima verosimilitud, que nos asegura que la media de los residuos siempre es cero.

PROCEDIMIENTO SECUENCIAL DE SELECCIÓN DE VARIABLES

Mediante el criterio AIC y el método Stepwise vamos a estudiar si nuestro modelo original se puede reducir en número de variables explicativas. El criterio AIC nos indica que valores pequeños son preferidos a valores grandes. Este método nos proporciona un valor AIC global (del modelo completo) y unos valores AIC para el caso de eliminar una cualquiera de nuestras variables explicativas. A continuación, detallamos los resultados obtenidos.

AIC Original = 216.94	
VARIABLE	AIC
Primconsult	218.64
Ingresos	220.59
Urgencias	227.56
Camas	241.10

Tabla 6.- Stepwise, paso 1

Excluir primconsult generan mejores modelos.

AIC Original= 216.94	
Variable	AIC
Ingresos	220.59
Urgencias	227.56
Camas	241.10

Tabla 7.-Stepwise, paso 3

El modelo resultante de aplicar el modelo será:

Médicos ~ Ingresos + Urgencias + Camas

El criterio AIC ya ha excluido la variable menos relevante y el nuevo modelo resultante es de la siguiente manera:

Médicos = $\beta_0 + \beta_1$ Ingresos + β_2 Urgencias + β_3 Camas

A continuación vamos a realizar el análisis pormenorizado del nuevo modelo reducido.

5.1 Estudio Econométrico Simplificado

Al igual que para el estudio completo de regresión múltiple vamos a analizar en las mismas cuatro fases el nuevo modelo simplificado:

- 1.-Análisis preliminar
- 2.-Ajuste del modelo
- 3- Bondad del ajuste
- 4- Diagnóstico del modelo

FASE 1: ANÁLISIS PRELIMINAR

Como en el modelo anterior vamos a llevar a cabo la inspección gráfica de los datos para este caso reducida a las variables Camas, Ingresos y Urgencias, para observar de qué tipo son las relaciones entre las variables si las hay mediante el gráfico de dispersión.

La Figura 6 nos muestra los gráficos de dispersión del modelo reducido y se puede observar de nuevo cierta relación lineal entre todas las variables.

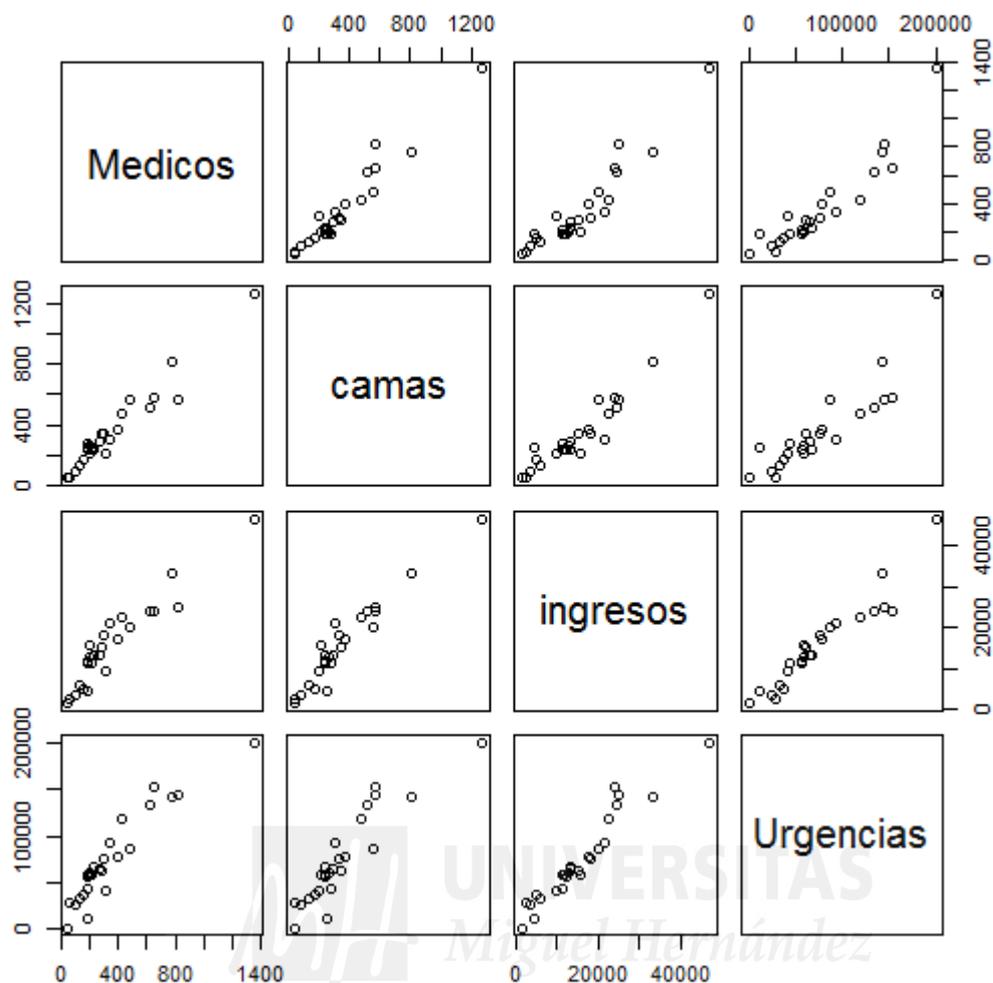


Figura 6.- Relación entre las variables explicativas y la variable respuesta del Modelo simplificado.

Realizamos a continuación el análisis numérico para hallar las correlaciones simples y parciales con el fin de corroborar si realmente existen las asociaciones lineales apreciadas en el gráfico.

Como podemos apreciar en los resultados de la Tabla 8, los valores obtenidos son iguales a los obtenidos en el modelo completo. Se confirma relación lineal entre la variable Médicos y las variables Camas (0.9686856), Ingresos (0.9416815), al obtener unos valores muy cercanos a 1. Se aprecia también relación lineal con la variable Urgencias (0.9363900).

VALORES DE CORRELACIÓN SIMPLE				
	Médicos	Camas	Ingresos	Urgencias
Médicos	1	0.9686856	0.9416815	0.9363900

Tabla 8.-Valores Correlación Simple, Modelo Simplificado

Las correlaciones parciales las tenemos representadas en la Tabla 9. La variable más importante en este caso sigue siendo Camas, con un valor de 0.77034, es la que mayor grado de relación lineal tiene con la variable respuesta Médicos. Le siguen en importancia la variable Urgencias (0.55496) e ingresos (-0.28667).

VALORES DE CORRELACIÓN PARCIAL				
	Médicos	Camas	Ingresos	Urgencias
Médicos	0.00000	0.77034	-0.28667	0.55496

Tabla 9.-Valores Correlación Parcial, Modelo Simplificado

FASE 2: AJUSTE DEL MODELO.



En la segunda fase veremos los valores que estima el modelo para β_0 , β_1 , β_2 y β_3 , sus respectivos intervalos de confianza al 95% y el contraste de hipótesis para responder a la cuestión de si cada coeficiente es igual a cero o no. Por lo tanto cuando rechazamos la hipótesis de coeficiente cero obtendremos las variables que son útiles, es decir, descubriremos las variables que contienen suficiente información para explicar la variable respuesta Médicos. Estimaremos el valor de los parámetros con el criterio de máximo verosímil.

En la Tabla 10, mostramos las estimaciones máximo verosímiles de los parámetros, así como el p-valor asociado a los contrastes anteriormente explicados. Como en el modelo completo también añadimos valores relacionados con la bondad del ajuste.

El modelo ajustado para el estudio reducido corresponde a la expresión siguiente:

$$\text{Médicos} = -4.850e+01 + 9.177e-01 * \text{Camas} + (- 8.554e-03) * \text{Ingresos} + 2.859e-03 * \text{Urgencias}$$

Con el nuevo modelo observamos una relación inversa con la variable Ingresos, mientras que la relación es directa con Camas y Urgencias.

Que exista relación inversa quiere decir, que si aumentamos la variable ingresos en una unidad, por ejemplo, manteniendo el resto de variables constantes (*ceteris paribus*), la variable respuesta Médicos disminuirá en $-8.554e-03$ unidades.

RESUMEN DEL MODELO				
	β_0	β_1	β_2	β_3
Parámetros	-4.850e+01	9.177e-01	-8.554e-03	2.859e-03
p-valor	0.04979	1.07e-05	0.17442	0.00488
Valor Residual Estándar				62.45
Múltiple R-Squared				0.9595
Ajuste R-Squared				0.954
P-Value				1.809e-15

Tabla 10.-Resumen del modelo. Modelo Simplificado

En la Tabla 11 se muestran los intervalos de confianza al 95% para cada uno de los coeficientes del modelo lineal.

INTERVALOS DE CONFIANZA		
	Extremo Inferior	Extremo Superior
Intercepto	-9.695568e+01	-0.049086977
Beta de Camas	5.818634e-01	1.253607437
Beta de Ingresos	-2.119335e-02	0.004085470
Beta de Urgencia	9.641568e-04	0.004754061

Tabla 11.-Intervalos de confianza. Modelo Simplificado

IC95% β_1 = (5.818634e-01, 1.253607437)

IC95% β_2 = (-2.119335e-02, 0.004085470)

IC95% β_3 = (9.641568e-04, 0.004754061)

A continuación, estudiaremos si cada coeficiente puede ser o no igual a cero. En la Tabla 10 aparecen los p-valores asociados a cada par de contraste-coeficiente. Por ejemplo, para el caso del intercepto, el contraste de hipótesis a resolver sería el siguiente

H0: $\beta_0 = 0$

H1: $\beta_0 \neq 0$

En este caso el p-valor es igual a 0.04979, como es inferior a 0.05, podemos rechazar H0. Por tanto, nos indica que en presencia de las demás variables ésta es relevante.

p-valor $\beta_1 = 1.07e-05 < 0,05$ Rechazamos H0 a favor de H1.

p-valor $\beta_2 = 0.17442 > 0,05$ No rechazamos H0, asumiremos que el parámetro es nulo.

p-valor $\beta_3 = 0.00488 < 0,05$ Rechazamos H0 a favor de H1.

FASE 3: BONDAD DEL AJUSTE

Como anteriormente cuando realizamos el ajuste en un modelo de regresión lineal debemos de verificar si dicho modelo proporciona un buen ajuste a la hora de predecir la variable respuesta, Médicos. Cuantificaremos esta viabilidad de diversos modos, en concreto tres:

1.-Error estándar residual: hacemos la medición de la bondad del ajuste relativa a la escala de medida utilizada. De los datos de la Tabla 10 obtenemos el valor residual estándar del modelo, el cual tiene un valor de 62.45. Para determinar si éste valor es grande o pequeño realizamos el cálculo del coeficiente de variación $cv = 100 * (62.45 / \text{la media de Médicos}) \%$ lo que da un resultado de $cv = 17.87035 (> 10\%$ recomendado por el instituto nacional de estadística). No aprueba el primer criterio del error estándar residual. El valor es prácticamente idéntico al valor del modelo anterior.

2-Tabla de Anova: en este criterio cuantificamos cuánta de la variabilidad contenida en los datos ha conseguido ser explicada por dicho modelo.

Damos respuesta al siguiente contraste de hipótesis:

H0: $\beta_1 = \beta_2 = \beta_3 = 0$

H1: lo contrario

El p-valor de nuestro modelo es $1.809e-15$ que como es menor que 0,05, rechazo H_0 y asumimos que el modelo es bueno al igual que en el modelo global. Las variables explicativas explican suficientemente bien la variable respuesta a través del modelo lineal propuesto.

3.-Coeficiente de determinación: obtenemos un valor que nos indicará el valor en tanto por ciento de la variabilidad total de los datos, será la proporción de la varianza que es explicada por la recta de regresión. Cuanto más cercano a 1 (100%) sea el valor, más parte de la varianza es explicada por la recta de regresión, y cuánto más se acerque a cero el valor, indica que prácticamente toda la variabilidad de los datos queda sin explicar por el modelo de regresión.

Como anteriormente, compararemos el valor del coeficiente de R_2 con el valor de coeficiente R_2 ajustado para ver si son parecidos, en caso de que no lo sean descartaremos el valor del coeficiente de R_2 por estar distorsionado y daremos por válido el coeficiente de R_2 ajustado.

El coeficiente de R_2 es 0.9595 (>65%), explica el 95,95% de la variabilidad total de la variable Médicos, por lo que es un modelo muy bueno.

El coeficiente ajustado de R_2 tiene un valor similar de 0.954 por lo que consideramos bueno el valor de coeficiente de R_2 . Por lo tanto, damos por válido el valor del coeficiente R_2 original.

FASE 4. DIAGNÓSTICO DEL MODELO:

En la cuarta fase vamos a verificar si el modelo satisface las hipótesis básicas del modelo de regresión lineal para los errores:

- Normalidad
- Varianza constante (Homocedasticidad)
- Incorrelación
- Media Cero

1.-Normalidad: haremos un estudio gráfico viendo el gráfico QQPLOT donde representamos los residuos ordenados versus los cuantiles correspondientes de una normal y el histograma como en el modelo anterior.

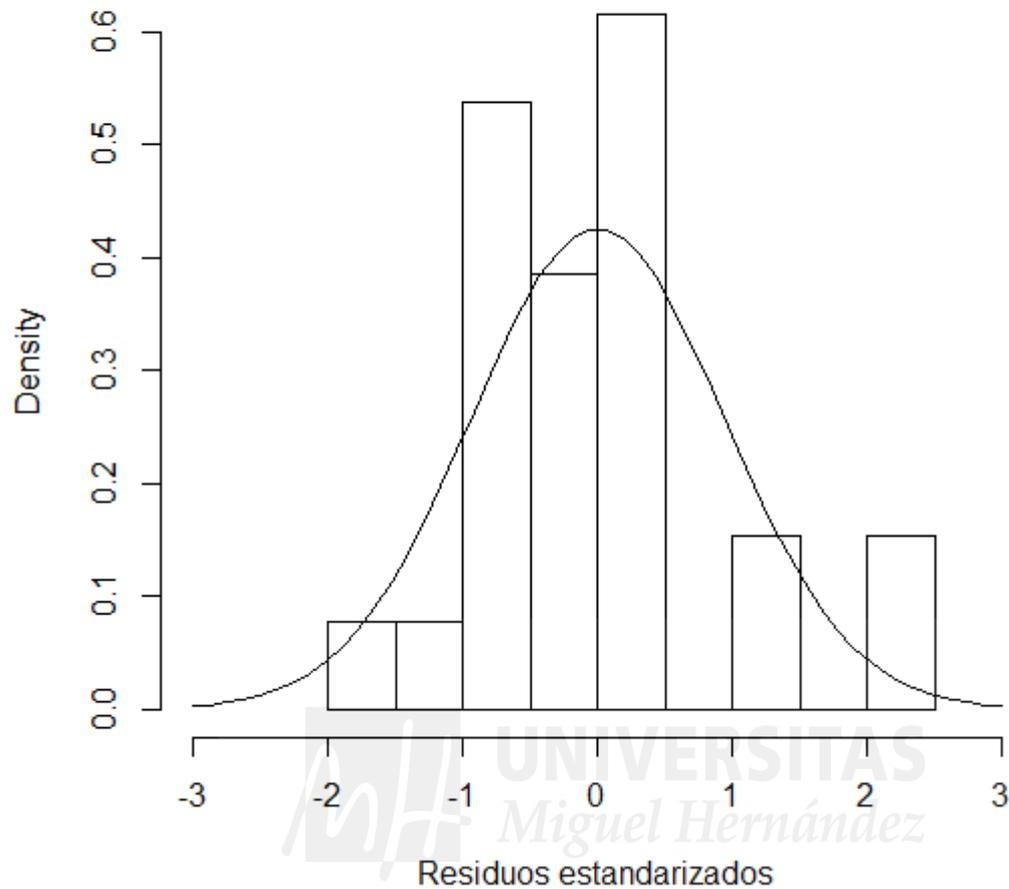


Figura 7.- Histograma Modelo Simplificado

La Figura 7 muestra, que al igual que sucedía en el modelo completo, en el modelo simplificado, el histograma no se asemeja a la campana de Gauss, no hay simetría.

En el gráfico qqplot que se muestra en la Figura 8, la mayoría de los puntos vuelven a estar sobre la línea recta, lo que indica que si hubiera problemas con la hipótesis de normalidad, estos no son muy graves.

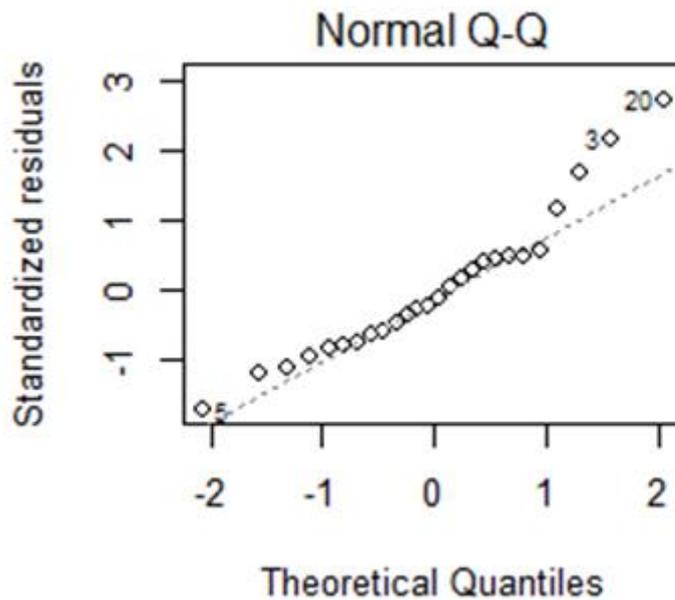


Figura 8.- Gráfico QQPLOT . Modelo Simplificado

La normalidad también será contrastada a través del contraste de hipótesis de Shapiro-Wilk:

H0: $e \sim \text{Normal}$

H1: lo contrario

En este caso, obtenemos un p-valor de $0.09568 > 0,05$, por lo que no podemos rechazar la hipótesis de normalidad y con lo cual asumiré que los errores son normales. En definitiva, esta hipótesis podemos asumir que se verifica.

2 -Homocedasticidad: (hipótesis de varianza constante). Realizaremos un análisis gráfico sobre el gráfico de residuos versus valores predichos de la variable respuesta Médicos y numéricamente resolveremos el contraste de hipótesis de Breush-Pagan.

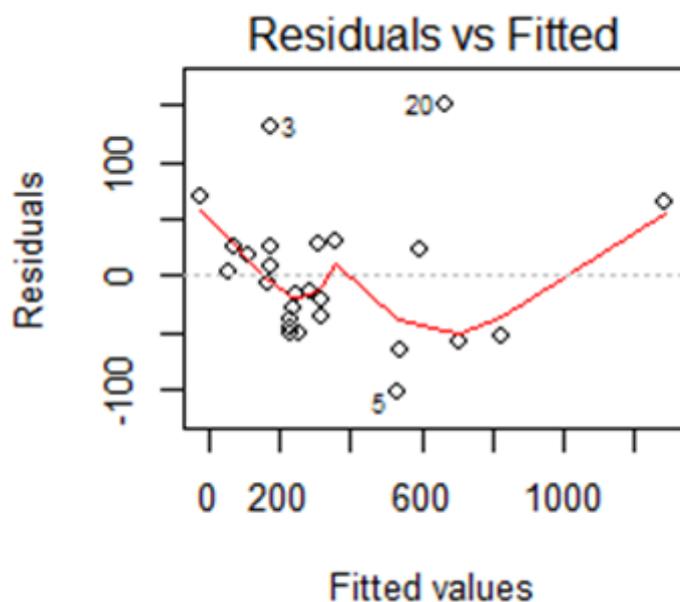


Figura 9.- Residuos frente a valores predichos. Modelo Simplificado



Resolvemos a continuación el contraste de hipótesis debido a Breusch y Pagan:

H0: Homocedasticidad

H1: Heterocedasticidad

Obtenemos en concreto un p-valor de $0.342 > 0.05$. No podemos rechazar H0. Por lo tanto, asumiremos la Homocedasticidad, es decir, que la varianza es constante. Esta hipótesis se verifica.

3.-Incorrelación: vamos a observar si los valores de unas variables afectan a las de otras, para comprobarlo construiremos el gráfico del residuo i frente anterior $i-1$ y realizaremos el contraste de hipótesis de Durbin y Watson para hacer la comprobación numérica.

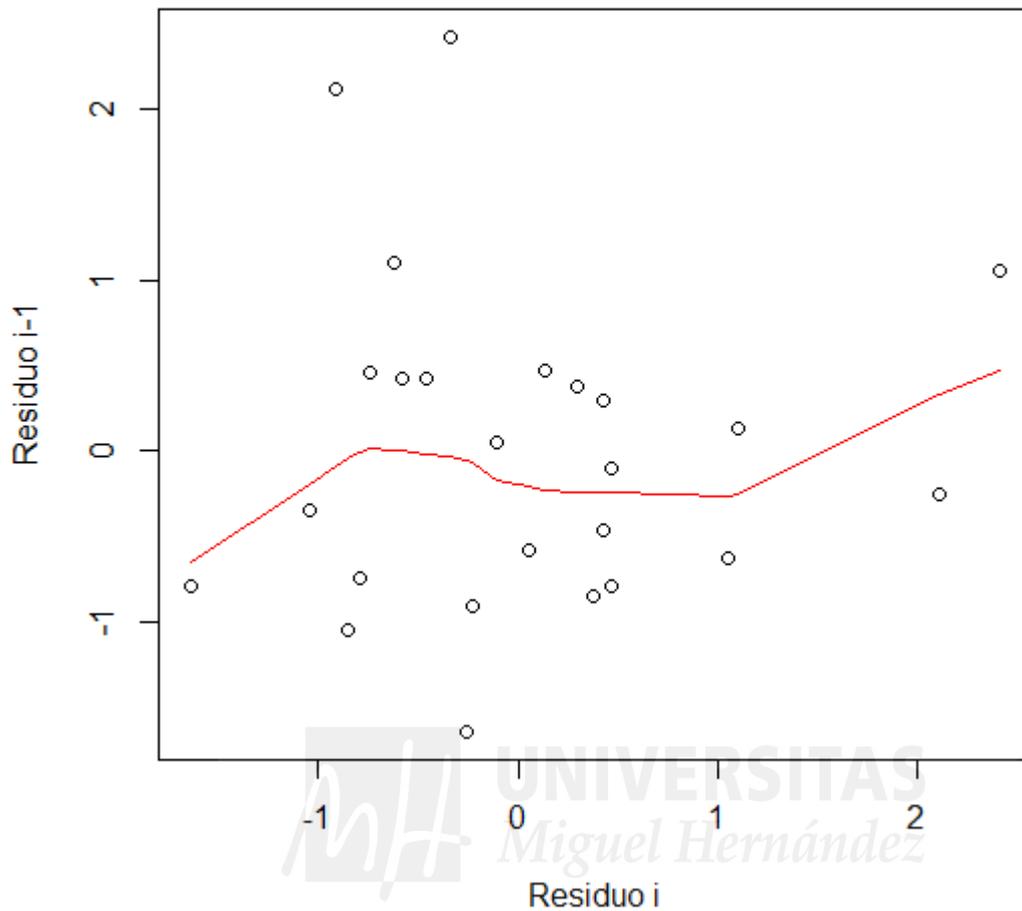


Figura 10.- Gráfico Incorrelación. Modelo Simplificado

En este sentido, el test que debemos de usar se debe a Durbin y Watson:

H0: $\rho = 0$ (incorrelación)

H1: $\rho \neq 0$

En este caso, obtendremos un p-valor = $0.8886 > 0.05$ No podemos rechazar H0, lo que indica que no se puede rechazar la hipótesis de Incorrelación el modelo es bueno para este criterio: hay Incorrelación.

4.- Media Cero: Esta hipótesis siempre se cumple porque usamos el método de máxima verosimilitud, que nos asegura que la media de los residuos siempre es cero.

5.3 Estudio Econométrico Simplificado 2

A la vista de los resultados obtenidos, quitamos la variable ingresos porque vuelve a ser no significativa en presencia de las restantes covariables. Volvemos a analizar las mismas cuatro fases en el nuevo modelo simplificado:

- 1.-Análisis preliminar
- 2.-Ajuste del modelo
- 3- Bondad del ajuste
- 4- Diagnóstico del modelo

FASE 1: ANÁLISIS PRELIMINAR

Como en el modelo anterior volvemos a llevar a cabo la inspección gráfica de los datos, en este caso reducida a las variables Camas y Urgencias. Mediante el gráfico de dispersión queremos observar de qué tipo son las relaciones entre las variables.

La Figura 11 nos muestra los gráficos de dispersión del modelo reducido y se puede observar de nuevo cierta relación lineal entre todas las variables.

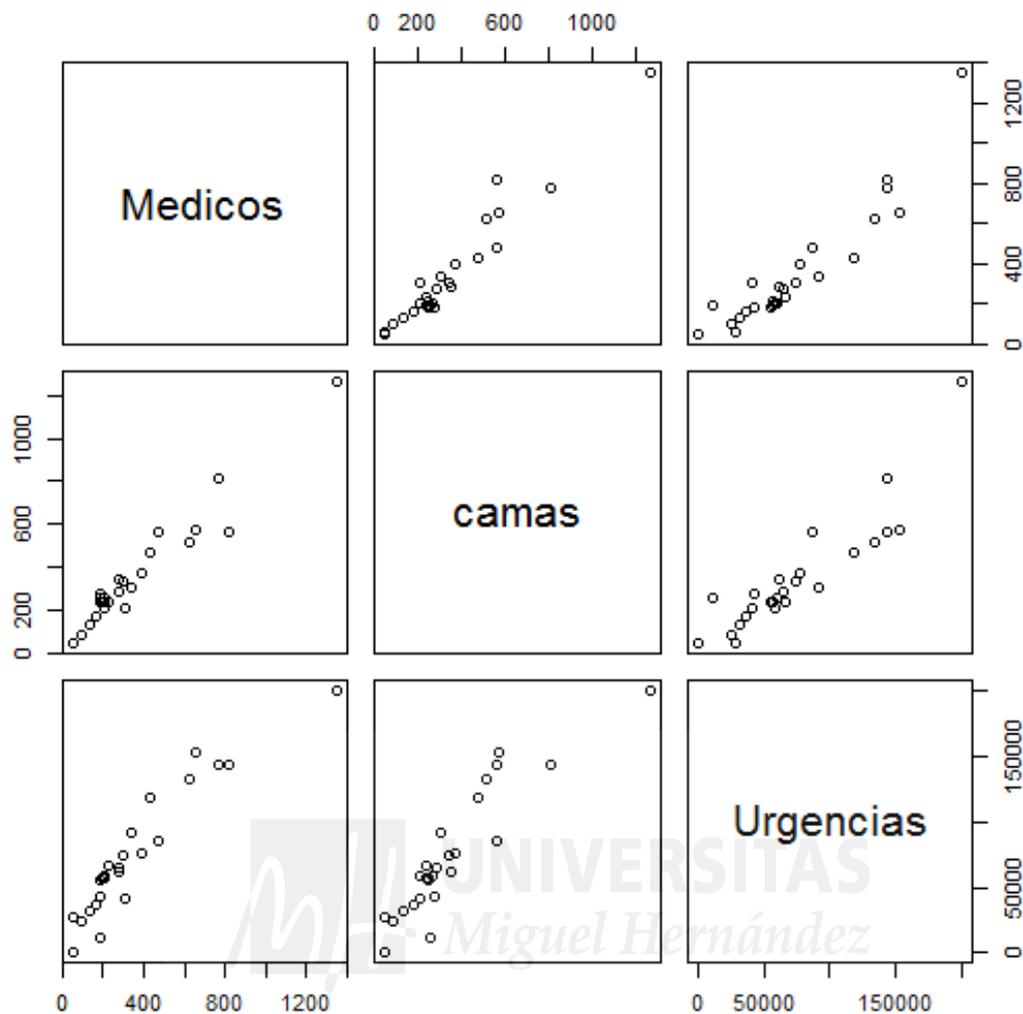


Figura 11.- Relación entre las variables explicativas y la variable respuesta del Modelo simplificado 2.

Realizamos el análisis numérico en busca de las correlaciones simples y parciales con el fin de corroborar si realmente existen las asociaciones lineales apreciadas en el gráfico.

Según los resultados de la Tabla 11, los valores obtenidos son iguales a los obtenidos en el modelo completo, valores muy cercanos a 1.

VALORES DE CORRELACIÓN SIMPLE			
	Médicos	Camas	Urgencias
Médicos	1	0.9686856	0.9363900

Tabla 11.-Valores Correlación Simple, Modelo Simplificado 2

Las correlaciones parciales las tenemos representadas en la Tabla 12. La variable más importante es Camas, con un valor de 0.80128, le sigue la variable Urgencias (0.53369)

VALORES DE CORRELACIÓN PARCIAL			
	Médicos	Camas	Urgencias
Médicos	0.00000	0.80128	0.53369

Tabla 12.-Valores Correlación Parcial, Modelo Simplificado 2

FASE 2: AJUSTE DEL MODELO.

En la segunda fase volvemos a ver los valores que estima el modelo para β_0 , β_1 y β_2 , sus respectivos intervalos de confianza al 95% y el contraste de hipótesis para responder a la cuestión de si cada coeficiente es igual a cero o no.

En la Tabla 13, mostramos las estimaciones máximo verosímiles de los parámetros, así como el p-valor asociado a los contrastes anteriormente explicados.

El modelo ajustado para el estudio reducido corresponde a la expresión siguiente:

$$\text{Médicos} = -5.650e+01 + 7.587e-01 * \text{Camas} + 1.918e-03 * \text{Urgencias}$$

RESUMEN DEL MODELO			
	β_0	β_1	β_2
Parámetros	-5.650e+01	7.587e-01	1.918e-03
p-valor	0.0227	1.49e-06	0.0060
Valor Residual Estándar	63.75		
Múltiple R-Squared	0.9559		
Ajuste R-Squared	0.9521		
P-Value	2.57e-16		

Tabla 13.-Resumen del modelo. Modelo Simplificado 2

En la Tabla 14 se muestran los intervalos de confianza al 95% para cada uno de los coeficientes del modelo lineal.

INTERVALOS DE CONFIANZA		
	Extremo Inferior	Extremo Superior
Intercepto	-1.043502e+02	-8.650329215
Beta de Camas	5.143387e-01	1.003035823
Beta de Urgencia	6.070741e-04	0.003229209

Tabla 14.-Intervalos de confianza. Modelo Simplificado 2

IC95% β_1 = (5.143387e-01, 1.003035823)

IC95% β_2 = (6.070741e-04, 0.003229209)

A continuación, estudiaremos si cada coeficiente puede ser o no igual a cero. En la Tabla 13 aparecen los p-valores asociados a cada par de contraste-coeficiente. Para el caso del intercepto, el contraste de hipótesis a resolver sería el siguiente:

H0: $\beta_0 = 0$

H1: $\beta_0 \neq 0$

En este caso el p-valor es igual a 0.0227, como es inferior a 0.05, podemos rechazar H0. Por tanto, nos indica que en presencia de las demás variables ésta es relevante.

p-valor β_1 = 1.49e-06 < 0,05 Rechazamos H0 a favor de H1.

p-valor β_2 = 0.0060 < 0,05 Rechazamos H0 a favor de H1.

FASE 3: BONDAD DEL AJUSTE

Verificamos si dicho modelo proporciona un buen ajuste a la hora de predecir la variable respuesta, Médicos. Cuantificaremos esta viabilidad de diversos modos, en concreto tres:

1.-Error estándar residual: hacemos la medición de la bondad del ajuste relativa a la escala de medida utilizada. De los datos de la Tabla 13 obtenemos el valor residual estándar del modelo, el cual tiene un valor de 63.75. Para determinar si éste valor es grande o pequeño realizamos el cálculo del coeficiente de variación $cv = 100 * (63.75 / \text{la media de Médicos}) \%$ lo

que da un resultado de $cv=18.24$ ($> 10\%$ recomendado por el instituto nacional de estadística). No aprueba el primer criterio del error estándar residual.

2-Tabla de Anova: este criterio cuantifica cuánta de la variabilidad contenida en los datos ha conseguido ser explicada por dicho modelo.

Damos respuesta al siguiente contraste de hipótesis:

$H_0: \beta_1 = \beta_2 = \beta_3 = 0$

H_1 : lo contrario

El p-valor de nuestro modelo es $2.57e-16$ que como es menor que $0,05$, rechazo H_0 y asumimos que el modelo es bueno al igual que en el modelo global. Las variables explicativas explican suficientemente bien la variable respuesta a través del modelo lineal propuesto.

3.-Coeficiente de determinación:

El coeficiente de R_2 es 0.9559 ($>65\%$), explica el $95,59\%$ de la variabilidad total de la variable Médicos, por lo que es un modelo muy bueno.

El coeficiente ajustado de R_2 tiene un valor similar de 0.9521 por lo que consideramos bueno el valor de coeficiente de R_2 . Por lo tanto, damos por válido el valor del coeficiente R_2 original.

FASE 4. DIAGNÓSTICO DEL MODELO:

En la cuarta fase vamos a verificar si el modelo satisface las hipótesis básicas del modelo de regresión lineal para los errores:

- Normalidad
- Varianza constante (Homocedasticidad)
- Incorrelación
- Media Cero

1.-Normalidad: haremos un estudio gráfico viendo el gráfico QQPLOT donde representamos los residuos ordenados versus los cuantiles correspondientes de una normal y el histograma como en el modelo anterior.

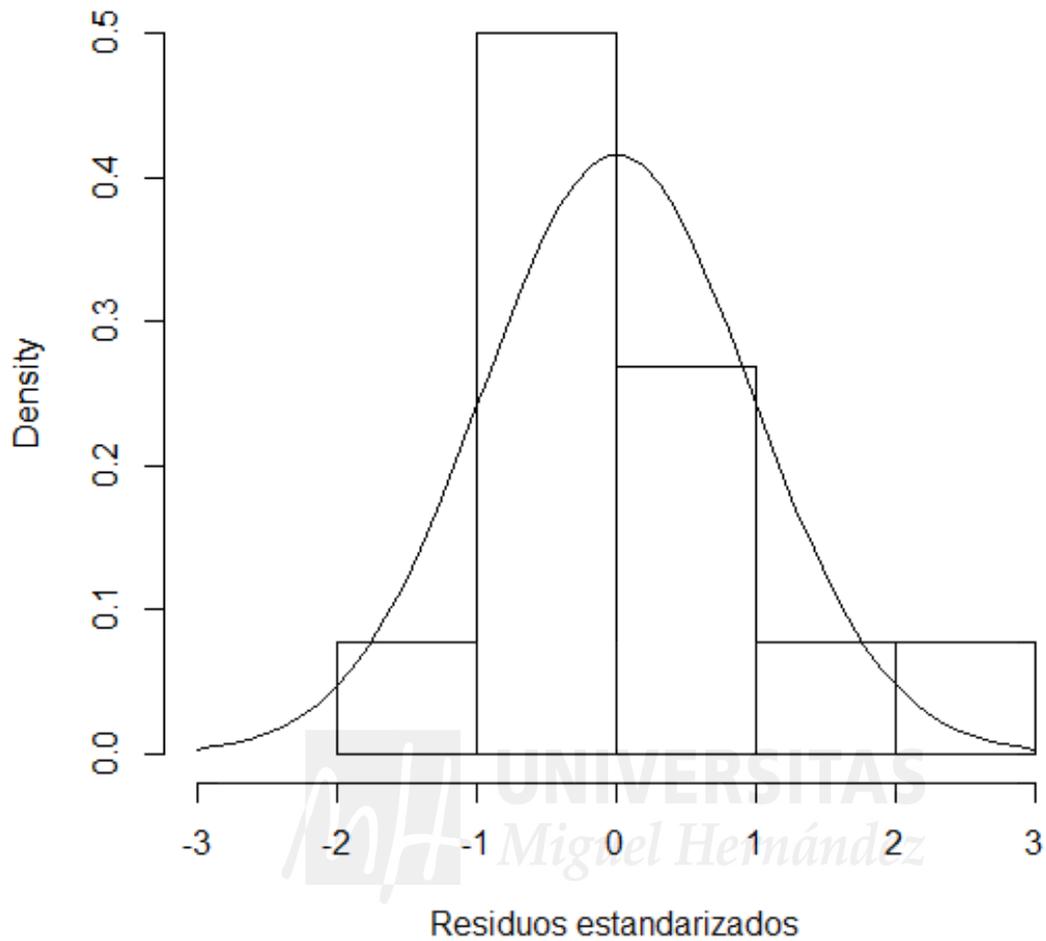


Figura 12.- Histograma Modelo Simplificado

La Figura 12 no se asemeja a la campana de Gauss, no hay simetría.

En el gráfico qqplot que se muestra en la Figura 13, la mayoría de los puntos están sobre la línea recta.

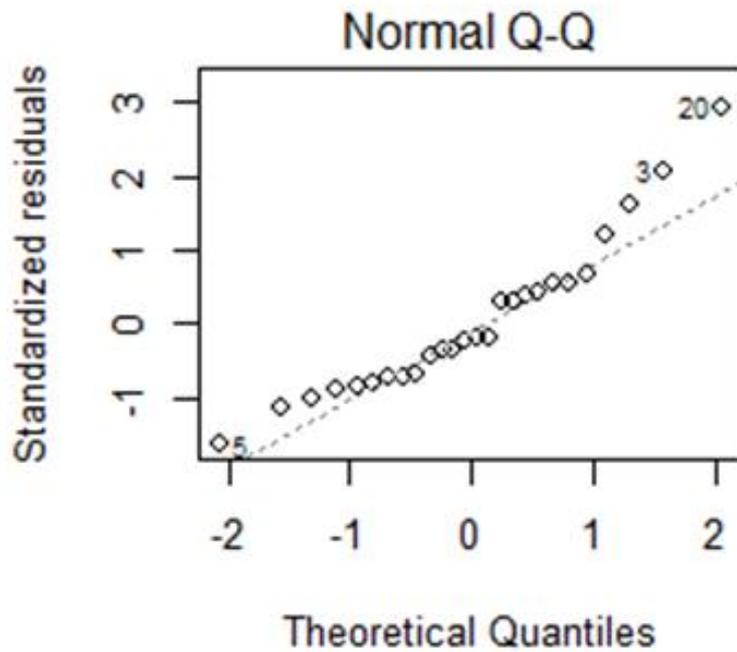


Figura 13.- Gráfico QQPLOT . Modelo Simplificado 2

La normalidad también será contrastada a través del contraste de hipótesis de Shapiro-Wilk:

H0: $e \sim \text{Normal}$

H1: lo contrario

En este caso, obtenemos un p-valor de $0.05703 > 0,05$, por lo que no podemos rechazar la hipótesis de normalidad y con lo cual asumiré que los errores son normales. En definitiva, esta hipótesis podemos asumir que se verifica.

2 -Homocedasticidad: (hipótesis de varianza constante). Realizaremos un análisis gráfico sobre el gráfico de residuos versus valores predichos de la variable respuesta Médicos y numéricamente resolveremos el contraste de hipótesis de Breush-Pagan.

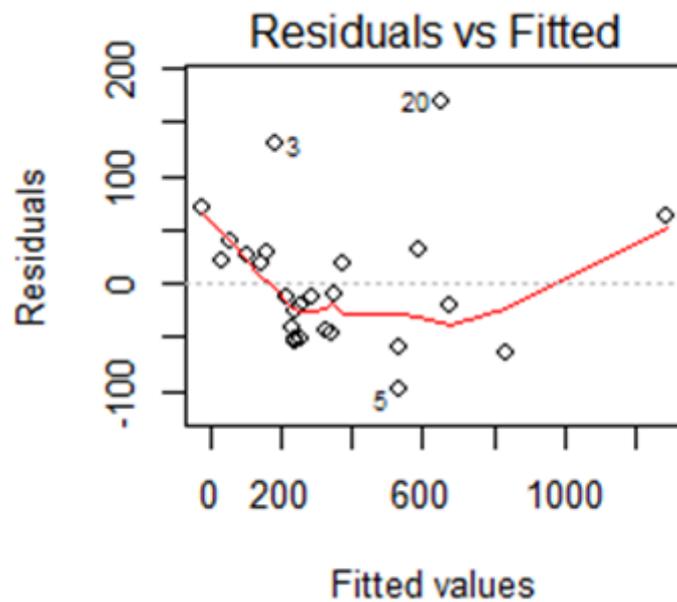


Figura 14.- Residuos frente a valores predichos. Modelo Simplificado 2.

Resolvemos a continuación el contraste de hipótesis debido a Breusch y Pagan:

H0: Homocedasticidad

H1: Heterocedasticidad

Obtenemos en concreto un p-valor de $0.3388 > 0.05$. No podemos rechazar H0. Por lo tanto, asumiremos la Homocedasticidad, es decir, que la varianza es constante. Esta hipótesis se verifica.

3.-Incorrelación: vamos a observar si los valores de unas variables afectan a las de otras, para comprobarlo construiremos el gráfico del residuo i frente anterior $i-1$ y realizaremos el contraste de hipótesis de Durbin y Watson para hacer la comprobación numérica.

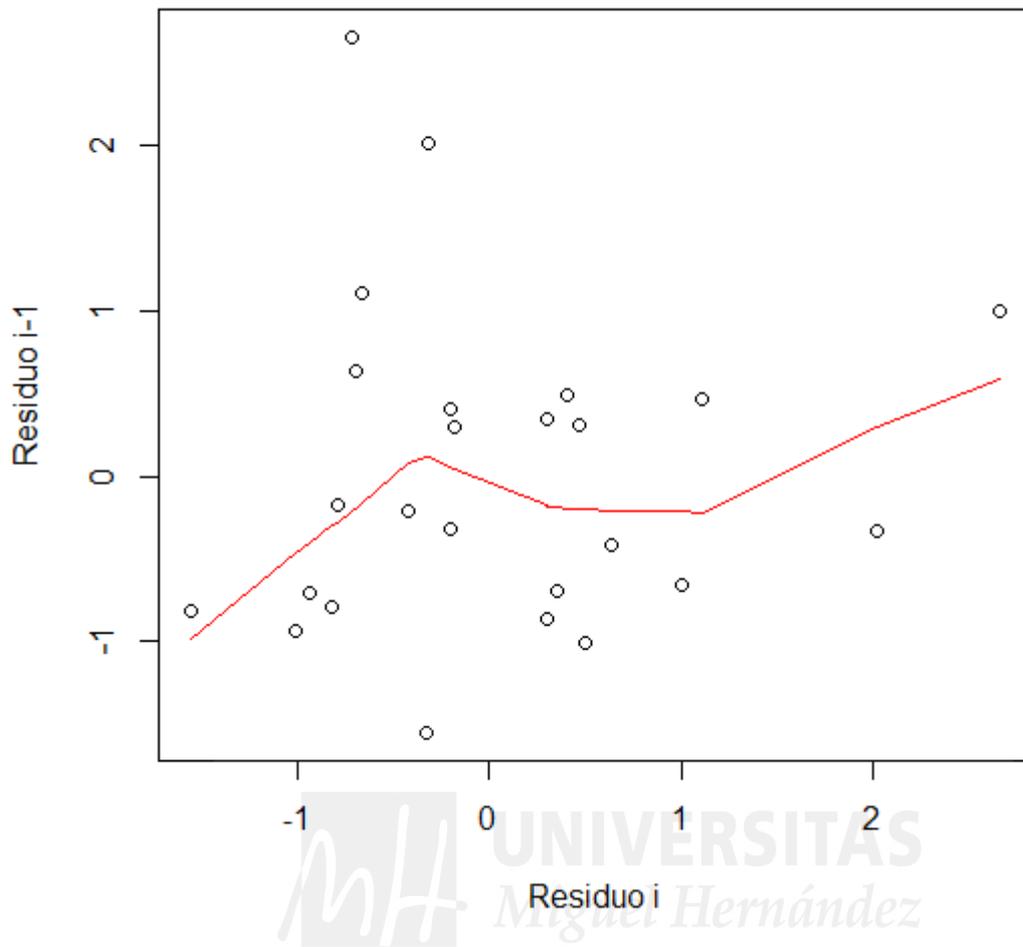


Figura 15.- Gráfico Incorrelación. Modelo Simplificado 2.

En este sentido, el test que debemos de usar se debe a Durbin y Watson:

$H_0: \rho = 0$ (incorrelación)

$H_1: \rho \neq 0$

En este caso, obtendremos un p -valor = 0.6413 > 0.05 No podemos rechazar H_0 , lo que indica que no se puede rechazar la hipótesis de Incorrelación el modelo es bueno para este criterio: hay Incorrelación.

4.- Media Cero: Esta hipótesis siempre se cumple porque usamos el método de máxima verosimilitud, que nos asegura que la media de los residuos siempre es cero.

6. CONCLUSIONES

En este apartado trataremos de exponer las principales conclusiones a las que hemos llegado después de realizar un estudio econométrico de la base de datos “Hospitales en la Comunidad Valenciana”. El objetivo desde el principio ha sido observar la relación existente, entre número de Médicos en los hospitales de la Comunidad Valenciana y otras variables relacionadas con el sector sanitario.

Con el fin de obtener unas conclusiones lo más fidedignas y realistas, hemos realizado tres modelos diferentes pero a la vez muy relacionados: el modelo lineal completo y dos modelos simplificados (con menor número de variables explicativas, en nuestro caso hemos pasado de cuatro variables a tres variables y de tres a dos). Para nuestro estudio, no ha sido necesario desarrollar el modelo de Transformación de la variable respuesta ya que hemos considerado que tiene un R^2 aceptablemente bueno y que además cumple con todas las hipótesis del modelo.

El procedimiento empleado para el análisis de los tres modelos ha sido el mismo, primero hemos realizado un análisis preliminar, a continuación procedemos a ajustar el modelo, comprobamos la bondad de los ajustes y por último verificamos que se cumplen las hipótesis asumidas por el modelo.

En el análisis preliminar, hemos analizado gráfica y numéricamente las relaciones existentes entre las variables explicativas y la variable respuesta. Según los gráficos de dispersión, no se observa una alta relación lineal entre la variable respuesta Médicos y el resto de variables explicativas. Con el estudio de las correlaciones simples, sin embargo, podemos comprobar que existen unas correlaciones simples altísimas entre número de médicos y Camas e Ingresos y Urgencias. Con el estudio de las correlaciones parciales confirmamos que la variable más importante es el número de camas hospitalarias, sin embargo otras que a priori parecían ser relevantes, después no lo han sido. Esta situación es muy similar en ambos modelos estudiados, siempre que tengamos en cuenta que en el modelo simplificado hemos suprimido la variable primconsult y después la variable ingresos

Una vez realizado el análisis preliminar, realizamos el ajuste del modelo para estimar los valores de los cuatro parámetros del modelo completo, los tres y posteriormente dos parámetros de los modelos simplificados, además calculamos intervalos de confianza al 95%. En el primer modelo concluimos que el coeficiente correspondiente a primconsult puede ser nulo, mientras que el resto de parámetros Camas, Ingresos y Urgencias tendrían información suficientemente útil para explicar la variable respuesta. En el modelo simplificado concluimos que el parámetro Ingresos puede ser nulo, mientras que Camas y

Urgencias tendrían información estadísticamente significativamente al rechazar la hipótesis de $H_0: \beta_i = 0$ y son útiles para explicar el modelo.

En ambos casos hemos realizado la bondad del ajuste para comprobar si se verifica que el modelo es bueno para explicar la variable respuesta. Para lograr cuantificar la bondad del ajuste, hemos utilizado tres criterios: Error Estándar Residual, la Tabla de Anova y el Coeficiente de Determinación. Tanto en el modelo completo como en los simplificados, se cumplen los criterios de la Tabla de Anova y el Coeficiente de Determinación, sin embargo no el del Error Estándar Residual. En los tres modelos las pruebas gráficas parecen no terminan de aclarar las hipótesis básicas, sin embargo, al realizar los contrastes de hipótesis, éstas se cumplen y por tanto podemos aceptar que los modelos son bastante buenos.

Aplicando el método de máxima verosimilitud, manteniendo constante el resto de variables, según el modelo simplificado 2, si aumenta en 1 unidades el número de camas, la plantilla de médicos en los hospitales para que el servicio ofrecido fuese similar al actual, debería aumentar en $7.587e-01$. En el caso de que aumentásemos en 1 unidad los pacientes que acuden a urgencias y mantuviésemos constante el resto de variables, en este caso, habría que aumentar en el número de médicos en $1.918e-03$.

Podemos concluir que los modelos son suficientemente buenos y existe cierta relación lineal entre las variables explicativas y el número de médicos en los hospitales de la Comunidad Valenciana. Con los modelos simplificados tratamos de reducir y eliminar las variables menos útiles para hacer un modelo más sencillo, y como hemos comprobado también han cumplido los objetivos previstos. No hemos considerado necesario realizar una transformación de la variable respuesta ya que el modelo simplificado tiene un R^2 muy bueno, 0.9559 (>65%) y cumple con todas las hipótesis del modelo.

7. BIBLIOGRAFÍA:

Aparicio, J., Martínez Mayoral, M., & Morales, J. (s.f.). Modelos Lineales Aplicados en R. (2004) (C. I. Hernández, Ed.) Dto. Estadística, Matemáticas e Informática.

Aparicio, J. (s.f.). ECONOMETRÍA. Obtenido de <http://umh3067.edu.umh.es/material/videos/>

Ministerio de Sanidad, Servicios Sociales e Igualdad. Sistema Nacional de Salud. España 2012. Madrid; 2012. Disponible en:
www.msssi.gob.es

Catálogo Nacional de Hospitales 2013, actualizado a 31 de diciembre de 2012. Ministerio de Sanidad, Servicios Sociales e Igualdad y las Consejerías de Sanidad de las Comunidades Autónomas, el Ministerio de Defensa, los órganos competentes de las Ciudades Autónomas de Ceuta y Melilla.

Catálogo general de publicaciones oficiales: <http://publicacionesoficiales.boe.es/>

Departamento de Estadísticas e Información de Salud. DEIS, Ministerio de Salud, Chile. Notas Técnicas. www.deis.cl

<http://vozpopuli.com/actualidad/30615-la-sanidad-espanola-la-quinta-mas-eficiente-del-mundo-segun-bloomberg>

<http://www.bloomberg.com/visual-data/best-and-worst/most-efficient-health-care-countries>

BOE» núm. 312, de 30 de diciembre de 1987, páginas 38171 a 38176 (6 págs.) www.boe.es

MAPFRE, Bienestar, Hábitos de vida saludable.
www.mapfre.es/salud/es/cinformativo/importancia-cuidar-salud.shtml

ESTADÍSTICA DE GASTO SANITARIO PÚBLICO 2012: Principales resultados. (Edición abril 2014) Ministerio de Sanidad, Servicios Sociales e Igualdad.

<http://www.elmundo.es/comunidad-valenciana/2013/11/19/528b802b61fd3da93f8b457f.html>
(19/11/2013)

Barómetro de la Sanidad Privada. Instituto para el desarrollo y la integración de la Sanidad.
www.fundacionidis.com/

«BOE» núm. 102, de 29 de abril de 1986, páginas 15207 a 15224 (18 págs.) En 1986 mediante la ley General de Sanidad

<http://www.bloomberg.com/visual-data/best-and-worst/most-efficient-health-care-2014-countries>