

Actividades mercantiles y profesionales en la provincia de Almería: el impacto del ferrocarril en 1931.

**M. Pilar Fernández Sánchez, Luis González Ruiz y Gregorio Núñez
(Universidad de Granada y Grupo de estudios históricos sobre la empresa).¹**

Resumen:

En este texto estudiamos la composición funcional y la estructura territorial de la oferta de bienes y servicios ofrecidos a los consumidores y a los empresarios de la provincia en la etapa inmediatamente anterior a la segunda república española.

Para ello estamos utilizando una fuente que nunca hasta ahora había sido utilizada de forma completa y sistemática: el Anuario General de España, en este caso la edición de 1931, una publicación que durante muchas décadas sucesivas imprimió la casa editorial Bailly-Bailliere y Riera. En ella aparecen largos listados de profesionales y empresarios almerienses ordenados por las actividades que ejercían y clasificados por municipios y provincias. Es una fuente de gran riqueza y de cierta complejidad cuyo aprovechamiento nos ha obligado a establecer métodos rigurosos de recopilación y análisis. Unos métodos que ahora pretendemos contrastar y evaluar con la intención de avanzar en la metodología y de ampliar el estudio al resto de España.

En particular deseamos evaluar aquí si será posible aplicar modelos econométricos simples a los datos estadísticos resultantes de nuestro estudio, para lo cual en este texto nos hemos preguntado si el conjunto de actividades empresariales y profesionales ofrecidos en cada localidad presenta características diferenciadas según nos encontremos en municipios servidos o no por el ferrocarril. Para ello hemos aplicado las pruebas de Kolmogorov-Smirnov (K-S) y de Mann-Whitney (M-W) para establecer si los municipios agrupados en los dos grupos deben ser considerados como dos poblaciones diferentes a los efectos indicados.

En una primera aproximación, considerando los datos "brutos", los resultados de la prueba no resultaban concluyentes de modo que no se podía rechazar la hipótesis nula de igualdad de poblaciones ni sostener que se tratara de dos grupos diferenciados. No obstante se ha apreciado también que si avanzamos en el análisis y tomamos en consideración grupos de actividades más depurados y homogéneos, si aislamos grupos de actividades con influencias previsiblemente dispares o incluso contrarias, el nivel de significación de las pruebas sube hasta poder rechazarse la hipótesis nula con determinadas condiciones. Sobre todo resulta claro que tomando en cuenta el efecto tamaño de los municipios los resultados aparecen ya más contrastados y significativos. Parece claro que el resultado, como el análisis mismo, ha de resultar más complejo de lo esperado.

Entendemos que dicha conclusión nos invita a seguir adelante para definir más y mejor los grupos de las actividades económicas, reclasificar los casos individuales con mejor criterio y determinar después cuales y en qué medida resultan ser más significativos respecto a la influencia del ferrocarril y de otras variables económicas importantes.

¹ M. Pilar Fernández Sánchez, Orcid ID 0000-0002-5834-1452. Luis González Ruiz, Orcid ID 0000-0002-9469-5442. Gregorio Núñez Romero-Balmas, Orcid ID 0000-0003-1793-5817. Los autores agradecen al profesor don Agustín Hernández Bastida, de la Universidad de Granada, su colaboración y acertadas sugerencias.

1. Planteamiento y objetivos

En las páginas que siguen pretendemos estudiar la configuración organizativa y territorial del tejido empresarial de la provincia de Almería en el marco de un estudio que venimos realizando sobre los profesionales y empresarios andaluces antes de la guerra civil. Para ello estamos vaciando publicaciones mercantiles de diferentes años del primer tercio del siglo XX y, más concretamente, los procedentes del *Anuario General de España*, publicación de la que la casa Bailly-Baillière y Riera publicó diferentes ediciones a lo largo de más de cincuenta años. Las posibilidades que encierra esta fuente para el estudio de las profesiones mercantiles y de la estructura comercial de España ya han quedado descritas en otras publicaciones.²

En las páginas que siguen utilizaremos los datos del *Anuario General de España* correspondiente a la edición de 1931 (en adelante AGE-1931). El conjunto nos proporciona un largo listado de los profesionales y empresarios inscritos en cada una de las actividades económicas que se ofrecían en cada municipio de la provincia. En realidad los listados son varios en cada término municipal: un listado alfabético de titulares para cada una de las actividades desarrolladas, de modo que un mismo titular puede aparecer en uno o varios segmentos de actividad. Dichos segmentos o grupos recogen la práctica entonces habitual y deben considerarse por tanto como un dato importante. Pero, a cambio, dado que no se aplica con un criterio sistemático de intención estadística, supone un factor de distorsión que es preciso corregir por medio de un complejo proceso de reclasificación cuya aplicación empieza a quedar ya claramente establecida con la práctica.

2. Variables y métodos.

Como ya se ha dicho, la fuente nos proporciona la relación nominativa de los profesionales y empresarios domiciliados en cada municipio, ordenados conforme a las diferentes actividades que anunciaban realizar. Esto nos obliga a establecer con claridad

² v. González Ruiz, Luis y Núñez Romero-Balmas, Gregorio. «Estadísticas para la historia: Uso estadístico de fuentes mercantiles». Herrerías Velasco, José Manuel y Callejón Céspedes, José, Eds. *Investigaciones en métodos cuantitativos para la economía y la empresa. Homenaje al profesor Rafael Herrerías Pleguezuelo*. Granada: Editorial Universidad de Granada; 2016; pp. 319-334. [<http://hdl.handle.net/10481/43516>]. Sobre un sector y una provincia concretas v. Gonzalez Ruiz, Luis y Núñez Romero-Balmas, Gregorio. «La fragua del turismo: Empresas y actividades turísticas en la Almería de 1931. en: González Ruiz, Luis; López-Burgos del Barrio, María Antonia, and Núñez Romero-Balmas, Gregorio. *La hostelería almeriense antes del turismo de masas*. Granada: Godel - Grupo de estudios históricos sobre la Empresa; 2015; pp. 37-111. [<http://hdl.handle.net/10481/38949>].

dos conceptos diferentes que dan lugar a dos variables distintas. Por un lado debemos considerar a los «titulares» de los diferentes negocios y/o profesiones, personas naturales o jurídicas que aparecen identificados con nombre y domicilio; unos titulares cuyo número total alcanza a 11.041 profesionales o empresarios en toda la provincia. Dichos titulares individuales son susceptibles de ser clasificados y agregados de diferentes modos, lo que nos proporcionan una primera variable estadística que indica el número de titulares (N_{tit} = suma de todos los profesionales y empresarios de cada localidad).

Las descripciones de los titulares en el AGE-1931 son mucho más precisas en las capitales de provincia que en los pueblos. En las primeras aparecen por lo general el nombre y dos apellidos, así como el domicilio o domicilios precisos referidos a las distintas actividades que ejercían. En los otros no suelen aparecer los domicilios y el nombre se ve a menudo reducido a un sólo apellido. Los compiladores de la fuente entendían sin duda que dicha precisión era suficiente para identificar en cada caso a los habitantes de pequeñas poblaciones. Nosotros suponemos que sin duda debía ser suficiente para evitar errores y duplicaciones, de modo que nuestra variable queda definida por medio de la concatenación de nombre, con uno o dos apellidos, y localidad.

Por el otro lado hay que distinguir las «actividades» que ofrecían dichos titulares; a menudo sólo una, más o menos o precisa, de modo que 8.236 (74,59%) de ellos aparecen como ejercientes de una sola actividad. Pero 2.805 (25,41%) de los titulares restantes anunciaban operar simultáneamente en dos o más segmentos diferentes, hasta un máximo de veintitrés.

Con todas las críticas que se pueden hacer a la hora de reclasificar las actividades en la base de datos utilizaremos los criterios del INE en la Clasificación Nacional de Actividades Económicas de 2009 [CNAE 2009]. Es claro que la clasificación de actividades económicas correspondientes a épocas y territorios muy diferentes encierra problemas específicos y aporta al análisis dificultades adicionales que no podemos tratar en estas páginas. Sólo insistiremos en que el amplio catálogo de actividades mercantiles «de época» que nos proporciona el AGE 1931 constituye un rico complemento terminológico para el estudio histórico de la evolución empresarial desde el punto de vista de la lexicografía.

Es lógico esperar que la proximidad o no al ferrocarril de una determinada comarca debería influir sobre el paisaje empresarial y profesional de la zona. Especialmente en una época en que la movilidad de personas y bienes era costosa y limitada pues ni la automoción ni la inversión en carreteras estaba aún ampliamente desarrollada.

Por tanto, parece relevante analizar si esa esperada influencia se manifestaba realmente en la configuración real del sistema empresarial de la provincia y, en caso positivo, qué diferencias significativas quedan en evidencia considerando la diversidad de sectores de actividad económica de la época. Igualmente habrá que determinar el signo de dicha influencia, que puede resultar menos obvio de lo cabría esperar.

En este trabajo nos proponemos realizar una primera aproximación a dicho tema, de modo que su objetivo será verificar si, en la estructura de la actividad mercantil entre los municipios cercanos al ferrocarril y los que quedaron alejados de él, aparecen diferencias significativas en la configuración del tejido empresarial. Igualmente limitado es el territorio y la época objeto de análisis: la provincia de Almería en 1931.

De confirmarse la hipótesis principal, intentaremos determinar también los perfiles de especialización de los diferentes municipios tomados en consideración atendiendo a esa característica.

Arrancaremos de la variable número de actividades (*NAct*) presentes en cada municipio, clasificadas por su sector económico moderno, y esperamos descubrir que, en lo que al tejido empresarial se refiere, los municipios próximos al ferrocarril constituyen una población diferenciada respecto a los municipios sin ferrocarril. Todo ello usando las variables que nos proporciona nuestra base de datos sobre profesionales y empresarios de la provincia de Almería.

Un primer paso es determinar qué entendemos por municipio con ferrocarril. Se trata de una decisión subjetiva en la que se ha tenido en cuenta una distancia que parece adecuada si atendemos a desplazamientos realizados a pie o en carro tirado por mulas. El criterio adoptado es que un municipio cuenta con ferrocarril si la estación está a menos de diez kilómetros de la población central del municipio. Con este criterio, el 55,3% de los municipios de la provincia de Almería no tenía acceso al ferrocarril. En la actualidad su número es muy superior pues la línea norte, entre Guadix y Almendricos, hace tiempo que fue suprimida.

De esta forma, el conjunto total de municipios de Almería lo vamos a segmentar en dos subpoblaciones y nos proponemos establecer si dichas subpoblaciones pueden ser consideradas diferentes entre sí. Es decir, ver si el comportamiento de la variable *NAct* en cada uno de los sectores establecidos puede ser considerada diferente atendiendo al hecho de tener o no ferrocarril en su entorno inmediato.

A poco que se avanza se aprecia que al trabajar con municipios es necesario tomar en cuenta el tamaño del mismo. No es lo igual comparar municipios grandes con municipios pequeños porque ese tamaño puede afectar al análisis. Para eliminar ese efecto tamaño vamos a acudir a otras variables que dan una idea del tamaño de los municipios: *Superficie del término* y su *Población de hecho*.

Seguidamente construimos otras dos variables que son, para cada municipio, la densidad de actividades ($NAct/km^2$) y actividades per capita ($NAct/hab$), variables que serán las que tomemos en consideración para el análisis estadístico. Se comprueba que, como era de esperar, existía una correlación significativa y positiva entre *NAct* en cada sector y las variables *Superficie* y *Población* de los términos respectivos, indicando que a más superficie o mayor población mayor es *NAct*.

Se han analizado cada uno de los ciento tres municipios de la época y se dispone de datos para diferentes sectores (en columnas) y municipios (en filas) en una tabla de doble entrada (ver Tabla Anexo 2) de la provincia de Almería. Para la variable $NAct/Km^2$ se ha eliminado el término de Garrucha que presenta unas características muy particulares que obligan a estudiarla de forma independiente.³

La metodología estadística que aplicaremos son contrastes no paramétricos de independencia y de diferencia de medias, así como el cálculo de medidas de análisis regional, más concretamente cocientes de localización. Para los detalles técnicos puede consultarse el Anexo.

3. Análisis de los datos

Iniciamos el estudio viendo qué grupos de actividades conforme al criterio CNAE 2009 estaban presentes en todos los municipios.⁴ Más precisamente vamos a ver qué porcentaje de municipios cuentan con un determinado sector de actividades. En general, sin distinguir entre municipios con y sin tren, hay cuatro sectores de actividad económica que figuraban en todos los municipios: la Industria Manufacturera, el Comercio, la Administración Pública y Defensa y, por último, los clasificados como Otros Servicios. Otros siete sectores de actividad figuran en más del noventa por ciento de los municipios: la Educación, la Agricultura, la Hostelería, la Construcción, el Transporte y Almacenamiento y las Actividades Sanitarias. En la práctica vienen a ser amplios cajones de sastre rellenos de actividades de efectos estadísticos en ocasiones contradictorios.

Actividades de los grupos del Suministro de Energía Eléctrica y las Actividades Profesionales aparecían en poco más de la mitad de los municipios (56 y 55 por ciento, respectivamente) mientras que el resto de sectores de actividades económicas no llegaban a estar presentes en la mitad de los municipios.

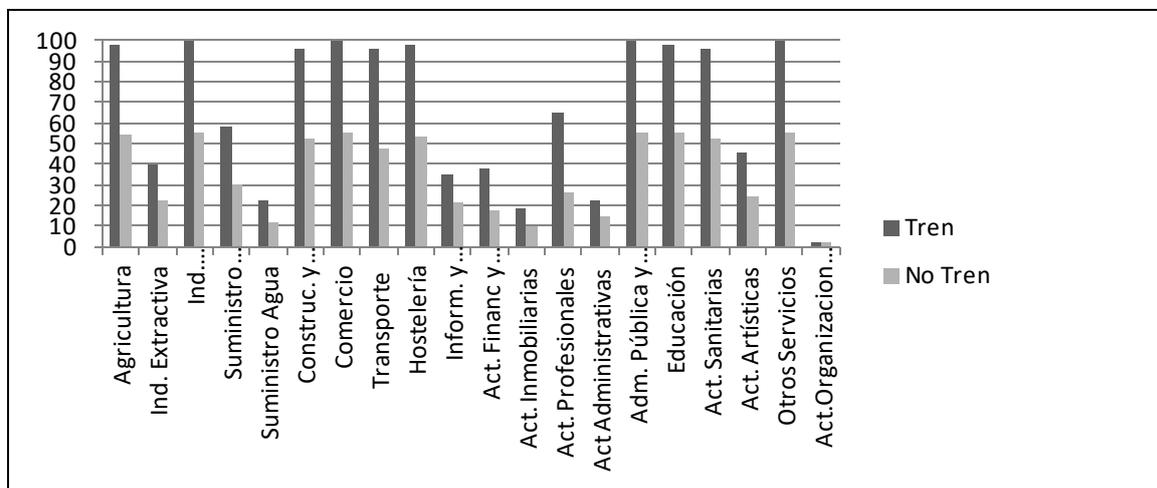
El Gráfico 1 muestra la distribución del porcentaje de municipios con un determinado sector presente, según si hay tren o no en dicho municipio.

Se observa que los porcentajes calculados difieren para los dos grupos de municipios y se comprueba mediante las Pruebas de Kolmogorov-Smirnov (K-S) y de Mann-Whitney (M-W) que esas diferencias son significativas (ver Anexo 7.2).

³ En términos superficiales el municipio de Garrucha es prácticamente irrelevante. Su superficie, de 0,37 km² había sido segregada de los términos municipales vecinos para favorecer las actividades portuarias. Por ello las diferentes variables demográficas y empresariales alcanzan en ella magnitudes propias de una gran ciudad y hace que deba ser estudiada de forma independiente

⁴ v. Pardo Merino, A. y Ruíz Díaz, M. (2010). *Análisis de datos con SPSS 13Base*. McGraw Hill.

Grafico 1
Frecuencia (%) por sector de actividad y tren de los municipios de la provincia de Almería en 1931.



A continuación, aplicamos esas pruebas a las tres variables definidas en el epígrafe anterior en función de la característica ferrocarril en el municipio (v. Tabla 1). La primera prueba contrasta la igualdad de poblaciones y la segunda la igualdad de medias. Presentamos los p-valores asociados a dichas pruebas. Se ha sombreado de gris oscuro aquellas situaciones en las que se rechazaría la hipótesis nula (de igualdad de poblaciones o de medias, respectivamente) al 5% y en gris claro las que se rechazarían al 10%.

Como puede observarse, para la variable *NAct* (que no toma en cuenta el tamaño del municipio), los p-valores son elevados con lo que no puede rechazarse la hipótesis nula de homogeneidad entre los dos grupos de municipios. Esta evidencia refuerza nuestra idea que sostiene que los datos brutos por municipios, sin depurar el efecto tamaño de la localidad, no son adecuados para el estudio.

Sin embargo, para las otras dos variables sí que encontramos diferencias en algunos sectores. En la variable ponderada por la superficie, hay nada menos que siete sectores de actividad (valores sombreados en gris obscuro) cuyo p-valor (para K-S) nos permite rechazar la hipótesis nula de homogeneidad de las dos poblaciones y, por tanto, podemos afirmar que la densidad de actividades ($NAct/Km^2$) no es el mismo para aquellos municipios con ferrocarril que para los que no lo tienen.

Los sectores para los que se han encontrado dichas diferencias son: la Agricultura, la Industria Manufacturera, el Comercio, el Transporte y Almacenamiento, la Hostelería, las Actividades Profesionales y los clasificados como Otros Servicios.

Tabla 1

<i>p-valores</i> pruebas Kolmogorov-Smirnov y Mann-Whitney para las variables <i>NAct</i> , <i>NAct/Km²</i> y <i>NAct/hab</i> por sectores de actividad.						
SECTOR	Número de actividades		Densidad de actividades		Actividades per capita	
	<i>NAct</i>		<i>NAct/Km²</i>		<i>NAct/hab</i>	
	K-S	M-W	K-S	M-W	K-S	M-W
Agricultura	0,238	0,115	0,003	0,002	0,565	0,310
Industria extractiva	1,000	0,749	0,998	0,963	0,856	0,647
Industria manufacturera	0,665	0,273	0,001	0,000	0,416	0,105
Suministro de energía eléctrica	0,995	0,524	0,066	0,038	0,322	0,205
Suministro Agua	1,000	0,665	0,967	0,679	0,947	0,659
Construcción	0,353	0,392	0,267	0,053	0,888	0,816
Comercio	0,121	0,250	0,002	0,000	0,016	0,041
Transporte y almacenamiento	0,336	0,286	0,006	0,002	0,112	0,158
Hostelería	0,643	0,528	0,009	0,002	0,379	0,392
Información y comunicaciones	1,000	0,893	0,997	0,702	1,000	0,890
Actividades Financieras y seguros	0,999	0,659	0,759	0,291	0,998	0,575
Actividades Inmobiliarias	1,000	0,825	1,000	0,812	1,000	0,833
Actividades Profesionales	0,392	0,200	0,018	0,011	0,132	0,081
Actividades Administrativas	1,000	0,990	0,752	0,782	0,924	0,965
Administración pública y defensa	1,000	0,811	0,073	0,018	0,884	0,786
Educación	0,707	0,421	0,090	0,020	0,848	0,715
Actividades Sanitarias	0,851	0,803	0,141	0,033	0,525	0,725
Actividades Artísticas	1,000	0,720	0,342	0,242	0,982	0,590
Otros Servicios	0,550	0,234	0,002	0,000	0,025	0,012

Hay tres sectores en los que el p-valor está entre el 5% y el 10% y por eso lo señalamos: el Suministro Energía Eléctrica, la Administración Pública y Defensa y la Educación. Dependerá del criterio del analista decidir si se rechaza o no la hipótesis nula en estos casos.

Para la variable actividades per capita (*NAct/hab*) sólo se han encontrado diferencias significativas en dos sectores, Comercio y Otros Servicios.

A continuación, recogemos los principales resultados descriptivos referentes a aquellos sectores en los que se ha encontrado diferencias significativas entre las dos subpoblaciones (v. Tabla 2).

Tabla 2

Descriptivos para las variables $NAct/Km^2$ y $NAct/hab$ en los sectores con diferencias entre subpoblaciones.							
			Media	Mediana	C.Var	Asimetría	Curtosis
$NAct/Km^2$	Agricultura	Tren	0,560	0,357	1,187	2,963	10,023
		No tren	0,288	0,238	1,213	2,379	5,829
	Industria Manufacturera	Tren	0,567	0,496	0,695	1,637	3,286
		No tren	0,341	0,227	1,211	5,620	37,308
	Suministro Energía E.	Tren	0,022	0,012	1,444	1,511	2,311
		No tren	0,009	0,003	1,254	1,727	2,666
	Comercio	Tren	1,138	0,849	0,764	1,472	1,592
		No tren	0,625	0,403	0,864	1,838	4,076
	Transporte	Tren	0,075	0,056	0,875	1,702	3,080
		No tren	0,046	0,028	1,106	1,747	2,508
	Hostelería	Tren	0,211	0,153	0,694	0,799	-0,563
		No tren	0,143	0,087	1,108	3,011	11,617
	Actividades profesionales	Tren	0,059	0,026	1,621	2,548	6,253
		No tren	0,023	0,000	1,813	2,612	7,084
	Ad. Pública y defensa	Tren	0,189	0,138	0,945	2,235	5,334
		No tren	0,124	0,097	0,995	3,190	13,906
	Educación	Tren	0,102	0,085	0,820	2,081	5,308
		No tren	0,067	0,056	0,589	0,942	0,250
Otros Servicios	Tren	0,198	0,155	0,717	1,828	4,871	
	No tren	0,104	0,075	0,729	2,192	8,065	
$NAct/hab$	Comercio	Tren	0,020	0,019	0,420	0,969	1,008
		No tren	0,018	0,014	0,635	1,590	3,630
	Otros Servicios	Tren	0,004	0,004	1,221	0,408	-0,620
		No tren	0,003	0,003	1,471	1,193	1,473

Observaciones parecidas pueden hacerse en el caso de la variable $NAct/hab$ en cuanto a la media y a la mediana, ya que ambas son mayores en los municipios con ferrocarril. En este caso hay, además, mayor dispersión, mayor asimetría y mayor curtosis en el caso de los municipios con tren.

Tabla 3
p-valores Pruebas Kolmogorov-Smirnov y Mann-Whitney para las variables $NAct$, $NAct/Km^2$ y $NAct/hab$ en el sector Comercio.

	$NAct$ Comercio		$NAct$ Comercio/ Km^2		$NAct$ Comercio/ hab	
	Mayoristas	Minorista	Mayorista	Minorista	Mayorista	Minorista
K-S	0,137	0,199	0,005	0,015	0,089	0,112
M-W	0,353	0,257	0,000	0,003	0,107	0,085

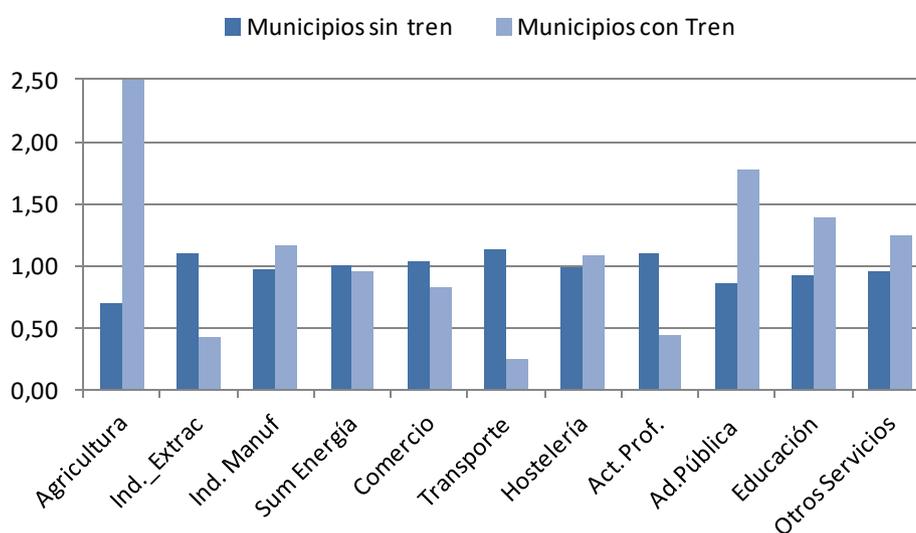
Visto que el Comercio aparece en las dos variables como un sector donde hay diferencias entre las dos subpoblaciones, vamos a desagregar dicho grupo para estudiar si se puede ver de dónde proceden esas diferencias. Para ello vamos a distinguir entre Comercio al por mayor y Comercio al por menor y volveremos a aplicar el análisis anterior. Por tanto consideramos seis nuevas variables y vemos si existen para ellas diferencias entre los municipios con y sin ferrocarril.

Como puede apreciarse y era de esperar no hay diferencia entre el comercio mayorista y el minorista cuando se toman en consideración los datos brutos. Pero son claramente poblaciones diferentes si tomamos en cuenta la superficie del municipio. Al 10%, se considerarían incluso diferentes para la variable *NAct* de Comercio mayorista y minorista si consideramos el tamaño de la población.

Por último, vamos a obtener los perfiles de especialización para los dos grupos de municipios.

Los gráficos correspondientes a los dos grupos de municipios ponen de manifiesto que los dos grupos tienen perfiles diferentes tanto en intensidad como en composición.

Gráfico 2
Perfil de actividad de los municipios de Almería en 1931.



4. Conclusiones y desarrollos futuros

En este trabajo se ha iniciado el estudio estadístico del número de actividades profesionales y mercantiles domiciliadas en los municipios de Almería en 1931, distinguiendo entre los municipios con ferrocarril y sin ferrocarril.

Podemos afirmar que la presencia del ferrocarril afecta a la configuración del paisaje empresarial y que su influencia es diferente según los sectores de actividad económica, a la vista de las siguientes conclusiones:

1. Para cualquier sector de actividad considerado el porcentaje de municipios en donde está presente dicho sector es mayor si el municipio tiene tren.
2. Para la variable número de actividades no existe evidencia empírica que permita concluir que el hecho de tener o no ferrocarril en la inmediata vecindad influya sobre esa variable. Creemos que esto pone de manifiesto la importancia de tener en cuenta el efecto tamaño en la variable de estudio.
3. Sin embargo esa diferencia sí que aparece cuando consideramos la densidad de las actividades en relación a la superficie. Son claramente subpoblaciones

diferentes en cuanto a sus medidas de posición, forma y dispersión. Además, podemos decir que esa variable en los sectores Agricultura, Industria Manufacturera, Comercio, Transporte y Almacenamiento, Hostelería, Actividades Profesionales y Otros Servicios, tiene un comportamiento diferente en los municipios con y sin ferrocarril. Las medias y las medianas son mayores en el primer grupo.

4. Para la variable número de actividades por habitante existen también diferencias, pero sólo para los sectores del Comercio y de Otros Servicios. En estos casos la media y la mediana también son mayores para los municipios con tren.
5. Centrándonos en el Comercio podemos decir que existen diferencias significativas al considerar por separado el comercio mayorista y minorista, pero únicamente para la variable ponderada por la superficie.
6. Los municipios con ferrocarril están especializados en Agricultura, Industria Manufacturera, Hostelería, Administración pública, Educación y Otros Servicios, mientras que el grupo de los municipios sin tren están especializados en Industria Extractiva, Suministro de Energía, Comercio, Transporte y Actividades Profesionales.

Como posibles líneas de estudio futuro sugerimos la necesidad de ampliar el análisis de las medidas a todos los municipios de la región; establecer categorías nuevas de municipios para ver si otras clasificaciones pueden determinar cambios en la configuración de ese paisaje empresarial; estudiar el cambio en los valores de las variables entre periodos diferentes de estudio y darle por tanto una dimensión evolutiva; y estudiar los sectores CNAE con un mayor nivel de desagregación de actividad.

7. Anexos

7.1. Clasificación Nacional de actividades Económicas (CNAE 2009) del INE.

A	Agricultura, ganadería, silvicultura y pesca;
B	Industrias extractivas;
C	Industria Manufacturera;
D	Suministro de energía eléctrica, gas, vapor y aire acondicionado;
E	Suministro de agua, actividades de saneamiento, gestión de residuos y descontaminación;
F	Construcción;
G	Comercio al por mayor y al por menor; reparación de vehículos de motor y motocicletas;
H	Transporte y almacenamiento;
I	Hostelería;
J	Información y comunicaciones;
K	Actividades financieras y de seguros;
L	Actividades inmobiliarias;
M	Actividades profesionales, científicas y técnicas;
N	Actividades administrativas y servicios auxiliares;
O	Administración pública y defensa; Seguridad social obligatoria;
P	Educación;
Q	Actividades sanitarias y de servicios sociales;
R	Actividades artísticas, recreativas y de entretenimiento;
S	Otros servicios;
U	Actividades de organizaciones y organismos extraterritoriales.

Esta última actividad (cuerpo consular) no se va tendrá en cuenta para el análisis puesto que sólo aparecen en los tres municipios portuarios.

7.2. Criterios metodológicos

Un contraste de hipótesis es un procedimiento que nos permite decidir si una afirmación o proposición sobre una población puede rechazarse o no, teniendo en cuenta el grado de compatibilidad con los datos obtenidos de una muestra. La afirmación se escribe en forma de hipótesis (denominada hipótesis nula, H_0) y su correspondiente alternativa (H_1).

Es necesario establecer una regla que nos permita decidir si la hipótesis nula puede o no ser rechazada. Para ello, usamos un estadístico de contraste y teniendo en cuenta los valores observados se rechazará o no la hipótesis nula. Se establecen dos zonas excluyentes: la zona de rechazo y la de aceptación. La zona de rechazo es el conjunto de valores del estadístico que se encuentran tan alejados de la afirmación establecida en H_0 , que es muy poco probable que ocurran si, como se supone, la hipótesis nula es verdadera. Su probabilidad se llama nivel de significación o de riesgo y se denota por α . La zona de aceptación es justo lo contrario, el conjunto de valores del estadístico que es probable que ocurran si la hipótesis nula es cierta.

La regla de decisión consiste en rechazar H_0 si el estadístico de contraste toma un valor perteneciente a la zona de rechazo ya que si la hipótesis nula planteada fuese cierta el estadístico es poco probable que tome ese valor. Si lo ha tomado, la conclusión es que la hipótesis planteada es falsa. Por el contrario, si no puede rechazarse solo puede afirmarse que no se dispone de evidencia empírica para rechazar esa hipótesis.

La mayoría de los paquetes estadísticos suministran el *p-valor*, que corresponde al nivel de significación más pequeño posible que puede escogerse, para el cual todavía se aceptaría la hipótesis alternativa con las observaciones actuales. Los valores pequeños indican que es muy infrecuente obtener una muestra como la actual. Tiene la ventaja de permitir que cualquiera decida qué hipótesis acepta basándose en su propio nivel de riesgo α . Lo usual es que si *p-valor* $< \alpha$ se rechaza la hipótesis nula. El software utilizado para los cálculos ha sido IBM SPSS 22.

Los contrastes se clasifican de forma general en paramétricos o no, atendiendo, entre otras cosas, al tipo de hipótesis o afirmación que se plantea en ellos. Si se conoce o se hace algún supuesto sobre la distribución que sigue la variable de la que proceden los datos se pueden hacer afirmaciones sobre algunas características de la población (media, varianza...) denominándose contrastes paramétricos. Sin embargo, lo usual, es no conocer esa distribución de partida con lo que se aplican métodos no paramétricos que permitirán contrastar hipótesis referidas a la aleatoriedad, forma de la distribución, independencia u homogeneidad entre variables, ..., etc.

Este tipo de contrastes tienen una serie de ventajas frente a los paramétricos:

- a) Son sencillos, ya que necesitan menos condiciones para su aplicación, al no depender de la forma en que se distribuya la variable en la población;
- b) Se pueden aplicar a más situaciones ya que pueden utilizarse con variables cualitativas con tal de que se puedan ordenar sus modalidades o incluso variables exclusivamente nominales;
- c) Son válidos en muestras muy cortas, en las que los procedimientos paramétricos pueden no serlo (es decir cuando los datos no puedan considerarse que provienen de una población que se distribuye según una Normal).

Como inconveniente, señalar que si la población de la que proceden es Normal son menos potentes que los contrastes de tipo paramétrico. Es decir si los datos son normales aumenta la probabilidad de que no se rechacen hipótesis falsas.

En nuestro caso se ha optado por este tipo de contrastes al comprobar la falta de normalidad (*p-valor* $< 0,001$ para prueba normalidad de Shapiro-Wilk) en todas las variables consideradas en cada uno de los sectores.

Dentro de los contrastes no paramétricos destacamos los de homogeneidad. Se utilizan para ver si dos muestras o subpoblaciones (de tamaño n y m respectivamente) proceden de la misma población, sin la necesidad de conocer la distribución de

probabilidad de la población de partida. Uno de los más utilizados es la Prueba de Kolmogorov-Smirnov, que permite contrastar las siguientes hipótesis:

$$\left. \begin{array}{l} H_0 : F_n(X_i) = F_m(X_i) \text{ (misma población)} \\ H_1 : F_n(X_i) \neq F_m(X_i) \text{ (distinta población)} \end{array} \right\}$$

Es decir, compara si las dos funciones de distribución (funciones de probabilidad acumuladas) para las dos muestras son las mismas. Para ello,

1. Se entremezclan y se ordenan los valores de las dos muestras de menor a mayor.
2. Se calcula cada una de las dos funciones de distribución empíricas, donde

$$F_j(x) = \frac{n^\circ \text{ observaciones} \leq x}{j}$$

3. Se calcula la diferencia entre las dos funciones de distribución y se coge el máximo en valor absoluto de las diferencias.
4. Por último se obtiene el estadístico

$$Z_{K-S} = \left(\max |F_n(x) - F_m(x)| \right) \sqrt{(nm)/(n+m)}$$

que sigue un modelo de probabilidad Normal (0,1).

Si la probabilidad de obtener una diferencia tan grande como la observada es muy pequeña (p-valor < 0,05), se puede rechazar la hipótesis nula de que ambas muestras proceden de la misma población.

Adicionalmente, se llevará a cabo otro contraste no paramétrico que permite comparar si las medias de las subpoblaciones difieren (prueba U Mann-Whitney). Es una alternativa a la prueba de diferencia de medias cuando se incumple el supuesto de normalidad. Se basa en la suma de los rangos asignados a las muestras de manera que si hay mucha diferencia entre ellos se puede rechazar la hipótesis de igualdad de medias. Para más detalle ver Pardo y Ruiz (2010) y referencias incluidas.

Por último, vamos a obtener una medida de localización espacial para obtener los perfiles de actividades de los dos grupos de municipios considerados para el estudio. En general se consideran L sectores repartidos en N regiones y se definen los cocientes de especialización de cada región como

$$L_{ij} = \frac{\frac{Y_{ij}}{Y_j}}{\frac{Y_i}{Y}}$$

donde,

Y_{ij} es el valor de la variable en la región i del sector j , $i=1,\dots,N$; $j=1,\dots,L$;

$Y_{i.} = \sum_{j=1}^L Y_{ij}$ es la suma de los valores de todos los sectores en la región i ;

$Y_{.j} = \sum_{i=1}^N Y_{ij}$ es la suma de los valores de todas las regiones en el sector j ;

$Y = \sum_{i=1}^N \sum_{j=1}^L Y_{ij}$ es la suma de los valores de todas las regiones y todos los sectores.

El cociente de localización regional del sector j en la región i , L_{ij} , es la relación que existe entre la participación de la región i en el sector j y la participación de la región en el total. Si $L_{ij} < 1$, indica una menor actividad del sector j en la región i en comparación con otros sectores, y al revés y nos permitirá obtener los perfiles de especialización de los dos grupos de municipios analizados.

7.3. Datos

Tabla 4, a.

NAct por municipio clasificadas según CNAE-2009 y presencia de ferrocarril en el municipio																				
N° Act	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	U
Municipios con Tren																				
Abia	57		26	1		6	40	3	15	2	2	3	4		6	4	3	2	7	
Abrucena	27		23			4	27	2	10				2		5	2	3	1	5	
Albánchez	20		38	1		3	43	2	10		1		3	1	5	3	3		9	
Albox	42		90	1	1	10	216	8	25	4	11		9	7	8	17	12	7	40	
Alcántar	7	3	16			2	19	1	6				1		5	5	1		4	
Alhabia	1		14	1		4	45	3	7				2		5	3	2		6	
Alhama de Almería	22		30	1	4	6	91	1	14	2	3		8		5	5	5	1	9	
Alicún	3		9			2	14	1	1						5	2	1		3	
Almería	41	28	251	4	3	11	1029	91	125	24	82	1	121	11	216	126	109	15	226	26
Alsodux	67		7	1			26		6				2		5	2	2		3	
Arboleas	23		38			10	50	4	21				1		5	8	4		5	
Armuña de Almanzora	7		4			1	5	1							5	2	2		3	
Bayarque	16	3	6	3		1	9	1	4						4	2	1		4	
Benahadux	15		10	1		3	21	4	8						5	1	1		5	
Cantoria	35		51	2		4	104	10	12	1	2		5		5	12	5	2	20	
Doña María - Ocaña	19		12			4	27	3	9				1		5	4	3		3	
Escúllar	9		15			3	23	1	4				1		5	2	1		3	
Fines	9		20			3	19	2	4						5	2	4		5	
Finana	28		36	1		5	56	6	12	1	1	3	5		5	4	4	2	7	
Gádor	40	2	21		2	5	86	4	10	1		2			6	2	4	3	12	
Gérgal	31	2	39		2	4	124	7	14	4	3	1	7		17	7	6	2	10	
Huécija	60		22		1	3	29	4	8					1	5	2	2	2	4	
Huércal de Almería	17	8	13	1		9	31	2	6				3		5	5	3		8	
Huércal Overa	46	2	63	1	1	7	189	13	25	3	3	1	9	12	15	17	10	2	23	
Laroya	10		10			2	12	1	4						5	2	1		4	
Líjar	8	4	9				9	1	2						5	2	1		3	
Lúcar	5		20			2	37	1	4						5	3	2	1	8	
Macael	15	1	26	1		3	36	3	7	1			2		6	2	5	1	9	
Nacimiento	9		11	1		1	35	2	6	1	1		2		5	2	2		6	
Olula de Castro	22		10			1	3	1	3				1		3	1			3	
Olula del Río	12		28	1		5	34	4	7		1			2	5	1	3		4	
Oria	79	1	53	1		4	60	2	14	2	1		5		5	15	3	1	11	
Partaloa	4		19		1	1	14	2	8						5	4	2	1	5	
Pechina	11	1	19	1		4	84	1	8	1	1		2		5	3	3		13	
Pulpí	23	2	22			3	63	7	12	1	1	1	2	5	10	4			7	
Purchena	6	1	37	1		2	40	4	7	2			15	4	10	5	2	2	9	
Rioja	36		13	1		3	33	2	5						5	2	3		7	
Santa Cruz de Marchena	18		4	1		1	25	2	5				1		4		1	1	4	
Santa Fe de Mondíjar	13	1	11	2		1	28	5	4				3		5	2	2		5	
Serón	22	5	48	2	5	4	137	8	25	2	5	1	7	4	6	19	11	2	19	
Sierro	11		16	1		4	25	1	2						5	3	1	1	3	
Somontín	20	1	17	1	1	1	23	1	7				1		5	2	3		4	
Sufí	13	3	20	1			12	30	5						5	2			3	
Terque			8		1	2	27	2	7				6		5	2	2		4	
Tíjola	21	2	41	2		9	72	4	11	2	1	1	3	1	5	3	3	3	9	
Urrácal	12		17	1		1	17	1	5						5	2	2	1	5	
Viator	2	2	15			5	68	2	10		2		1		5	3	2		9	
Zurgena	14		18			4	32	2	3		1			3	5	6	1	1	10	

Tabla 4, a.

NAct por municipio clasificadas según CNAE-2009 y presencia de ferrocarril en el municipio																				
Nº Act	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	U
Municipios Sin Tren																				
Adra	4		31	2		3	107	11	20	2	4				1	6	8	4	1	17
Alboloduy	8		27	2		3	38		4							5	3	2		5
Alcolea	49		24		1	3	46	3	10	1	1				3	1	5	3	4	6
Alcudia de Monteagud	19		12			3	12	1	4							5	2	3		3
Almócita	20		4	1			5	1								6	2	1		2
Antas	15	1	17	1		2	38	4	6						2	5	8	2		7
Bacares	7	2	18			2	13	1	10							5	2	2		7
Bayárcal	3		6			2	14		5							5	2	2		2
Bédar	12		10			3	12	1	6							5	4	3		2
Beires	8		9			3	21	1	2	1						5	2	4		2
Benífar	25		12	1		3	7		4							5	3	1		3
Benitagla			20			3	18	1	6							5	1	2		3
Benizalón	6		3			2	9	1	2							5	2			2
Bentarique	4		6			2	23	1	2							5	2	1		2
Berja	47	1	60	3		9	154	3	16	4	8	3	27	5	15	15	13	3		28
Canjáyar	19	2	23	1	1	2	83	5	12	3	1	2	12	10	3	6	2			12
Carboneras	11	2	13				26	1	3					4		5	7	3		4
Castro de Filabres	8		5			1	3		3							4	1			2
Chercos	6	2	7			4	10	2	4							5	2	2		3
Chirivel	18		17			4	23	5	10	1		1				5	4	1		7
Cóbdar	9	2	12			4	19	2	4							5	2	2		3
Cuevas de Almanzora	16	17	49	1	2	3	169	11	22	5	12	2	26	2	17	28	11	6		26
Dalias	15		44	2		6	120	2	23	3	3	2	6		5	10	6	3		18
Darrical	11		9			1	3	1								5	2	1		1
Enix	23	1	16			3	39	3	9							5	4	4	1	6
Félix	35		25	1	1	3	26	1	3							5	2	2		3
Fondón	24	10	11	1		2	21	1	2							5	4	2	1	8
Gallardos, Los	25	1	15	1		6	41	7	10					1	1	4	3	2		6
Garrucha	8	7	37	2		6	121	31	15	3	6		13	2	6	5	2	4	12	6
Illar	17		11	1		2	6	1	12							4	1	1		3
Instinción	4	2	14	1	2	2	63	6	4		2			1	1	5	2	2	2	6
Laujar de Andarax	30		47	1	1	8	66	1	8	1				1		5	4	3	1	12
Lubín	11	2	22	1		4	46	2	16	2	2			5	4	4	8	4	1	10
Lucainena de las Torres	16	18	30	1		2	79	5	10	1	1			1		5	5	7	1	6
María	22		48	1		5	69	3	15		2	1	13	1	5	5	3	3		13
Mojácar	9		8	1	1	2	29	1	5							5	7	8	1	7
Níjar	74	18	92	2		10	187	13	17	4	3			8	1	6	10	10		15
Ohanes	11		16	1		3	23		8							5	4	4	1	7
Padules	8		5			2	20	1	1							5	1	2		1
Paterna del Río	11		16	1	3	2	34	8	11			1	1			5	4	4	2	4
Rágol	2	2	9	1		1	39	2	8	1	1			1		5	2	1	1	6
Rooquetas de Mar	18	2	17	1	1	8	75	2	26	1				2		5	6	5		7
Senés	11		17			3	9		3							5	2			7
Sorbas	11		49		1	9	158	11	19	3	3			10		15	13	9	5	17
Tabernas	33	3	28	1		8	89	4	11	3	4			6	1	5	8	4	2	8
Taberno	23		8				14	1	2							5	5	1		5
Tabal	6		11			1	15	3	3					1		5	4	3		6
Turre	29		33	1	1	9	104	6	13	1	1	1	3			5	5	8	3	6
Turrillas	10	1	6			3	15	1	3							5	3	1	1	3
Uleña del Campo	10		24			6	38	1	7					3	1	5	1	4		4
Veñefique	19		26			5	12	9	6							5	2	7		3
Vélez Blanco	26		44	4		3	60	8	18	2		2	7	1	5	9	5	2		14
Vélez Rubio	28	3	55	2	4	2	117	7	25	6	5	1	15	3	16	8	9	6		13
Vera	8	4	36	1		5	91	4	14	2	5			12	5	8	8	4	2	10
Vicar	16		3			1	8		3							5	2	1		3