

Una aproximación sencilla al uso de las matemáticas en epidemiología

VIRGILIO GÓMEZ RUBIO
M.^a JOSÉ HARO DELICADO

Cesde épocas remotas la terna vida-enfermedad-muerte ha preocupado mucho al ser humano. Hay que conservar la vida y requisito imprescindible para ello es saber medir la salud, pero ¿qué significa medir la salud? Según la Organización Mundial de la Salud, se trata de establecer el nivel de salud y de bienestar de una población detectando la presencia y causas de las enfermedades y muertes, así como su expectativa de vida. La epidemiología es el procedimiento utilizado para ello por investigadores y médicos.

Epidemiología y medicina aparecen relacionadas desde sus orígenes. Ya Hipócrates analizaba conjuntamente causas de enfermedades con factores personales y ambientales. Sin embargo, hasta mediados del siglo XVII no se cuantificaron los patrones de nacimientos y muertes ni se analizó la distribución de enfermedades en poblaciones con unas determinadas características. Fue un comerciante en tejidos, Graunt, el primero en hacerlo. Hay ejemplos de gran relevancia que ponen de manifiesto la potencia y eficacia de los métodos epidemiológicos. Entre ellos cabe destacar los estudios de J. Snow en la epidemia de cólera en Londres en 1848-1849, las pruebas de asociación entre el hábito de fumar y el cáncer de pulmón desarrolladas por Doll y

Resumen Castellano

Palabras clave: palabra, palabra.

English title

English Abstract

Key words: word, word.

Peto y ensayos clínicos de campo como el de la vacuna de la poliomielitis que incluyó en torno a un millón de niños en edad escolar y que estableció, sin dejar lugar a dudas que la vacuna era segura y eficaz. En nuestros días, la epidemiología es un instrumento fundamental en investigación biomédica y salud pública, así como en la evaluación de los servicios sanitarios.

La serie de actividades que presentamos a continuación está destinada a estudiantes del segundo ciclo de la ESO y con ella se pretende:

- Afianzar conocimientos previos referentes a contenidos de Estadística Descriptiva y probabilidad.
- Introducir terminología básica y procedimientos propios del trabajo en epidemiología.
- Aprender nuevos contenidos referentes a Estadística Descriptiva, Inferencial y Probabilidad.
- Aprender a través de la experimentación y del descubrimiento.
- Aplicar los conocimientos a situaciones prácticas reales.
- Aprender a utilizar software para tratamiento estadístico que los ponga en contacto con técnicas estadísticas actuales y permita comprender y profundizar más en los contenidos desarrollados.
- Conmemorar el año Mundial de la Estadística poniendo de manifiesto sus aplicaciones en un tema de vital importancia como es el tema de la Salud.

Propuesta de trabajo en el aula

El desarrollo histórico de los procesos con los que se va a trabajar en el aula debe formar parte de cualquier propuesta didáctica. Poner a los estudiantes en contacto con su evolución histórica les ayudará a percibir la ciencia como algo vivo, que tuvo que atravesar gran número de dificultades hasta llegar a la situación actual. Las actividades en contextos reales son otra herramienta a utilizar. Ambos elementos pueden motivarles y animarles a perseverar con el fin de vencer las dificultades que surjan en el proceso de aprendizaje de los diversos conceptos y

procedimientos, a la vez que les hace ver la importancia de lo que aprenden.

Uno de los aspectos del enfoque epidemiológico es la observación de una serie de hechos en la población. Estos hechos son, entre otros, la presencia de enfermedad y la exposición a determinados factores (Organización Panamericana de la Salud. Organización Mundial de la Salud, 2002) Actuando sobre estos factores se puede modificar positivamente la realidad observada. Después de la recogida de datos, es fundamental ordenar, clasificar y agrupar la información.

Actividad 1. La siguiente tabla representa las edades de 120 personas que sufrieron una gastroenteritis después de asistir a la celebración de la cena de Nochevieja en un restaurante de la ciudad de Albacete (ver anexo). Agrupa los datos en intervalos y obtén la tabla de frecuencias. Representa gráficamente los datos de diferentes formas y extrae conclusiones sobre la distribución de casos en función de la edad. ¿Qué crees que ocurriría si agrupáramos en un número mayor de intervalos? ¿Cuál es la media de edad? ¿Es representativa? ¿Qué opinas sobre la simetría de la distribución?

Objetivos:

- Repasar y aplicar conceptos previos de estadística descriptiva.
- Profundizar en la importancia de realizar un buen trabajo de organización y representación de la información.
- Manipular fórmulas de mayor nivel conceptual en las que se relacionan medidas de centralización, posición y dispersión con el fin de obtener diferentes representaciones de una misma realidad.

Observaciones. Se entregará a los alumnos una tabla como la que se recoge en el anexo. La visualización de los resultados será una herramienta de gran ayuda para analizarlos, comprender la información presente en los datos y llegar a conclusiones e interpretaciones correctas.

Se repasarán los contenidos aprendidos en cursos anteriores según vayan surgiendo.

Se trabajará la agrupación de la información en un número cada vez mayor de intervalos de manera algebraica y gráfica, analizando el sentido de las fórmulas que se manejan y los pros y los contras de las diferentes formas de agrupamiento.

Otro aspecto de interés es la cuantificación o asignación de números a los hechos. Medir su ocurrencia o frecuencia es algo fundamental y, para ello hay que valerse de medidas y procesos estadísticos. Gran parte de estos procesos ya han sido trabajados por los estudiantes en 1.º y 2.º de la ESO. Se trata de aplicarlos a situaciones tratables con técnicas epidemiológicas, profundizar más en ellos y ampliarlos con la incorporación de nuevos conceptos y formas de actuación. Con las actividades 2, 3 y 4 se pretende lograrlo.

Actividad 2. ¿Cuál es la diferencia entre prevalencia e incidencia? ¿Cuál es la relación entre ambas?

Explica el siguiente gráfico:



¿Qué ocurrirá si aumenta la incidencia pero el número de muertes y recuperados se mantiene sin cambios? ¿Qué ocurrirá si, por el contrario, aumenta la mortalidad o más gente se recupera, pero sin aumentar la incidencia?

¿Qué ocurrirá, si se introduce una nueva prueba que detecte la presencia de enfermedad tempranamente? ¿Y si se introduce un medicamento que posponga o evite la mortalidad prematura, pero que no cure definitivamente, como ocurre con el SIDA? ¿Qué relación existe entre prevalencia, incidencia y duración de una enfermedad?

Actividad 3: En la tabla que puedes ver a continuación (ver anexo) se recogen los datos de muertes por asma en las diferentes provincias españolas en el año 2010. A la vista de los datos, responde las siguientes preguntas:

¿Dónde hay más casos?

¿Es el número de casos indicativo del riesgo, entendiéndose éste como probabilidad de morir por asma?

¿En qué provincias encontramos las tasas más altas?

¿Cómo calcularías la tasa de mortalidad nacional?

¿Qué provincias están por encima de la tasa de mortalidad nacional?

¿Qué ocurre cuando la tasa de mortalidad estandarizada es igual a uno? ¿Y cuándo es mayor que uno?

¿Crees que los casos están distribuidos de manera uniforme por todo el país? ¿Cómo crees que se podría justificar este hecho numéricamente? ¿Y gráficamente?

Objetivos:

- Repasar y aplicar conceptos adquiridos previamente como el de recuento, porcentaje, frecuencias absolutas y relativas, diferentes tipos de representaciones gráficas de datos categóricos y concepto clásico de probabilidad.
- Introducir algunos conceptos básicos propios de las técnicas epidemiológicas como los de morbilidad, prevalencia, incidencia, tasa de mortalidad estandarizada, riesgo, ...
- Trabajar con datos reales de modo que se perciba la importancia de las matemáticas en el mundo en que vivimos.

Observaciones. Se planteará a los estudiantes la actividad y se les facilitarán los datos en una tabla en papel y en un archivo para Hoja de Cálculo. Se les dejará algo de tiempo para que desarrollen estrategias y, a continuación, las presentarán y defenderán. Se discutirá sobre ellas, para posteriormente trabajar sobre las más acertadas, introduciendo paulatinamente los conceptos básicos necesarios para desarrollar la actividad de acuerdo a los procedimientos utilizados en bioestadística y en epidemiología.

Actividad 4. Los datos de mortalidad y población que se presentan a continuación corresponden al año 2012 en dos países latinoamericanos. El país A es un país escasamente industrializado; el país B es un país altamente industrializado.

A la vista de los siguientes datos, ¿Dónde crees que es el riesgo de morir más alto?

País A		País B	
Defunciones	Población	Defunciones	Población
1 269 166	68 386 000	5 564 944	1 9825 0000

Las autoridades del país B sospechan que los resultados no son muy fiables, dado que se han obtenido sobre poblaciones totales y que, como poco, habría que tener en cuenta la distribución de edad de la población, ya que el riesgo de morir o de contraer una enfermedad está relacionado con la misma. A la vista de los datos de la tabla siguiente, calcula:

- La tasa específica de mortalidad por cada 1000 habitantes para cada franja de edad y en cada país.
- El número de defunciones esperadas en cada país por franjas de edades y totales
- ¿Cuál es la tasa de mortalidad ajustada en cada país?
- ¿Cuál es el riesgo de morir en cada país?
- Extrae conclusiones.

Grupos de edad (años)	País A		País B	
	Defunciones	Población	Defunciones	Población
<15	317 308	19 831 740	94 169	24 781 250
15-44	338 100	35 218 790	380 430	79 256 250
45-64	270 261	10 941 760	1 223 875	61 501 250
>64	343 497	2 393 710	3 866 470	32 711 250
Total	12 69 166	68386000	5 564 944	198 250 000

Tabla 1. Datos obtenidos de la Organización Mundial de la Salud

Objetivos:

- Profundizar en los conceptos aprendidos.
- Introducir nuevos conceptos como los de tasas específicas y tasas ajustadas de mortalidad.
- Reconocer la importancia de reflexionar sobre los resultados obtenidos y la conveniencia de no aceptarlos sin más.

Observaciones. Se realizará la primera parte del ejercicio y se expondrán los resultados, invitándoles a la reflexión para, finalmente, llegar a una conclusión. Posteriormente se presentará la segunda parte del ejercicio, se resolverá y se compararán los resultados con los obtenidos en la primera parte de la actividad. Se les hará ver la importancia de utilizar el sentido común y analizar críticamente los resultados en cualquier situación.

Actividad 5. En los años 1853 y 1854 Londres se enfrentó a su tercera epidemia de cólera. Los habitantes de ciertos distritos del sur de la ciudad extraían el agua directamente de pequeños afluentes del Támesis o bien la obtenían de numerosas bombas de agua de uso público, abastecidas por dos compañías, Southwark and Vauxhall Water Company y Lambeth Water Company. En la tabla que aparece a continuación se muestran las muertes por cólera por cada lugar de abastecimiento de agua:

Compañía de agua	Hogares	Muertes por cólera
Southwark and Vauxhall Water Company	40 046	1 263
Lambeth Water Company	26 107	98
Resto de Londres	256 423	1 422

- ¿Cuál es el número de muertes por cada 1000 hogares?
- ¿Cuál es el número de defunciones esperadas en cada caso?
- ¿Cuál es la tasa de mortalidad estandarizada?
- ¿Crees que existe mayor riesgo de morir en un distrito que en otro?

Objetivos:

A los objetivos anteriores añadimos el de trabajar con datos reales y sobre un acontecimiento histórico de gran relevancia, ya que marcó el inicio de las técnicas epidemiológicas.

Las medidas de asociación permiten determinar la existencia de relación entre la exposición a un factor, considerado de riesgo y la ocurrencia de enfermedad en una población. Dos variables están asociadas cuando existe entre ellas una relación de dependencia estadística, es decir, si la probabilidad de ocurrencia de un evento depende de la ocurrencia de otro o de otros. En epidemiología, investigando la búsqueda de una posible asociación, muchas veces, es posible establecer una relación de causa-efecto entre exposición y enfermedad. Por otra parte, los factores de riesgo están asociados a determinados comportamientos, estilos de vida, genética, herencia y exposición ambiental.

Hasta el momento, sólo se han utilizado porcentajes, procedimientos de estadística descriptiva y una tímida alusión al concepto clásico de probabilidad, sin embargo, el análisis epidemiológico también requiere de la realización de inferencias para toda una población a partir de la información extraída de una muestra.

Actividad 6. En una localidad rural de 760 habitantes, situada en una zona muy húmeda, se observó que en el último año, la incidencia acumulada de casos de asma en los campesinos que trabajaban en los humedales fue de 88,2 por mil, mientras que en los habitantes con otro tipo de actividad laboral, lejos del campo, fue 55,8 por mil. De acuerdo al censo local más reciente, en la localidad hay 204 campesinos. ¿Podríamos asociar la ocupación en los humedales con el asma?

	<i>Padecen de asma</i>	<i>No padecen de asma</i>	<i>Totales</i>
Campeños	18	186	204
Resto de habitantes	31	525	556
<i>Totales</i>	49	711	760

Actividad 7. Se realiza una investigación sobre una nueva vacuna contra la gripe. Se elige una muestra aleatoria de 900 individuos y se clasifica a cada uno de ellos según haya contraído la gripe o no durante el último año y según haya sido o no vacunado. Se obtiene la siguiente información:

<i>Vacunado</i>	<i>Contraído la gripe</i>	
	<i>Sí</i>	<i>No</i>
<i>Sí</i>	150	200
<i>No</i>	300	250

¿Existe relación entre haber recibido o no la vacuna y contraer la gripe? (Milton, 2004)

Objetivos:

- Introducir el concepto y técnicas básicas del contraste de hipótesis.
- Introducir nuevas formas de actuación en epidemiología.
- Aprender a usar una herramienta informática en las técnicas básicas del contraste de hipótesis.
- Aplicar lo aprendido en situaciones reales.

Observaciones. Se introducirán previamente los conceptos básicos propios del contraste de hipótesis. El ejercicio se realizará primeramente sin la ayuda de ningún programa informático, apoyándose en los valores recogidos en las tablas de la χ^2 . Posteriormente, se podrá utilizar software apropiado para trabajar contraste de hipótesis.

Actividad 8. Los adictos a la cocaína necesitan esta droga. Quizás dándoles una medicación que combatiere la depresión se les ayudaría a dejarla. Un estudio de tres años comparó un antidepresivo denominado Desipramina con el litio (tratamiento habitual para combatir la adicción a la cocaína) y con un placebo. Los sujetos experimentales eran 72 cocainómanos. Se asignaron al azar 24 sujetos a cada tratamiento. En la tabla adjunta se muestran los recuentos y las proporciones de sujetos que no recayeron en el consumo de la cocaína durante el estudio.

<i>Tratamiento</i>	<i>Sujetos</i>	<i>No recayeron</i>	<i>Sí recayeron</i>
Desipramina	24	14	10
Litio	24	6	18
Placebo	24	4	20

¿Hay diferencias entre las tres situaciones? Si es así, estudia entre cuáles de ellas. (Moore, 2000)

Objetivos:

A los objetivos de las dos actividades anteriores agregamos el de trabajar con más de dos tratamientos, lo que nos permite adentrarnos un poco más en el contraste de hipótesis.

Si con las anteriores pruebas se podía detectar la existencia de relación entre dos variables, en epidemiología interesa también detectar la fuerza de esa relación, por ejemplo, es muy importante establecer el riesgo que tienen de contraer una enfermedad personas expuestas a la misma. Con las siguientes actividades se pretende trabajar con diversas medidas que cuantifican esa fuerza como son los riesgos absolutos y relativos y la odds ratio.

Actividad 9. En la siguiente tabla (ver anexo) se muestra la distribución de muertes por asma en las diferentes provincias españolas, así como la clasificación de las mismas en «Muy alto», «Alto», «Medio» y «Bajo», según los niveles de humedad. ¿Existe un exceso de riesgo en los grupos de niveles «Muy Alto» y «Alto» con relación al grupo formado por los niveles «Medio» y «Bajo»? ¿Crees que el factor humedad favorece el aumento del número de muertes por asma?

Actividad 10. Los datos que se muestran en la tabla siguiente exponen la relación entre el hábito de fumar y la presencia de problemas vasculares en una muestra de 240 sujetos

	<i>Con problemas</i>	<i>Sin problemas</i>	<i>Totales</i>
Fuman	23	81	104
No fuman	9	127	136
<i>Totales</i>	32	208	240

- Obtén la proporción de casos con problemas vasculares entre los fumadores.
- Obtén la proporción de casos con problemas vasculares entre los no fumadores.
- Obtén el riesgo relativo. ¿Qué indica?
- Obtén la razón de posibilidades. ¿Qué indica?
- Compara ambas medidas

Actividad 11. Se ha realizado un estudio sobre la edad de la madre en el momento del nacimiento de su hijo como factor de riesgo en el desarrollo del síndrome de la muerte súbita del lactante (SMSL). Se seleccionaron para el estudio un total de 7 330 mujeres que estaban por debajo de los 25 años en el momento del nacimiento del niño. De ellas, 29 tuvieron niños afectados de SMSL. De las 11 256 mujeres seleccionadas para el estudio que tenían 25 años o más en el momento del nacimiento de sus hijos, 15 tuvieron niños con SMSL. Estos datos se muestran en la tabla siguiente:

Edad	Sí	No	Totales
Menos de 25 años	29	7 301	7 330 (Fijo)
25 años o más	15	11 241	11 256 (Fijo)

- ¿Crees que existe un exceso de riesgo?
- Obtén la razón de posibilidades, ¿Qué indica?
- Analiza los resultados.
(Milton, 2004)

Objetivos:

- Trabajar con conceptos más elaborados como el de probabilidad condicionada.
- Introducir nuevos conceptos básicos en epidemiología.
- Trabajar sobre problemas reales.
- Favorecer la reflexión y el análisis crítico de los resultados obtenidos con el fin de llegar a conclusiones válidas.

Los test de diagnóstico sirven para detectar la presencia de condiciones específicas en unidades experimentales. Uno de los usos más habituales de este tipo de test es la detección de enfermedades o de factores genéticos. En la base del método empleado para ello aparece la teoría de la probabilidad.

Actividad 12. El suero de una mujer embarazada puede ser analizado por medio de un procedimiento llamado electroforesis en gel de almidón. Este procedimiento permite detectar la presencia de una zona proteínica llamada zona de embarazo, la cual se supone que es un indicador de que el niño es una hembra. Para investigar las propiedades de este test se seleccionaron 300 mujeres para su estudio. En la siguiente tabla se dan los resultados del test y los sexos de los niños nacidos.

Zona de embarazo	Varón (realidad -)	Mujer (realidad +)	
Presente (test +)	51	78	129 (aleatorio)
Ausente (test -)	96	75	171 (aleatorio)
	147 (aleatorio)	153 (aleatorio)	300 (fijo)

- ¿Cuál es la probabilidad de que siendo varón el test indique que es hembra?
- ¿Cuál es la probabilidad de que siendo hembra el test indique que es varón?
- ¿Cuáles son los valores predictivos positivos y negativos?
- ¿Cuál es la especificidad del test?
- ¿Cuál es la sensibilidad?
(Milton, 2004)

Actividad 13. En una Comunidad Autónoma española hay 4 millones de habitantes y la enfermedad X afecta en cada momento a 4 000 de ellos. La enfermedad tiene un pronóstico muy malo si no se consigue un diagnóstico precoz. Se aplica un test de diagnóstico que tiene una sensibilidad del 99% y una especificidad del 80%. Se somete a este test a todos los habitantes de la Comunidad y aquellos a los que da positivo son citados para ser sometidos a un estudio más completo, con el fin de averiguar si efectivamente tienen la enfermedad. Entre los convocados reina nerviosismo, pues temen que la prueba definitiva confirme la presencia de la enfermedad mortal. Pero uno de ellos se muestra tranquilo y confiado en no tener la enfermedad. Queremos averiguar si este sujeto es:

- a) Un inconsciente que no se percata de su alto riesgo de enfermedad.
- b) Un desesperado de la vida al que no le importa morir.
- c) Un matemático que calculó correctamente su riesgo de enfermedad
(Prieto y Herranz, 2008)

Objetivos:

- Introducir los teoremas de la probabilidad total y de Bayes.
- Introducir nuevos conceptos epidemiológicos como especificidad, sensibilidad y valores predictivos.
- Aplicar los conceptos anteriores en situaciones reales.

Una de las funciones más importantes de la epidemiología es determinar la relación existente entre frecuencia de la enfermedad y exposición a determinados factores

que se asocian con la enfermedad. Se denominan marcadores de riesgo a aquellos atributos que se asocian con un mayor riesgo de ocurrencia de una determinada enfermedad y que pueden ser controlados y prevenidos. Tres son los elementos principales a tener en cuenta cuando aparecen casos de una determinada enfermedad: quién la padece, dónde aparece y cuándo.

Actividad 14. Un niño está incubando una gripe cuando va a clase un lunes. Cada día que pasa en contacto con sus compañeros, tiene una probabilidad de contagiar a cada uno de ellos de 0.5. Supongamos que en la clase son 20 niños.

¿Cuál es la probabilidad de que el martes no haya ningún niño más infectado? ¿Y de que haya alguno? ¿Y el miércoles? ¿Y el jueves? ¿Y el viernes?...

Actividad 15. Supongamos que después de un tiempo prolongado, al colegio llega un niño enfermo que no es detectado hasta la hora del recreo. Vamos a suponer que el periodo de incubación de la enfermedad que padece el niño es de un día y que sólo durante ese día se puede contagiar la enfermedad. Si la probabilidad de ser contagiado es de 0,7 ¿Cuál es el número de niños más probable que se sentirán enfermos al día siguiente en el colegio? ¿Cuál es esa probabilidad?

Nota: Los niños que son infectados ese día no empezarán a contagiar a los demás hasta el día siguiente. Con lo cual, sólo hacemos referencia al contagio generado por el primer niño

Actividad 16. En una familia de tres hijos, uno de ellos es contagiado el lunes de una enfermedad vírica por un compañero de colegio. La probabilidad de que contagie a cualquiera de sus hermanos es de 0.3. ¿Cuál es la probabilidad de que para el sábado se hayan contagiado los otros dos niños? ¿Cuántos días se piensa que deberían pasar para que dicha probabilidad sea igual a uno?

Nota: El enunciado de este ejercicio es independiente del enunciado del ejercicio anterior.

Objetivos:

- Introducir los conceptos de experimento de Bernoulli y modelo de distribución binomial.
- Introducir el concepto de cadena de Markov finita con probabilidades de transición estacionarias.

- Dar un paso más en la aplicación de las matemáticas a problemas de epidemiología
- Saber apoyarse en la visualización y en el uso de herramientas dinámicas para resolver problemas que de otro modo serían difíciles de resolver.

Observaciones. Lo más probable es que los estudiantes con los que se vaya a trabajar ya sepan utilizar Geogebra, si no es así, se les entregará un guión con las opciones básicas del programa. Para realizar la actividad, necesitarán conocer algunos conceptos muy básicos sobre matrices y producto de matrices. Para calcular las potencias sucesivas de una matriz se utilizará el programa Geogebra.

La epidemiología espacial también representa un importante recurso para estudiar, entre otras cosas, la existencia de puntos geográficos relacionados con una determinada enfermedad. Cabe destacar entre las diferentes técnicas existentes el estudio de procesos puntuales, que son apropiadas cuando se sospecha que el riesgo aumenta cerca de la fuente. Tres son los propósitos fundamentales para los que se usan los procesos puntuales: Analizar la distribución que presentan los datos espaciales, con el fin de concluir si están distribuidos al azar, uniformemente o formando clústeres; analizar la densidad (número de individuos por unidad de área); comparar.

Las actividades siguientes están relacionadas con este tipo de técnicas y el objetivo fundamental de las mismas es el de introducir a los estudiantes en un procedimiento más, dentro del terreno de la epidemiología.

Actividad 17. En los mapas que se adjuntan de la ciudad de Albacete aparecen reflejados (puntos) casos de asma en la población, distribuidos geográficamente. Aparece marcada en el mismo una zona en la que estuvieron instaladas varias fábricas de harinas. Además, la ciudad se ha dividido en hexágonos, de manera que para cada uno de ellos aparece el número de casos de asma y la población residente (ambos datos han sido simulados). ¿Crees que la localización de los casos es completamente al azar? Si no, ¿de qué factores crees que depende? ¿Se parece la tasa en un hexágono a las tasas de sus vecinos (es decir, los hexágonos con los que comparte frontera)?

Analiza el mapa y calcula el riesgo por 1 000 habitantes de contraer la enfermedad en cada hexágono. ¿Cómo estudiarías si hay un riesgo mayor de acuerdo a la distancia a la Fábrica de Harinas? Calcula la distancia en hexágonos

hasta el foco de riesgo y representa en una gráfica el riesgo en cada hexágono y su distancia al foco, ¿a qué conclusiones llegas? ¿Cómo calcularías el riesgo en función de la distancia a la Fábrica de Harinas?

Objetivos:

- Introducir el concepto de proceso puntual aleatorio uniforme y no uniforme.
- Determinar de manera geométrica si un proceso puntual es completamente aleatorio y uniforme.
- Introducir el concepto de correlación espacial (Bivand, Pedesma y Gómez-Rubio, 2013).
- Aplicar estos conceptos a problemas de salud pública.

Observaciones. Se proporcionan tres mapas de la ciudad de Albacete (ver anexo) dependiendo de cómo se han generado los casos:

- En una malla regular (mapa-reg.png).
- Igual que la distribución de la población (mapa-unif.png).
- Con incremento de casos alrededor de la Fábrica de Harinas (mapa-cluster.png).

De esta manera exploramos tres situaciones distintas. Se recomienda que los alumnos trabajen por parejas para hacer los cálculos a mano de manera más rápida y que luego pongan los resultados en común ya que al haber varios mapas los resultados y las conclusiones son distintas.

Se recomienda que las tasas se calculen sobre 1 000 habitantes, es decir, $tasa = 1000 \cdot \text{casos}/\text{población}$. De esta manera los números son más manejables. Las tasas se pueden escribir en los hexágonos, debajo de los otros números. Además, se recomienda calcular primero la tasa global para usarla como referencia.

Para estudiar la autocorrelación espacial se puede comparar cada tasa con el promedio de las tasas en los vecinos y ver si se parecen o no.

A la hora de estudiar la distribución al azar de los casos hay que tener en cuenta que, si el proceso es totalmente aleatorio, el número de casos sólo depende de la población. Por tanto, se puede calcular el número de casos esperado en cada hexágono

como la población en ese hexágono por la tasa bruta global. Con esto podemos calcular un estadístico Chi-cuadrado y ver si es significativo o no.

Para calcular la influencia del foco de riesgo se calcula una tasa sumando los casos y la población de los hexágonos que están a la misma distancia del foco de riesgo. Hay un hexágono a distancia 0, 6 hexágonos a distancia 1, etc. La distancia máxima es de 6 hexágonos. Luego, se puede crear una gráfica con la distancia al foco de riesgo (x) y la tasa de riesgo (y).

Actividad 18. En el siguiente mapa (ver anexo) se representa la distribución de casos de muerte por cólera en Londres en los años 1853 y 1854. Calcula las densidades en cada una de las regiones ¿crees que existen en el mapa zonas de mayor riesgo? Justifica tus respuestas.

Objetivos:

- Introducir a los estudiantes en técnicas de análisis espacial de datos.
- Enlazar estadística con geometría.
- Potenciar la reflexión y el análisis crítico.
- Aplicar las técnicas de datos espaciales a casos prácticos reales.
- Utilizar el desarrollo histórico de la epidemiología como fuente de motivación.

Observaciones. Los estudiantes reflexionando sobre lo que observan tendrán que determinar cómo se distribuyen los datos (al azar, uniformemente o en grupos), posteriormente utilizarán diferentes técnicas (cálculo de riesgos en regiones o cálculo de densidades) para establecer la existencia de relación o no entre determinados focos de riesgo y el desarrollo de la enfermedad.

Tanto en un caso como en otro necesitarán echar mano de la geometría, ya sea

para el cálculo de distancias como para el cálculo de áreas de regiones planas.

Conclusiones

Las actividades presentadas en este trabajo muestran cómo es posible introducir numerosos conceptos de Probabilidad y Estadística utilizando ejemplos muy prácticos sacados del campo de la Epidemiología. Pensamos que estos ejemplos serán de mucho interés para los alumnos (entre otras cosas, porque el brote actual de ébola en África aparece con asiduidad en las noticias). Hemos dejado fuera de este trabajo otros ejemplos que creemos que pueden ser interesantes y que esperamos poder desarrollar en breve.

Aunque no hemos podido trabajar, con estudiantes de secundaria obligatoria, todas las actividades presentadas en este artículo, sí que lo hemos hecho con algunas de ellas y el nivel de interés de los estudiantes ha sido muy alto, implicándose en la resolución de los problemas con ahínco, a pesar de que el conocimiento previo que poseían sobre las ideas que se pretendía desarrollar era prácticamente nulo. El trabajar con casos relacionados con la realidad y el entorno de los chicos ha hecho que se sintieran más motivados.

Además de los ejercicios planteados aquí, los profesores interesados en ampliar información y ejemplos pueden consultar algunas propuestas interesantes en inglés. Hablamos de los programas

- *Epidemiology in the Classroom* del CDC
<<http://www.cdc.gov/excite/classroom/>>.
- *Detectives in the Classroom*
<<http://www.montclair.edu/Detectives>>.

El trabajar con casos relacionados con la realidad y el entorno de los chicos ha hecho que se sintieran más motivados.

- *Epidemiology in Action*
<http://www.collegeboard.com/prod_downloads/yes/holmstrom.pdf>.

Referencias bibliográficas

- MILTON, S. (2004), *Estadística Para Biología y Ciencias de la Salud*, Mc Graw Hill Interamericana, Madrid
- MOORE, D. (2000), *Estadística Aplicada Básica*, Antoni Bosch Editor, Barcelona
- ORGANIZACIÓN PANAMERICANA DE LA SALUD. ORGANIZACIÓN MUNDIAL DE LA SALUD (2002), *Módulo de principios de epidemiología para el control de enfermedades (MOPECE)*, OPS, Washington D.C.
- PRIETO, L., e I. HERRANZ (2000), *Bioestadística: Diseño de experimentos y análisis de datos en investigación biomédica*, Compañía Española de Reprografía y Servicios, Madrid
- BIVAND, R.S., E. PEDERSEN y V. GÓMEZ-RUBIO (2008), *Applied Spatial Data Analysis with R*, Springer, Nueva York.

Agradecimientos

Agradecemos a Juan Carlos Orengo Valverde, Director Asociado de Investigación Clínica en Puerto Rico MSD, sus comentarios y consejos sobre esta propuesta docente y a M.^a Llanos Pérez Haro, licenciada en medicina por la Universidad Complutense de Madrid, sus orientaciones sobre las técnicas más habituales en epidemiología. También nos gustaría agradecer a los participantes en el VI Seminario Nacional de Estalmat celebrado en Medina del Campo, los días 12 y 13 de abril de 2013 por sus comentarios tras la presentación de este trabajo. Por último, nos gustaría agradecer a los profesores José Javier Orengo y José Luis Martínez de la Torre por organizar una sesión con estudiantes de secundaria en el IES Andrés de Vandelvira de Albacete, para chequear estas actividades en el aula. Por último, agradecemos a un revisor anónimo sus ideas para mejorar este artículo.

Tabla para la actividad 1

27	32	58	44	24	32	29	50
28	36	38	48	38	47	29	39
40	37	35	36	36	36	42	45
32	36	48	42	46	35	32	54
30	29	36	44	30	36	27	37
35	33	38	63	37	53	35	46
36	27	34	32	46	38	43	20
25	41	27	53	40	31	47	43
29	40	50	34	47	36	38	24
30	51	43	46	38	49	47	30
29	33	54	40	28	63	36	41
46	51	49	37	41	37	39	38
30	35	36	34	43	43	37	55
29	44	38	42	43	35	42	50
35	47	32	54	41	41	35	40

Edades de 120 personas que sufrieron gastroenteritis

Tabla para las actividades 3 y 9

Provincia	Población	Asma	Grado de humedad
Albacete	396444	5	Bajo
Alicante	1891875	25	Medio
Almería	690487	19	Medio
Álava	311048	17	Muy alto
Asturias	1057215	101	Muy alto
Ávila	167967	11	Bajo
Badajoz	677181	7	Bajo
Baleares	1082091	11	Alto
Barcelona	5364876	109	Alto
Vizcaya	1136674	42	Muy alto
Burgos	362862	14	Alto
Cáceres	405620	5	Bajo
Cádiz	1221911	15	Medio
Cantabria	578713	6	Muy alto
Castellón	591477	4	Medio
Ciudad Real	519145	9	Bajo
Córdoba	787502	39	Bajo
La Coruña	1124438	30	Muy alto
Cuenca	213944	7	Bajo
Guipúzcoa	690966	12	Muy alto
Gerona	726700	12	Alto
Granada	909761	50	Bajo
Guadalajara	243717	3	Bajo
Huelva	507575	12	Medio
Huesca	221292	8	Alto
Jaén	654450	11	Bajo
León	481058	22	Alto
Lérida	430911	8	Alto
Lugo	341312	13	Muy alto
Madrid	6352983	66	Bajo

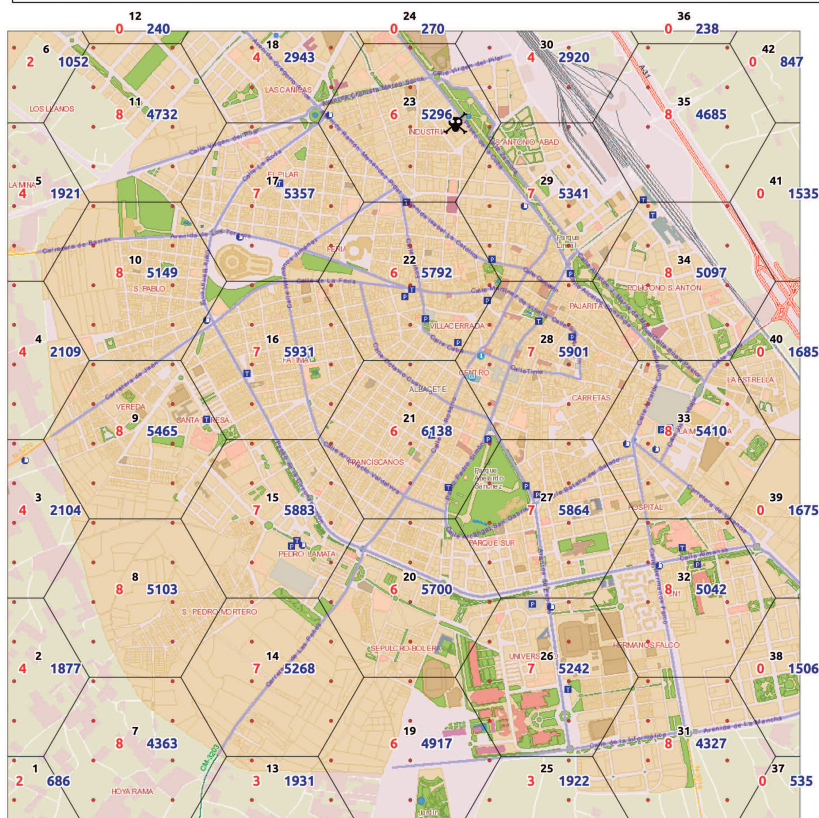
Provincia	Población	Asma	Grado de humedad
Málaga	1 588 896	28	Medio
Murcia	1 464 715	23	Bajo
Navarra	620 573	13	Alto
Orense	325 023	22	Alto
Palencia	168 935	5	Alto
Las Palmas	1 077 981	16	Alto
Pontevedra	946 856	17	Muy alto
La Rioja	314 106	7	Alto
Salamanca	345 277	7	Alto
Santa Cruz de Tenerife	1 015 296	15	Alto
Segovia	159 694	5	Medio
Sevilla	1 871 819	18	Bajo
Soria	92 872	7	Medio
Tarragona	794 426	22	Alto
Teruel	143 011	5	Medio
Toledo	666 998	10	Bajo
Valencia	2 512 218	56	Medio
Valladolid	525 292	10	Alto
Zamora	191 640	12	Alto
Zaragoza	949 108	17	Alto
Ceuta	75 207	0	Alto
Melilla	73 574	1	Alto

Datos de población y casos de asma obtenidos de la web del INE

Mapas correspondientes a la actividad 17

En los mapas, el número de casos aparece en rojo, la población en azul y el número de hexágono en negro.

Investigación sobre los casos de asma alrededor de la Fábrica de Harinas



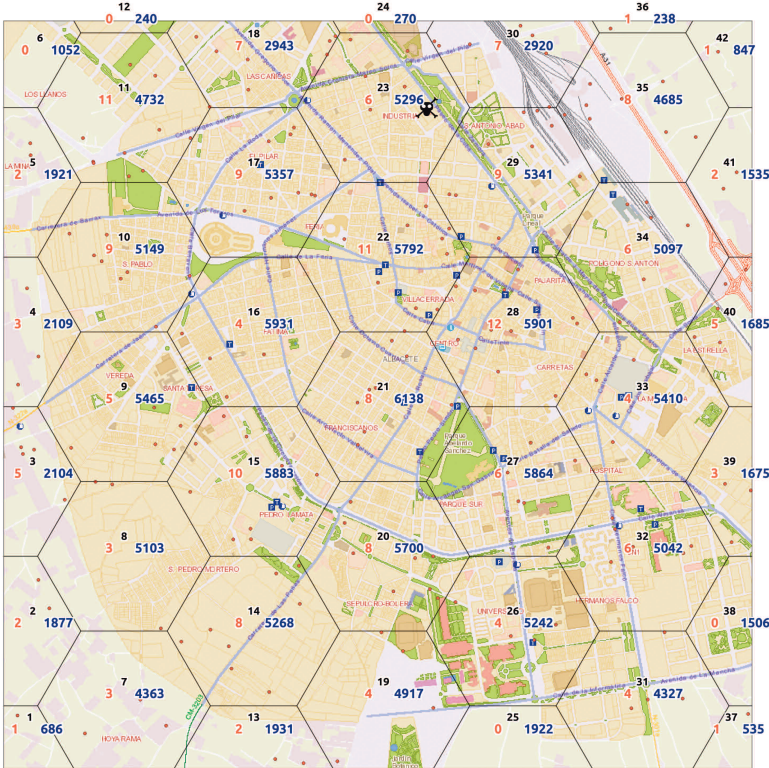
Casos: 200

Población: 150000

Hexágonos: 42

Mapa con una distribución de casos de acuerdo a una malla regular (es decir, aproximadamente el mismo número de casos por hexágono independientemente de la población) (Fichero mapa-reg.png)

Investigación sobre los casos de asma alrededor de la Fábrica de Harinas



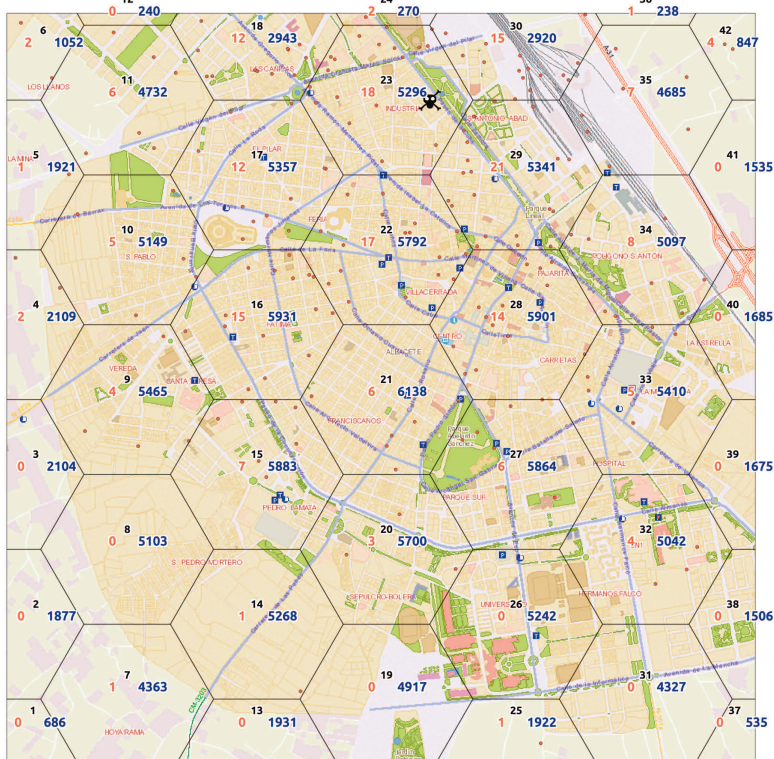
Casos: 200

Población: 150000

Hexágonos: 42

Mapa con una distribución de casos según un riesgo uniforme (es decir, el número de casos es proporcional a la población del hexágono) (Fichero mapa-unif.png)

Investigación sobre los casos de asma alrededor de la Fábrica de Harinas

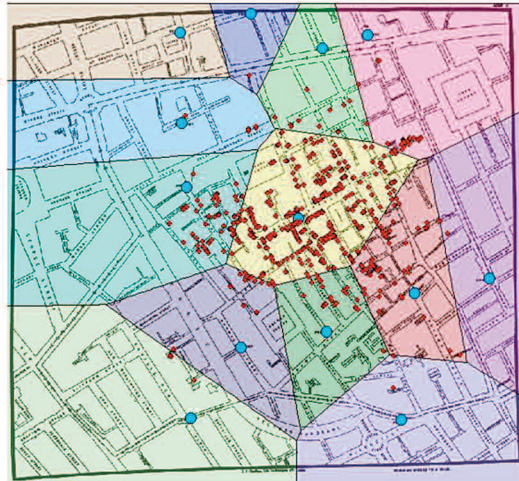


Casos: 200

Población: 150000

Hexágonos: 42

Mapa con una distribución de casos más alta alrededor del foco de riesgo (es decir, hay proporcionalmente más casos alrededor del foco de riesgo). (Fichero mapa-cluster.png)



Mapa del barrio del Soho, Londres (Reino Unido)
Los puntos azules representas fuentes de agua
mientras que los rojos representan casos de cólera