

ADQUISICIÓN DEL SENTIDO DE LA CORRELACIÓN A PARTIR DE RECURSOS VIRTUALES

María Magdalena Gea¹, Emilse Gómez² y Gustavo Cañadas¹

¹Universidad de Granada (España) - ²Universidad Nacional de Colombia
mmgea@ugr.es, egomez@unal.edu.co, grcanadas@ugr.es

A pesar de la importancia de la idea de correlación, la investigación sobre el tema sugiere pobres capacidades de los adultos en la estimación de la misma, así como existencia de sesgos asociados. Son también numerosas las dificultades que permanecen después de la enseñanza del tema. En este trabajo analizamos los componentes del sentido de la correlación y regresión y algunos ejemplos de recursos interactivos en Internet que podrían utilizarse para desarrollar este sentido en los estudiantes.

PALABRAS CLAVE

Correlación, recursos virtuales, sentido estadístico.

INTRODUCCIÓN

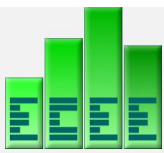
La estadística es actualmente uno de los temas más enseñados en todos los niveles educativos. Sin embargo muchas las investigaciones nos alertan de la variedad de errores en la interpretación y aplicación de conceptos estadísticos elementales, incluso después de la instrucción.

Batanero, Díaz, Contreras y Roa (2013) explican esta situación contradictoria por la falta de sentido estadístico en los estudiantes y proponen un modelo de dicho sentido. En lo que sigue, aplicamos su definición de sentido estadístico al caso particular de la correlación y seguidamente, sugerimos algunas actividades con recursos interactivos en Internet que pueden ayudar a adquirir este sentido por parte de los estudiantes.

MARCO DE REFERENCIA

En el modelo de sentido estadístico de Batanero Díaz, Contreras y Roa (2013), se incluye la comprensión de las ideas estadísticas fundamentales (Burrill y Biehler, 2011), el razonamiento estadístico (Wild y Pfannkuch, 1999) y actitudes, lo cual aplicamos a para describir el sentido de la correlación. Las ideas fundamentales sobre correlación son las siguientes:

- *Datos y distribución bivariada.* En el estudio de la correlación y regresión se manejan datos bivariantes, es decir, una variable estadística bidimensional y la correspondiente distribución bidimensional. Asociada a la misma, es necesario que los estudiantes diferencien las frecuencias absolutas y relativas conjuntas, marginales y condicionales.
- *Representación tabular y gráfica.* En el estudio de la correlación y regresión los datos se organizan en una tabla de doble entrada. La representación gráfica más



utilizada es el diagrama de dispersión, aunque también encontramos histogramas o gráficos tridimensionales y gráficos de burbuja. El diagrama de burbujas también permite visualizar, simultáneamente, hasta tres variables (representando la tercera mediante el diámetro) o incluso cuatro, si mediante el color pudiera representarse una cuarta variable.

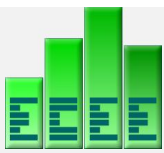
- *Dependencia funcional, estadística e independencia.* Mientras que en una dependencia funcional a cada valor de una variable x (independiente) corresponde un solo valor de otra variable y (dependiente), en la dependencia estadística a cada valor de x corresponde una distribución de valores de y , por lo que este concepto amplía el de dependencia funcional.
- *Covarianza y correlación.* Con objeto de medir el signo y la intensidad de la dependencia entre dos variables estadísticas, se introducen otros dos conceptos importantes: la covarianza y la correlación. Es de gran utilidad la explicación de su significado mediante la división en cuatro cuadrantes de la nube de puntos, por medio de las rectas correspondientes a las medias de cada variable, que Holmes (2001) propone. Así mismo, este análisis propicia que el estudiante comprenda más significativamente el cálculo del coeficiente de correlación.

También es necesario desarrollar en los estudiantes el razonamiento sobre la correlación; en primer lugar, mejorar su estimación de la correlación a partir de diversas representaciones de datos. Para ello, conviene presentar a los estudiantes ejemplos diversos de diagramas de dispersión teniendo en cuenta las variables que influyen en la dificultad de la estimación de la correlación (Sánchez Cobo, Estepa y Batanero, 2000): signo, intensidad de la correlación, número de datos y si la correlación en los datos coincide o no con la esperada por el estudiante.

Es también importante desarrollar la comprensión de la diferencia entre correlación y causalidad, mediante la discusión con los estudiantes sobre la explicación de la correlación, donde además de una relación causa-efecto, otras posibles explicaciones de la correlación (Barbancho, 1973) son: la interdependencia, cada variable afecta a la otra; dependencia indirecta, donde una tercera variable es la que determina la correlación; concordancia o coincidencia en preferencia u ordenación de una misma serie de datos por dos personas, y covariación casual o espúrea.

Finalmente, el sentido de la correlación requiere unas buenas actitudes, donde los alumnos se interesen por el tema y valoren su importancia para la modelización matemática en diversas áreas de conocimiento. Asimismo han de superar sus creencias erróneas, entre las que encontramos las siguientes (Estepa, 1994):

- *Concepción causal*, confundir correlación y causalidad.
- *Concepción unidireccional*, no considerar la correlación negativa o estimando una intensidad menor de la misma.
- *Concepción local*, basar la estimación de la correlación en el análisis de sólo una parte de los datos.
- *Concepción determinista*, considerar sólo la dependencia funcional y no la dependencia estadística.



ASPECTOS METODOLÓGICOS

Con objeto de proponer recursos para desarrollar los componentes descritos del sentido de la correlación se realizó una búsqueda de recursos en Internet que pueden contribuir a ello. Como se expone en Contreras (2009), estos recursos permiten crear micromundos virtuales, en el que los conceptos y objetos matemáticos abstractos pueden concretizarse. Proporcionan también a los estudiantes experiencia estocástica que es difícil de conseguir con otros medios. Dichos medios están hoy día al alcance de alumnos y profesores a partir de ordenadores (computadoras) y dispositivos móviles, no tienen coste y son un buen recurso educativo.

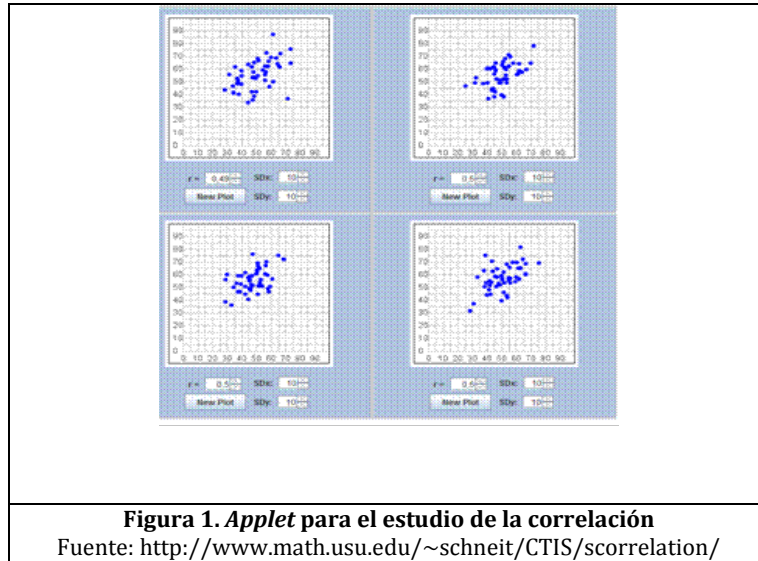
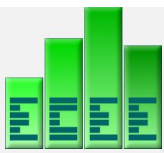
Cada uno de estos recursos se estudió con detalle por medio del análisis semiótico, para identificar los objetos matemáticos propuestos por Godino (2002) (conceptos, propiedades, lenguaje, procedimiento, argumentos), que el alumno debe usar en su trabajo. También se realizaron propuestas de actividades con el recurso.

A continuación se presenta el análisis detallado de uno de estos recursos y se comenta brevemente sobre otros disponibles.

RECURSOS EN INTERNET PARA EL ESTUDIO DE LA CORRELACIÓN

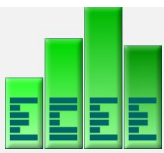
El recurso mostrado en la Figura 1, de acceso libre en Internet, presenta varias posibilidades:

- Variando el valor del coeficiente de correlación se obtienen nubes de puntos con dicho valor del coeficiente, bien en forma de animación (varía aleatoriamente el valor del r , desde -1 hasta 1); o a petición del usuario (se da un valor de r y el *applet* produce un diagrama con dicho valor de la correlación). Se puede también cambiar el valor de la desviación típica de las variables x e y , observando cómo afectan éstas la tendencia de la nube de puntos, para r fijo.
- Se pueden comparar sobre la misma pantalla hasta cuatro diagramas de dispersión simultáneos. Para cada diagrama el usuario puede ajustar el valor de r o la desviación típica de una o las dos variables, para comparar el efecto sobre los diagramas.
- El *applet* presenta un diagrama de dispersión y pide estimar el coeficiente de correlación r ; una vez se da la estimación, el *applet* proporciona el valor correcto.
- Se presentan varios diagramas y varios coeficientes de correlación que hay que emparejar; se puede trabajar con tres niveles de dificultad y como siempre, al final se da la respuesta correcta.
- Se pide ordenar cuatro gráficos, en función de su correlación, proporcionando también una lista con los valores de los coeficientes.



Tipo de objeto	Objetos matemáticos implicados en cada <i>applet</i>
Lenguaje	L1.Verbal: palabras como 'correlación', 'ordenar', 'gráfico'
	L2.Simbólico: símbolos como r, sd
	L4.Gráfico: diagramas de dispersión
	L5.Numérico: valores del coeficiente y desviación típica
Conceptos	C1.Variable estadística bidimensional; distribución
	C2.Dependencia funcional/estadística/independencia
	C3.Correlación
	C4.Medidas de tendencia central y posición
	C5.Medidas de dispersión: rango, desviación típica
Propiedades	P1. Correlación: signo y sentido de variación
	P2.Intensidad de la dependencia y valor de la correlación
	P3.Dependencia lineal y no lineal
	P4. Dispersión de la nube, correlación
Argumentos	A1.Análisis de ejemplos y contraejemplos
	A3.Visualizaciones
	A4.Generalización
	PC2.Representación gráfica
	PC3.Traducción entre representación gráfica y numérica
	PC4. Estimación e interpretación de la correlación
	PC7.Uso de tecnología
Problemas	P1.Analizar la existencia de relación entre variables
	P2. Diferenciar correlación directa e inversa
	P3. Ordenar correlación según intensidad

Tabla 1.Análisis semiótico del *applet* sobre correlación



En la Tabla 1 presentamos el análisis semiótico del *applet*, donde se presentan los objetos matemáticos que se trabajan a lo largo de las actividades descritas. Observamos que el recurso permite trabajar los principales objetos ligados al tema, partiendo de problemas como analizar la dependencia, diferenciar entre correlación directa o inversa y ordenar correlaciones. Se utilizan conceptos como variable estadística bidimensional, distribución, medidas de posición central y dispersión, dependencia y sus tipos, y correlación, así como las propiedades y procedimientos asociados, lo que favorece el aprendizaje de las ideas fundamentales en el tema. Así mismo, los tipos de argumentos como visualización, generalización y análisis de ejemplos, favorecen el desarrollo del razonamiento estadístico.

En la Tabla 2 listamos otros recursos útiles para la enseñanza de la correlación y regresión. Con ellos el profesor puede proponer actividades similares a las descritas o reforzar el aprendizaje de conceptos y propiedades.

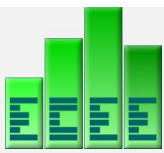
Fuente	Dirección de la página web
Aula virtual bioestadística	http://e-stadistica.bio.ucm.es/mod_regresion/regresion_applet_ghost.html
Biblioteca Nacional de Manipuladores Virtuales	http://nlvm.usu.edu/es/nav/frames_asid_144_g_4_t_5.html?open=activities&from=category_g_4_t_5.html
NCTM	http://illuminations.nctm.org/Activity.aspx?id=4187
Roger Webster	http://www.stat.sc.edu/~west/javahtml/Regression.html
Rossmann y Chance	http://www.rossmanchance.com/applets/guesscorrelation/GuessCorrelation.html
Stat Istics	bcs.whfreeman.com/ips4e/cat_010/applets/CorrelationRegression.html
Thales-CICA	thales.cica.es/rd/Recursos/rd99/ed99-0527-04/regresion.html
W. H. Freeman	bcs.whfreeman.com/ips4e/cat_010/applets/CorrelationRegression.html

Tabla 2. Otros recursos para el estudio de la correlación

CONCLUSIONES

La descripción realizada del sentido de la correlación y de sus componentes, puede motivar a los profesores para tratar de desarrollar dicho sentido en los estudiantes, de forma que adquieran una comprensión profunda del tema. Pensamos que los recursos descritos disponibles en Internet y otros semejantes, junto con las actividades sugeridas puedan ayudar al profesorado en esta labor.

También se podrían complementar estas actividades con los ejercicios propuestos en los libros de textos y pequeños proyectos estadísticos, en los que los estudiantes recogiesen sus propios datos para analizar la correlación entre ellos. Por ejemplo, los datos de medidas físicas de los estudiantes, como peso, talla, perímetro de cintura, longitud del antebrazo o la pierna, son fáciles de tomar y les permitirían realizar una introducción al tema; analizando las diferencias entre chicos y chicas, se podría conseguir enlazar con otros contenidos, como las distribuciones univariantes e incluso una introducción intuitiva a la inferencia.



En este sentido, conviene también proponer tareas con datos tomados de Internet (por ejemplo, de institutos de estadística), ya que permitirán al estudiante adquirir un sentido profundo de lo que implica un estudio estadístico de fenómenos sociales, tal y como se propone en las concreciones curriculares de diversos países, como en España (MEC, 2007). Un ejemplo detallado se presenta en Batanero, Díaz y Gea (2011).

Agradecimiento

Proyecto EDU2010-14947 (MICINN-FEDER) y grupo FQM126 (Junta de Andalucía).

REFERENCIAS

- Barbancho, A.G. (1973). *Estadística elemental moderna*. Barcelona: Ariel.
- Batanero, C., Díaz, C., Contreras, J.M. y Roa, R. (2013). El sentido estadístico y su desarrollo. *Números*, 83, 7-18.
- Batanero, C., Díaz, C. y Gea, M.M. (2011). Estadísticas de la pobreza y desigualdad. En C. Batanero y C. Díaz (Eds.), *Estadística con proyectos* (pp. 97-124). Granada: Departamento de Didáctica de la Matemática, Universidad de Granada.
- Burrill, G., y Biehler, R. (2011). Fundamental statistical ideas in the school curriculum and in training teachers. En C. Batanero, G. Burrill y C. Reading (Eds.), *Teaching statistics in school mathematics. Challenges for teaching and teacher education. A joint ICMI and IASE study* (pp. 57-69). Dordrecht, The Netherlands: Springer.
- Contreras, J.M. (2009). Recursos en Internet para la enseñanza de la probabilidad condicionada. Tesis de maestría. Granada: Universidad de Granada.
- Estepa, A. (1994). Concepciones iniciales sobre la asociación estadística y su evolución como consecuencia de una enseñanza basada en el uso de ordenadores. Tesis de doctorado. Granada: Universidad de Granada.
- Godino, J. (2002). Un enfoque ontológico y semiótico de la cognición matemática. *Recherches en Didactique des Mathématiques*, 22 (2-3), 237-284.
- Holmes, P. (2001). Correlation: From picture to formula. *Teaching Statistics*, 23 (3), 67-71.
- MEC (2007). *Real Decreto 1467/2007, de 2 de noviembre, por el que se establece la estructura del bachillerato y se fijan sus enseñanzas mínimas*. Madrid: Autor.
- Sánchez Cobo, F.T., Estepa, A. y Batanero, C. (2000). Un estudio experimental de la estimación de la correlación a partir de diferentes representaciones. *Enseñanza de las Ciencias*, 18 (2), 297-310.
- Wild, C. y Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, 67 (3), 223-265.