



# El uso de gráficos interactivos en Excel para facilitar la comprensión de conceptos básicos de Estadística<sup>1</sup>



**Vicente Coll Serrano**

Profesor del Departament d'Economia Aplicada de la Universitat de València  
[vicente.coll@uv.es](mailto:vicente.coll@uv.es)



**Olga Blasco Blasco**

Profesora del Departament d'Economia Aplicada de la Universitat de València  
[olga.blasco@uv.es](mailto:olga.blasco@uv.es)

| Fecha presentación: 05/04/2010 | Aceptación: 16/07/2010 | Publicación: 17/12/2010

## Resumen

En este trabajo se muestra una serie de aplicaciones realizadas en Excel que pretenden apoyar la docencia de la asignatura de Estadística. Concretamente, se presentan distintos gráficos interactivos que persiguen reforzar el contenido desarrollado en una clase y facilitar la comprensión conceptual por parte del estudiante de los principales aspectos analizados. Los estudiantes pueden relacionar, fácilmente, la teoría con una representación gráfica dotada de cierto grado de dinamicidad.

**Palabras clave:** gráfico dinámico, Excel, Estadística

## Resum

En aquest treball es mostra una sèrie d'aplicacions realitzades en la fulla de càlcul Excel que pretenen recolzar la docència de l'assignatura de Estadística. Concretament, es mostren diversos gràfics interactius que persegueixen reforçar el contingut desenvolupat en una classe i facilitar la comprensió conceptual per part de l'estudiant dels principals aspectes analitzats. Els estudiants poden relacionar, fàcilment, la teoria amb una representació gràfica dotada d'un cert grau de interactivitat.

**Paraules clau:** gràfic interactiu, Excel, Estadística

## Abstract

In this paper we offer several applications made in Excel which are intended to support the teaching of the subject Statistics. Specifically, various interactive graphics are shown which seek to strengthen the content developed in a class and facilitate conceptual understanding by the student of the main aspects analyzed. Students can, therefore, easily relate theory to practice by means of a dynamic graphical representation.

**Keywords:** interactive graphic, Excel, Statistics

<sup>1</sup> Una versió prèvia i reduïda de este treball se presentó a la I Jornada de Intercanvi de Experiències de Innovació Educativa en Estadística, celebrades en la Facultat de Economia de la Universitat de València el 20 de enero de 2010, y organizadas por el departamento de Economía Aplicada, el Servei de Formació Permanent y la Unitat d'Innovació Educativa.



## 1. Introducción

Tradicionalmente, la enseñanza ha estado centrada básicamente en la figura del profesor. Sin embargo, de acuerdo con la filosofía que subyace en el proceso de Bolonia y la construcción del Espacio Europeo de Educación Superior (EEES), el estudiante es el verdadero protagonista del proceso enseñanza-aprendizaje (Cal y Verdugo, 2009), lo que necesariamente debe implicar cambios en las estrategias y diseños docentes, que deben estar orientados al desarrollo de competencias (conocimientos, habilidades y actitudes), y subrayar el papel activo y responsable del estudiante en su propio proceso de aprendizaje.

Para conseguir los objetivos propuestos por el proceso de Bolonia, la Universitat de València, a través del Vicerrectorado de Convergencia Europea, presentó en el curso 2003-04 un proyecto de experimentación e innovación metodológica, que fue desarrollado mediante las convocatorias de Innovación Educativa. Al amparo de esta primera convocatoria y, con posterioridad, con financiación del Ministerio de Educación y Ciencia, en la Facultad de Economía se llevan a cabo dos Proyectos de Innovación Educativa (PIE) -PIE de Economía y doble titulación ADE-Derecho-, cuyos ejes fundamentales pueden sintetizarse de los siguientes (Dasí et al., 2007):

- Potenciar el trabajo del estudiante
- Atención a los objetivos competenciales
- Coordinación entre el profesorado de un mismo curso
- Atención personalizada (tutorías de transición)
- Incorporación de nuevas tecnologías

Relacionado con este último punto, la introducción de las TIC puede ayudar a transmitir de una forma más dinámica los conocimientos que tradicionalmente han sido expuestos mediante clases magistrales, al tiempo que permiten reforzar las competencias<sup>2</sup> que se quiere que desarrollen los estudiantes (Barroso, 2007; Ferro et al., 2009). Así, la combinación, entre otras, de clase magistral participativa, utilización de la metodología concepto-aplicación o de resolución de problemas y TIC puede contribuir a un cambio en la motivación y en la actitud del estudiante hacia la asignatura de Estadística, en la percepción que tiene sobre la misma -de forma que perciba que se trata de una materia útil en el ámbito de los estudios que cursa y en su futuro desarrollo profesional- y, en consecuencia, a una mejora de los resultados académicos. Con tal finalidad, en este artículo se muestran diversos gráficos interactivos diseñados en Excel que pueden utilizarse para reforzar el contenido teórico desarrollado en una clase y facilitar la comprensión conceptual por parte del estudiante de los principales aspectos analizados, al permitir relacionar la teoría con una representación gráfica dotada de cierto grado de interactividad.

## 2. Aprendizaje de Estadística y el uso de la hoja de cálculo Excel.

En el nuevo contexto de enseñanza-aprendizaje en que se desarrolla la actividad del docente, marcado por una reduc-

ción de presencialidad, la potenciación del trabajo autónomo del estudiante y el diseño de sistemas de evaluación y autoevaluación acordes con las metodologías docentes puestas en práctica, la incorporación de las nuevas tecnologías de la información y la comunicación es una herramienta esencial para orientar al estudiante en el aprendizaje efectivo y el logro de las competencias a adquirir por los estudiantes, resultado del proceso de aprendizaje.

En el ámbito del aprendizaje de la Estadística, la hoja de cálculo Excel “puede convertirse en una poderosa herramienta para crear ambientes de aprendizaje que enriquezcan la representación (modelado), comprensión y solución de problemas (...) y ofrece funcionalidades que van más allá de la tabulación, cálculo de fórmulas y graficación de datos, permitiendo crear y hacer uso de simulaciones que posibilitan a los estudiantes para realizar representaciones que permiten construir un puente entre las ideas intuitivas y los conceptos formales” (López et al., 2006).

El uso de Excel puede favorecer el proceso de aprendizaje de la Estadística al: (a) trasladar el contenido teórico estudiado al ámbito de la aplicación práctica sin necesidad de tener que realizar, cuando se trabaja con gran cantidad de datos, los engorrosos cálculos necesarios para obtener las medidas estadísticas requeridas; y (b) promover el tratamiento de forma autónoma y crítica de información, al poder dedicar más tiempo al proceso de planificación de resolución de los problemas y a la interpretación económico-empresarial de los resultados obtenidos.

Si bien Excel no es un programa propiamente estadístico y puede plantear algunos problemas de cálculo (MacCullough y Wilson, 2005; Heiser, 2006; Simonoff, 2008; MacCullough y Heiser, 2008), muchos de los conceptos tratados en el primer curso de la asignatura de Estadística que se imparte en la Facultad de Economía (Universitat de València), y gran parte de la materia de segundo curso, pueden resolverse a través de las diversas funciones estadísticas que ofrece esta hoja de cálculo o mediante las diferentes opciones que proporciona la herramienta de Análisis de Datos (Nash, 2008; McClave et al., 2008; Levine et al., 2010). Sin embargo, en determinadas ocasiones una representación gráfica puede facilitar la comprensión de los conceptos tratados. Por esta razón, para facilitar el proceso de enseñanza-aprendizaje de *Estadística I*, se han diseñado en Excel una gran variedad de gráficos interactivos. Se trata de gráficos que, a diferencia de los que se usan habitualmente, cambian, modifican su estructura, en función de los datos disponibles, y/o están dotados de cierto grado de dinamicidad, de movimiento.

A continuación se presentan, a modo de ejemplo, algunos de los gráficos dinámicos utilizados actualmente en la docencia de *Estadística I* para reforzar los contenidos explicados en clase relacionados con dos de los principales temas de la asignatura: el análisis de la regresión lineal y la distribución normal.

### 2.1 Análisis de regresión lineal

En el tema dedicado al análisis de regresión, en el que se hace especial énfasis en el caso lineal, se destacan tres objetivos de aprendizaje básicos: (i) suponiendo la existencia

<sup>2</sup> Competencia es el conjunto interrelacionado e interdependiente de conocimientos (saber los conocimientos requeridos para desempeñar adecuadamente una actividad y/o para generar nuevo conocimiento, capacidad de conocer y comprender), destrezas-habilidades (saber cómo hacer, ya sea para resolver problemas o para realizar cualquier tipo de tarea, académica o no, aptitudes), actitudes (saber cómo estar para adaptarse, participar y contribuir al desarrollo sostenible de su entorno) y valores (saber cómo ser, asumiendo los valores como parte integrante de la forma de ser, de percibir a los otros y de vivir en un contexto social y ambiental) (González y Wagenaar, 2003).

de dependencia (estadística) entre las variables objeto de estudio, determinar la forma o estructura que mejor ajuste a dicha relación; (ii) aplicar el modelo estimado para generar pronósticos o previsiones; y (iii) ser capaz de evaluar la calidad o bondad del ajuste realizado.

Relacionado con el primero de los objetivos de aprendizaje, con la regresión (lineal) se pretende explicar el comportamiento de una variable (dependiente o endógena) en función de la otra variable (independiente o exógena) a través de un determinado tipo de función (lineal), teniendo como soporte de la relación causa-efecto el marco teórico que proporciona la literatura<sup>3</sup>. Para estimar la recta que mejor ajusta a la nube de datos suele utilizarse el método de mínimos cuadrados ordinarios (MCO), que minimiza la suma de los residuos o errores al cuadrado. Las expresiones de las rectas de regresión que se obtienen son:

$$\text{Recta de regresión de Y/X: } Y^* - \bar{y} = \frac{S_{XY}}{S_X^2} (X - \bar{x}) \quad [1]$$

$$\text{Recta de regresión de X/Y: } X^* - \bar{x} = \frac{S_{XY}}{S_Y^2} (Y - \bar{y}) \quad [2]$$

donde:  $Y^*$ ,  $\bar{y}$  y  $S_Y^2$  son, respectivamente, el valor pronosticado o predicho, la media y la varianza de Y. De manera análoga,  $X^*$ ,  $\bar{x}$  y  $S_X^2$  representan el valor pronosticado o predicho, la media y la varianza de la variable X, respectivamente; y  $S_{XY}$  es la covarianza entre X e Y.

Para evaluar la fiabilidad del ajuste efectuado, y vinculado con el tercer objetivo del tema, se justifica el cálculo del coeficiente de determinación ( $R^2$ ), que mide la proporción de la varianza de la variable dependiente que explica la variable independiente en el modelo de regresión.

Para familiarizarse tanto con los procedimientos de cálculo implicados en la regresión como con la interpretación de los parámetros estimados, a medida que se van introduciendo los conceptos, los estudiantes realizan diversas prácticas a par-

tir de un conjunto de datos reducido. Prácticas más completas son realizadas utilizando la hoja de cálculo Excel que se ha diseñado con este propósito, y que se ilustra en la Figura 1.

En esta hoja de cálculo (Figura 1), para el conjunto de datos introducido a partir de la fila 6 en las columnas A (variable X) y B (variable Y), se muestran: (i) las ecuaciones de las rectas de regresión de X/Y y de Y/X; (ii) sus representaciones gráficas; y (iii) el coeficiente de determinación.

Las principales diferencias entre un diagrama de dispersión tradicional, al que se puede agregar una línea de tendencia, y el de la Figura 1 son que éste último: (a) incorpora la recta de regresión de X/Y; (b) es un gráfico interactivo, porque cuando el conjunto de datos cambia, la representación gráfica de las rectas, las ecuaciones estimadas y el coeficiente de determinación cambian, mostrando así la nueva estructura ajustada a los nuevos datos; y (c) se ha automatizado el título del gráfico y de los ejes de valores.

Otro de los aspectos que se estudian en la regresión es la relación que puede establecerse entre el coeficiente de correlación lineal y las rectas de regresión estimadas. Despejando la covarianza de la expresión del coeficiente de correlación y sustituyendo en [1] y [2], las rectas de regresión en función del coeficiente de correlación pueden escribirse, respectivamente, de la siguiente forma:

$$\text{Recta de regresión de Y/X: } Y^* - \bar{y} = \frac{r_{XY} S_Y}{S_X} (X - \bar{x}) \quad [3]$$

$$\text{Recta de regresión de X/Y: } X^* - \bar{x} = \frac{r_{XY} S_X}{S_Y} (Y - \bar{y}) \quad [4]$$

Donde  $r_{XY}$  es el coeficiente de correlación lineal entre las variables X e Y.

A partir de las expresiones [3] y [4], si se considera que:

· No existe relación lineal ( $r_{XY} = 0$ ); las rectas de regresión son perpendiculares.

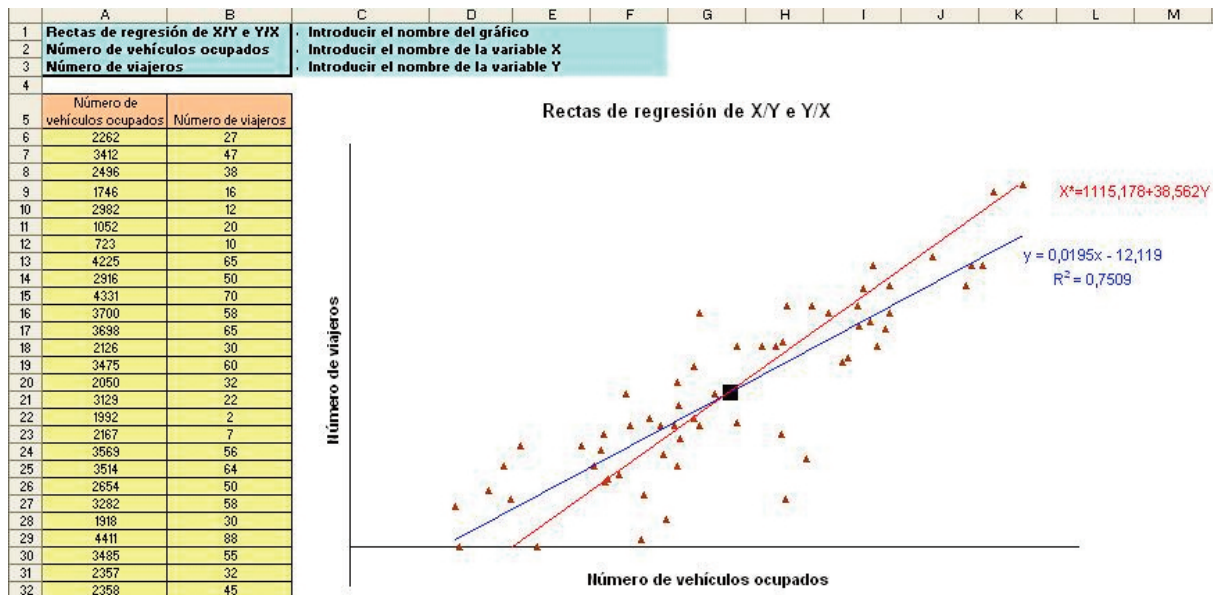


Figura 1. Rectas de regresión

<sup>3</sup> La relación causal entre dos variables debe fundamentarse en la teoría. La regresión resume la información que contienen los datos y los resultados obtenidos pueden apoyar o no la teoría (Newbold et al., 2007).

- La relación lineal es perfecta ( $r_{XY} = \pm 1$ ); las rectas de regresión coinciden.

Entre estas dos situaciones extremas, la intensidad de la relación lineal entre dos variables puede ser mayor o menor (positiva o negativa), cruzándose siempre las rectas de regresión en el centro de gravedad.

Para ilustrar estos conceptos, en la Figura 2 se muestran distintas representaciones gráficas que tratan de capturar la idea presentada en los dos párrafos anteriores; sin embargo, no puede percibirse la dinamicidad incorporada en el gráfico.

Para conseguir el efecto dinámico de *apertura de las rectas*, hasta llevarlas a la posición de perpendicularidad ( $r_{XY} = 0$ ), o de *cierre* de las mismas, hasta hacerlas coincidir ( $r_{XY} = \pm 1$ ), el gráfico está dotado con una barra de desplazamiento horizontal que actúa a modo de control; permitiendo aumentar o disminuir la intensidad de la fortaleza de la relación lineal entre las variables (coeficiente de correlación), lo que cambia la estructura de los datos y, por consiguiente, la representación de las rectas. El estudiante, junto con este efecto, puede observar que las rectas de regresión siempre se cortan en el centro de gravedad.

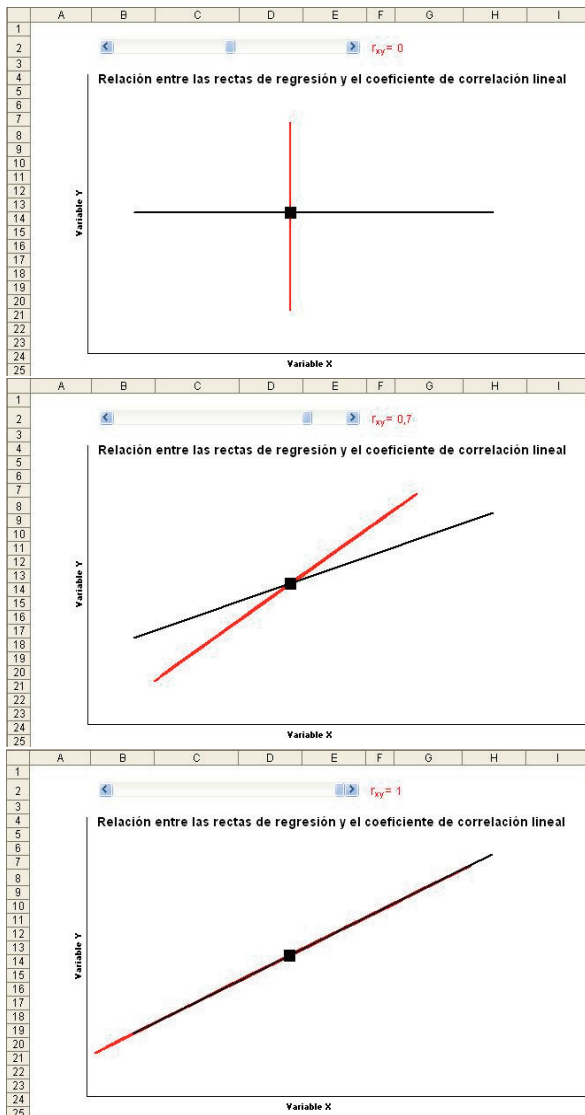


Figura 2. Relación entre el coeficiente de correlación y las rectas de regresión

### 2.2 Distribución Normal

Dentro del programa de *Estadística I*, en el tema 8 se trata la distribución normal, sin duda la distribución de probabilidad más importante y de más aplicación en Estadística. El estudio de la distribución normal se inicia justificando las razones por las que se utiliza con tanta frecuencia (Newbold et al., 2007) y con la siguiente definición: La variable aleatoria de tipo continuo X sigue una distribución normal con media  $\mu$  y desviación típica  $\sigma$  si X tiene una distribución de probabilidad cuya función de densidad es:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \forall x \in \mathfrak{R} \quad [5]$$

y se denota por  $X \sim N(\mu, \sigma)$ , donde el parámetro  $\mu$  es la media de la variable aleatoria y el parámetro  $\sigma$  es su desviación típica.

A continuación se señalan, utilizando la representación gráfica de la función de densidad, las principales características de la distribución normal (véase Esteban et al., 2005) y se destaca que no hay una única distribución normal, sino una familia completa de distribuciones que resultan de especificaciones concretas de  $\mu$  y  $\sigma$ . Así, si cambia  $\mu$  la distribución normal, manteniendo la forma, se desplaza a la

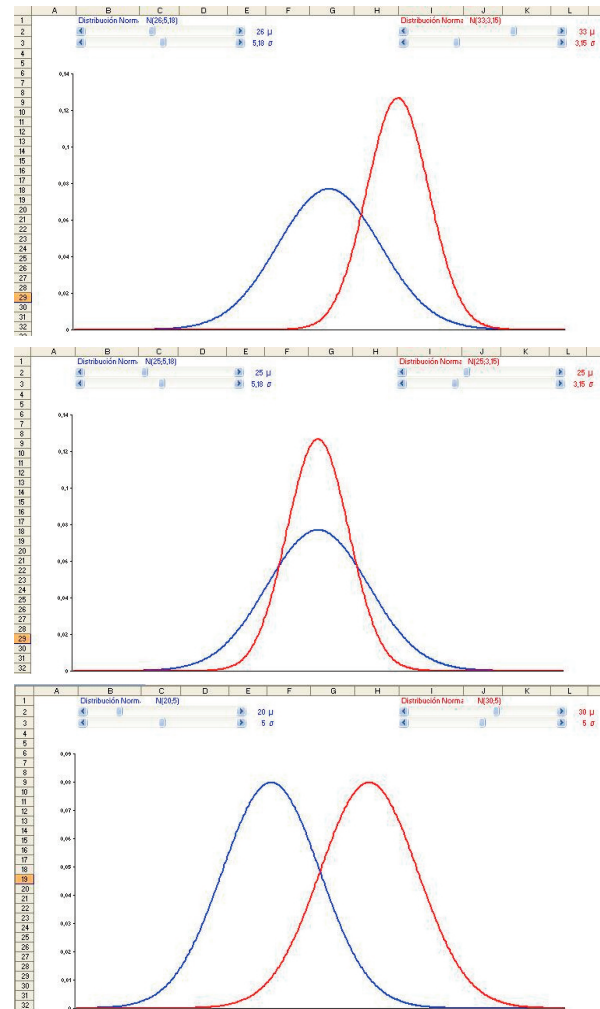


Figura 3. Efecto de cambio de  $\mu$  y  $\sigma$  en la distribución Normal



derecha o a la izquierda dependiendo de que la media aumente o disminuya, y si cambia el parámetro  $\sigma$  se modifica la amplitud de la distribución. Para ilustrar el efecto que los cambios en estos parámetros tiene en la función de densidad normal, se ha diseñado un gráfico interactivo (Figura 3) en el que se representan dos distribuciones normales, cuya forma puede ser modificada a través de dos barras horizontales que controlan los parámetros  $\mu$  y  $\sigma$ .

En la gráfica superior de la Figura 3 se representan dos distribuciones normales con distinta media y distinta dispersión; en la gráfica central, las normales tienen la misma media pero distinta desviación típica, la normal azul tiene mayor amplitud que la roja; y, finalmente, en la gráfica inferior las normales tienen distinta media y misma dispersión.

El uso de los controles vinculados con las dos distribuciones normales dota a la gráfica de movimiento, de forma que su manipulación por parte del estudiante facilita que éste comprenda la idea que se pretendía transmitir: no hay una única distribución Normal sino una familia completa de distribuciones que resultan de especificaciones concretas de  $\mu$  y  $\sigma$ .

### 3. Conclusiones

La hoja de cálculo Excel ofrece una gran cantidad de funciones estadísticas y opciones de análisis de datos que posibilitan abordar, a nivel de aplicación práctica, muchos de los conceptos que se estudian en la asignatura *Estadística I*. Así, el uso de Excel por parte de los estudiantes para la resolución de prácticas evita que éstos tengan que realizar manualmente los cálculos, que en muchas ocasiones resultan farragosos y, por tanto, permite prestar más atención al diseño de la estrategia de resolución del problema y a la interpretación (económico-empresarial) de los resultados obtenidos.

La contribución de Excel a la resolución de actividades prácticas es clara. Sin embargo, suele descuidarse la potencialidad que esta misma herramienta ofrece para facilitar que los estudiantes comprendan un determinado concepto teórico. En este sentido, es habitual que en la exposición de una determinada idea se utilicen gráficas para facilitar al estudiante su comprensión al promover la relación visual entre ambas (idea-gráfica). Frente a las tradicionales gráficas, estáticas, que suelen utilizarse con esta finalidad, en este trabajo se han mostrado diversos ejemplos de diseño de gráficos interactivos en Excel, y el contexto en el que se utilizan, cuya manipulación, por parte del profesor y/o estudiante, permite la modificación de la estructura de los datos y presentar situaciones diversas. Otros gráficos dinámicos que se utilizan en clase para apoyar el proceso enseñanza-aprendizaje de conceptos estadísticos relativamente complejos, y que no han sido presentados en este trabajo por cuestión de espacio, son, entre otros, la función de densidad de probabilidad, la convergencia de la distribución binomial a la distribución normal o la aproximación de la distribución de Poisson a la normal.

### 4. Bibliografía

- Barroso, Carlos (2007). La incidencia de las TICs en el fortalecimiento de hábitos y competencias para el estudio. *EduTec: Revista electrónica de tecnología educativa*, 23. <http://edutec.rediris.es/Revelec2/revelec23/cbarroso/EDUTEC2007-CarlosBarroso.pdf> Fecha de última consulta, 08.11.2010.
- Cal, M<sup>a</sup>. Isabel; Verdugo, M<sup>a</sup>.Victoria (2009). Una experiencia b-learning en Econometría. *Revista de Formación e Innovación Educativa Universitaria*, 2 (4), pp. 227-233. [http://webs.uvigo.es/refiedu/Refiedu/Vol2\\_4/arti\\_2\\_4\\_5.pdf](http://webs.uvigo.es/refiedu/Refiedu/Vol2_4/arti_2_4_5.pdf)
- Dasí, Angels et al. (2007). *Innovación Educativa en la Universidad: ADE-Derecho*. Valencia: Publicaciones de la Universidad de Valencia.
- Esteban, Jesús et al. (2005). *Estadística Descriptiva y nociones de Probabilidad*. Madrid: Thomson.
- Ferro, Carlos; Martínez, Ana Isabel; Otero, M<sup>a</sup>. Carmen (2009). Ventajas del uso de las TICs en el proceso de enseñanza aprendizaje desde la óptica de los docentes universitarios españoles. *EduTec: Revista Electrónica de Tecnología Educativa*, 29. [http://edutec.rediris.es/Revelec2/revelec29/edutec29\\_ventajas\\_TIC\\_docentes\\_universitarios.html](http://edutec.rediris.es/Revelec2/revelec29/edutec29_ventajas_TIC_docentes_universitarios.html) Fecha de última consulta, 08.11.2010.
- González, Julia; Wagenaar, Robert (eds) (2003). *Tuning Educational Structures in Europe*. Bilbao: Universidad de Deusto.
- Heiser, D. A., (2006). Microsoft Excel 2000 and 2003 faults, problems, workarounds and fixes. *Computational Statistics & Data Analysis*, 51 (2), pp. 1442-1443.
- Levine, David M., et al. (2010). *Statistics for Managers using MS Excel*. New Jersey: Prentice Hall.
- López, Mirna; Lagunes, Cristina; Herrera, Santa (2006). Excel como una herramienta asequible en la enseñanza de la Estadística. *Teoría de la Educación*, 7. [http://www.usal.es/~teoriaeducacion/rev\\_numero\\_07/n7\\_art\\_lopez\\_lagunes\\_herrera.htm](http://www.usal.es/~teoriaeducacion/rev_numero_07/n7_art_lopez_lagunes_herrera.htm) Fecha de consulta: 05.04.2010.
- McClave, James T.; Benson, P. George; Sincich, Terry (2008). *Statistics for business and economics*. New Jersey: Prentice Hall.
- MacCullough, B.D. y Wilson, B. (2005). On the accuracy of statistical procedures in Microsoft Excel 2003. *Computational Statistics & Data Analysis*, 44, pp.1244-1252.
- MacCullough, B.D. y Heiser, D. A. (2008). On the accuracy of statistical procedures in Microsoft Excel 2007. *Computational Statistics & Data Analysis*, 52, pp.4570-4578.
- Nash, John C., (2008). Teaching statistics with Excel 2007 and other spreadsheets. *Computational Statistics and Data Analysis*, 52, pp. 4602-4606.
- Newbold, Paul; Carlson, William L.; Thorne, Betty (2007). *Estadística para administración y economía*. Madrid: Prentice Hall.
- Simonoff, Jeffrey S. (2008). Statistical analysis using Microsoft Excel. <http://pages.stern.nyu.edu/~jsimonof/classes/1305/pdf/excelreg.pdf> Fecha de consulta: 05.04.2010.

#### | Cita recomendada de este artículo

Coll Serrano, Vicente y Blasco Blasco, Olga (2010). El uso de gráficos interactivos en Excel para facilitar la comprensión de conceptos básicos de Estadística. @tic. revista d'innovació educativa. (nº 5). URL. Fecha de consulta, dd/mm/aaaa.