

Exploración de conceptos de probabilidad con Geogebra

Mercado Martínez, Miguel

miguellmm@unam.mx, FES Acatlán, CCH Naucalpan – UNAM

Resumen

Se presentan hojas de trabajo en GeoGebra para la exploración y análisis de conceptos básicos de probabilidad. Primero se destaca el papel creciente del software dinámico en el escenario educativo. Este software nos permite explorar de manera dinámica conceptos y procesos matemáticos que favorecen la construcción de significados matemáticos por parte de los alumnos. En el proceso de aprendizaje de conceptos de probabilidad, no basta con conocer las diferentes reglas de la misma, es necesario realizar actividades de simulación de experimentos. A pesar de que Geogebra es un software para álgebra y geometría, también se pueden realizar simulaciones para contribuir a la comprensión de conceptos probabilísticos. En este trabajo presentamos tres hojas de trabajo en Geogebra, relacionadas respectivamente con los conceptos 1) diagramas de Venn, probabilidad condicional y Teorema de Bayes, 2) el tablero de Galton y 3) la aproximación de la distribución normal a la binomial.

Palabras clave: Palabras clave: Software dinámico, Probabilidad, Bayes, Simulación, Distribución.

1. Introducción

Los ambientes de aprendizaje basados en tecnología computacional han tenido impacto en los diferentes dominios matemáticos, entre ellos, la aritmética, el álgebra, la geometría, la estadística, el cálculo, etc. (Balacheff & Kaput, 1996). Sin embargo, se reconoce que el acceso a estos dispositivos electrónicos no garantiza que todos los estudiantes vayan a adquirir una cultura matemática; son herramientas que pueden simplificar la tarea que tienen en sus manos, pero no la resuelven (NCTM, 1989). No obstante, se considera que los estudiantes de todos los niveles educativos deberían tener acceso a las computadoras para ser usadas cuando resuelvan problemas matemáticos (NCTM, 2000). En el bachillerato de la UNAM se reconoce que el conocimiento matemático está influido por los avances tecnológicos, entre otros; y que las nuevas tecnologías contribuyen a investigar, conjeturar y verificar modelos matemáticos en el alumno (Consejo Académico del Bachillerato-UNAM, 2001).

Así, la constante creación y renovación de los recursos tecnológicos para la enseñanza, particularmente la enseñanza de las matemáticas, han influido en gran medida en la enseñanza. Si bien es cierto que los paquetes de cómputo se realizan con cierta intencionalidad, el desarrollo de su potencial educativo depende de su funcionamiento informático y de la creatividad y destreza de los usuarios, investigadores y profesores, principalmente, en lo que se refiere al diseño de actividades de aprendizaje para los estudiantes.

Noss y Hoyles (1996) ven en la computadora una ventana para mirar la manera en que los estudiantes producen significados matemáticos. En este sentido se cree que deben ser buscadas y construir actividades a desarrollar con la computadora, actividades que permitan observar características importantes de los procesos que los estudiantes llevan a cabo para darle significado a los distintos conceptos y procesos matemáticos.

La presencia de calculadoras, computadoras y software educativo modifica el entorno de la acción docente y establece nuevas condiciones para el aprendizaje, en particular de las matemáticas.

En la actualidad se han desarrollado softwares Dinámicos, softwares que permiten la exploración matemática de manera dinámica, entre ellos se encuentran GeoGebra, Fathom, Cabri-Géomètre, Sketchpad, etc., aparentemente cada uno para la exploración en un campo matemático específico, por ejemplo, geometría, estadística, etc., sin embargo los campos de aplicación de algunos de los softwares es muy diversa, como lo observaremos con GeoGebra, que pareciera es para la exploración del algebra, geometría y calculo, donde exploraremos conceptos probabilísticos.

En Geometría, estos programas computacionales tienen herramientas que permiten la construcción de figuras geométricas muy diversas, desde representaciones simples como puntos y líneas, figuras más complicadas como triángulos, círculos y polígonos, hasta construcciones más elaboradas donde se pueden necesitar rectas paralelas, perpendiculares, simetrías, etc. Una vez realizada una construcción, el software tiene la facilidad de desplazar sobre la pantalla del monitor puntos, segmentos, o figuras geométricas (triángulos, círculos,...); también se pueden generar lugares geométricos y sus ecuaciones; por ejemplo, se puede observar qué figura describe el punto medio de una cuerda a un círculo cuando uno de los extremos de la cuerda se mueve alrededor del círculo. También es posible la exploración de funciones y sus gráficas, por ejemplo dada una función $y = f(x)$, se puede tabular para diferentes valores de x , presentarlos en una tabla o observar la gráfica que se genera, otra posibilidad es la de exploración del efecto de los distintos parámetros sobre la grafica de la función.

En estadística, a partir de una base de datos se puede obtener representaciones graficas como histogramas, polígonos de frecuencias, diagramas de dispersión, etc., encontrar estadísticas básicas como media, mediana, cuartiles, etc., o llevar a cabo inferencias sobre los datos, como Intervalos de Confianza y Pruebas de Hipótesis. Todo ello de manera dinámica, o sea, cuando se agregan o eliminan datos o bien se modifican uno o más de ellos todo se actualiza de inmediato. También es posible realizar la simulación de experimentos, por ejemplo, lanzamientos de monedas o dados, toma de muestras de poblaciones, etc.

Las acciones posibles y los retornos a las construcciones, se amplían y resultan de naturaleza diferente. Se integran conocimientos matemáticos y se amplía el campo de experimentación posible, se pueden esperar nuevas posibilidades de organización de las situaciones de aprendizaje así como cambios en las conductas de los estudiantes.

Los estudiantes pueden realizar variaciones precisas e instantáneas de sus propias representaciones visuales que se producen con el uso de este tipo de software (Balacheff & Kaput, 1996), lo que les permite, en principio, realizar muchas exploraciones, probar ideas y conjeturar en una forma visual, eficiente y dinámica. Una actividad importante en el estudio de las matemáticas es que el sujeto desarrolle estrategias que le permitan formular conjeturas.

Con lo anterior en mente, en este trabajo se presenta la exploración de algunos conceptos de Probabilidad, usando el software GeoGebra. Tomaremos como punto de partida, algunos ejemplos:

2. Diagramas de Venn

Ejemplo 1: Se examinaron las tarjetas de registro de 200 estudiantes en relación a ciertos idiomas. Se encontró que 100 estudiaban francés, 80 estudiaban español y 60 ambos idiomas o sea que $P(F) = 0.5$; $P(E) = 0.4$ y $P(E \cap F) = 0.3$. Se toma un estudiante al azar:

- Encontrar la probabilidad de que estudie francés o español.
- $P(f \cup e) = p(f) + p(e) - p(e \cap f) = 0.5 + 0.4 - 0.3 = 0.6$
- Encontrar la probabilidad de que de que no estudie francés ni español.

- d. $P(f \cup e)^c = 1 - p(f \cup e) = 1 - 0.6 = 0.4$, que también es $p(e^c \cap f^c) = 0.4$
 e. Si sabemos que estudia francés, ¿cuál es la probabilidad de que estudie español?

$$P(E|F) = \frac{P(F \cap E)}{P(F)} = \frac{20}{60} = 0.33.$$

- f. Si sabemos que estudia español ¿Cuál es la probabilidad de que estudie francés

$$P(F|E) = \frac{P(F \cap E)}{P(E)} = \frac{20}{80} = 0.25.$$

Los datos pueden ser colocados en una tabla; la Tabla 1 contiene los datos dados por el enunciado; los números resaltados con “negritas” son los valores dados, los otros números se ajustan a los totales marginales. En la Tabla 2 se presentan las probabilidades correspondientes a la tabla No 1, por ejemplo $P(E^C) = 0.6$ y $P(E \cap F^C) = 0.1$. A partir de la tabla se pueden verificar varias reglas de probabilidad

Por ejemplo: $P(E \cap F^C) = P(E) - P(E \cap F) = 0.1$, y se observa que $P(E^C \cap F^C) = 0.4$, que nos permite verificar indirectamente una de las leyes de Morgan, o sea *que* $(F \cup E)^c = E^c \cap F^c$.

Tabla 1. Datos del problema

	F	F ^C	Total
E	60	20	80
E ^C	40	80	120
Total	100	100	200

Tabla 2. Probabilidad de los eventos

	P(F)	P(F ^C)	Prob
P(E)	0.3	0.1	0.4
P(E ^C)	0.2	0.4	0.6
Prob	0.5	0.5	1.0

Los resultados se pueden ver en la gráfica de la Figura 1, que se realizó con GeoGebra. Lo que es de destacar, es que con la hoja de trabajo hecha en GeoGebra, es posible realizar variaciones a las probabilidades en el ejercicio y de inmediato se ajustan los valores que resultan al aplicar las correspondientes reglas de probabilidad, como se podrá experimentar al usar la hoja en GeoGebra.

En el programa que genera el diagrama se pueden controlar los datos correspondientes a: $P(A)$, $P(B)$ y $P(A \cap B)$, o sea, $P(F)$, $P(E)$ y $P(E \cap F)$ en el ejercicio, el programa regresa $P(A|B)$, $P(B|A)$, $P(A' \cap B')$; $P(F|E)$, $P(E|F)$, $P(F' \cap E')$ del ejercicio. En la figura No 1 los valores son: $P(A) = 0.5$, $P(B) = 0.4$ y $P(A \cap B) = 0.299$, los resultados con el programa son: $P(A|B) = 0.747$, $P(B|A) = 0.598$, $P(A' \cap B') = 0.399$.

Lo interesante de la hoja de trabajo en GeoGebra es que podemos hacer variaciones a los datos de entrada, $P(A)$, $P(B)$ y $P(A \cap B)$, al desplazar de izquierda a derecha los “deslizadores” que representan los valores de probabilidad, ajustándose los valores de $P(A|B)$, $P(B|A)$, $P(A' \cap B')$, como se observa de las figuras. En la Figura 2 los valores dados, son: $P(A) = 0.3$, $P(B) = 0.25$ y $P(A \cap B) = 0.08$, los resultados con el programa son: $P(A|B) = 0.267$, $P(B|A) = 0.321$, $P(A' \cap B') = 0.53$.

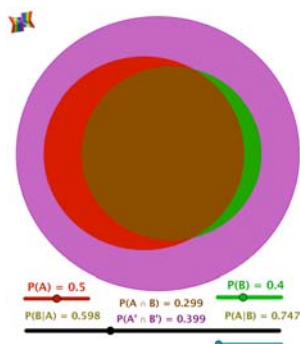


Figura 1. Diagrama de Venn

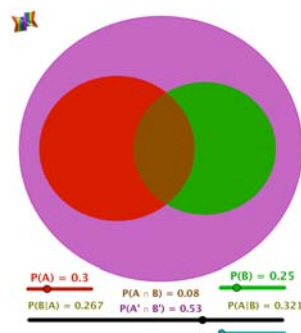


Figura 2. Diagrama de Venn

Ejemplo 2: Para una mujer americana de entre 40 y 50 años sin síntomas, tiene una probabilidad de cáncer de pecho de 8%. Si una mujer tiene cáncer la probabilidad de detectarlo es de 90%, aunque también el 7% de mujeres sanas dan positivo. Se toma una mujer americana entre 40 y 50 años, en forma aleatoria, si la prueba resulta positiva, ¿cuál es la probabilidad de que la mujer realmente tenga cancer de pecho? (Tomado de una presentación de la Dra. Carmen Batanero)

Los datos se pueden representar en el diagrama de arbol de la Figura 3:

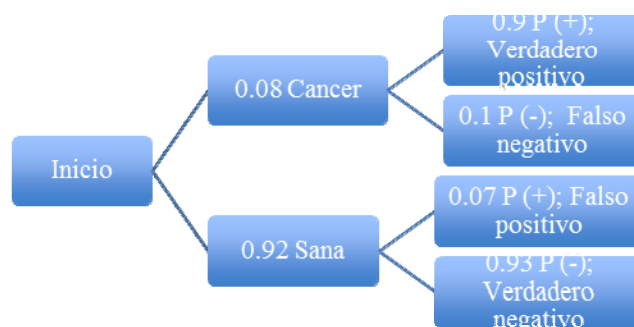


Figura 3. Diagrama de árbol del ejercicio 2

Si una persona es elegida al azar, la probabilidad de que al hacerse una prueba resulte positiva es: $P(+)=P(C)P(+/C)+P(S)P(+/S)=(0.08)(0.9)+(0.92)(0.07)=0.1364$

Mientras que si la prueba resulta positiva, la probabilidad de que realmente tenga cáncer es: $P(C/+)=\frac{P(C)P(+/C)}{P(+)}=\frac{(0.08)(0.9)}{0.1364}=0.5279$

También se puede realizar una tabla con los valores dados, Tabla 3, donde los valores de las intersecciones se obtienen de la regla de probabilidad condicional: $P(A/B)=\frac{P(A \cap B)}{P(B)}$, entonces $P(A \cap B)=P(B)P(A/B)$, por lo que $P(C \cap +)=P(C)P(+/C)=(0.08)(0.9)=0.072$. Esta situación se puede organizar en una hoja de trabajo con GeoGebra como se muestra en la Figura 4

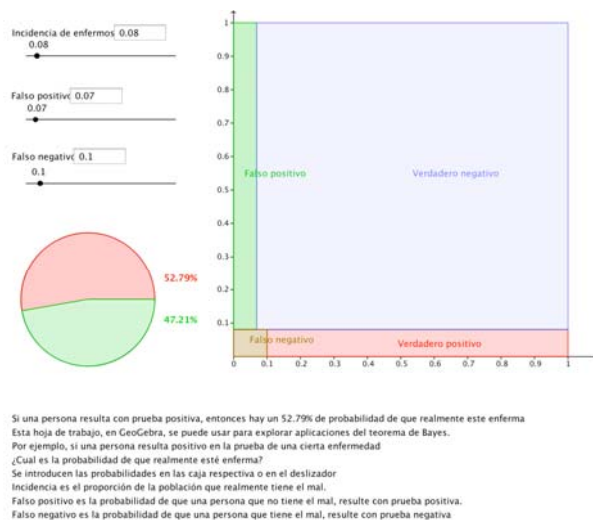


Figura 4. Aplicación del teorema de Bayes

En la anterior representación los datos de entrada son: incidencia de la enfermedad; $P(C)$, falso positivo, probabilidad de prueba positiva y no tener la enfermedad; $P(+/C')$, y falso negativo, probabilidad de prueba negativa y tener la enfermedad; $P(-/C)$. Estos datos de entrada pueden ser colocados en las vantage correspondientes o desplazando los deslizadores de izquierda a derecha, obteniéndose de salida: $P(-/C)$ y $P(+/C)$, información relevante en este tipo de problemas, en forma gráfica y numérica.

Tabla 3. Datos

	P(C)	P(C')	Total
P(+)	0.072	0.0644	0.1364
P(-)	0.008	0.8556	0.8636
Total	0.08	0.92	1

La hoja en GeoGebra permite la exploración de una gran cantidad de situaciones que involucran el teorema de Bayes, de manera dinámica, al desplazar los deslizadores de izquierda a derecha, como se podrá observar al interaccionar con la hoja y el GeoGebra.

3. Simulación de la Máquina de Galton

La máquina de Galton, también conocida como *quincence*, es un dispositivo inventado por Sir Francis Galton, para demostrar que la distribución binomial se aproxima a la distribución normal. Está constituida con una tabla con numerosas filas de clavos intercalados sucesivamente que evocan al triángulo de Pascal. Desde la parte superior se dejan caer bolas que botan a la izquierda o a la derecha, en forma aleatoria, cuando chocan con los clavos. En el fondo, las bolas son coleccionadas en recipientes verticales, figura No 5. Se cuenta el número de canicas en cada celda y se calcula la frecuencia relativa, se tiene una aproximación al valor de probabilidad que se obtiene al usar la distribución binomial

Se puede considerar que el número de canicas que caen en los recipientes verticales sigue una distribución binomial porque al chocar una canica con un clavo, i) puede desplazarse a la izquierda o a la derecha, éxito o fracaso, ii) la canica chocará con n clavos, la probabilidad de desplazamiento a la izquierda (I) o a la derecha (D) es la misma en cada clavo y los eventos son independientes ya que las trayectorias I D, no dependen unas de otras, y iii) X = número de desplazamientos a la derecha, aunque también podría ser hacia la izquierda.

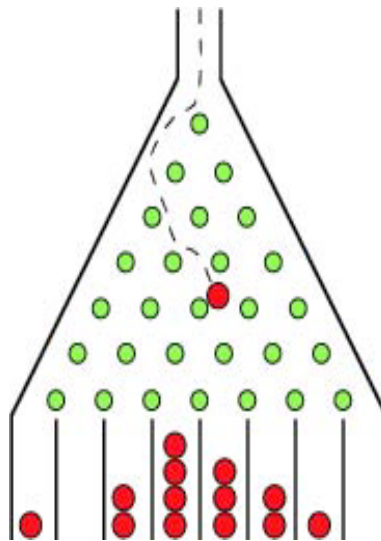


Figura 5. Máquina de Galton con 7 recipientes de canicas

En la Figura 5, Se puede deducir que el número de bolas que cae en las celdas del fondo sigue una distribución binomial con $n = 7$ y $p = 0.5$, con la variable aleatoria $X =$ número de canicas que se desplaza hacia la derecha al chocar con los clavos, que se puede tomar como número de canicas que cae en las celdas $0, 1, 2, \dots, 7$. Así, $x=0$ significa que la trayectoria de la canica fue I I I I I I I; $x=1$, posible trayectoria I I I I I D ó I I I I D I ó I I I I D I I ó I I I D I I I ó I I D I I I I ó I D I I I I I ó D I I I I I I; hasta $x=7$ donde la trayectoria es D D D D D D D, porque la canica, por igual, puede ir a la izquierda (I) o a la derecha (D) al chocar el clavo. La probabilidad para cada celda está dada por: $P(X = x_i) = \frac{n!}{x_i!(n-x_i)!} p^{x_i} q^{n-x_i}$. Entonces, la probabilidad de que la canica caiga en la celda cero, es: $P(X = 0) = \frac{7!}{0!(7-0)!} 0.5^0 0.5^{7-0}$; de que caiga en la celda uno, es $P(X = 1) = \frac{7!}{1!(7-1)!} 0.5^1 0.5^{7-1}$, y así sucesivamente. En un experimento se observará la frecuencia relativa del número de canicas que cae en cada celda y este valor se puede comparar con el valor que proporciona la distribución Binomial.

Máquina de Galton, simulada con Geogebra. La hoja de trabajo se simuló el lanzamiento de 250 bolas, Figura 6. En la parte inferior de la máquina se presenta un histograma construido con el número de bolas que cayó en cada celda, en la derecha se tiene la misma información en forma numérica, número de celda contra frecuencia. La figura presenta una ventana con la leyenda inicio, ahí se da inicio a la simulación, si hay datos de una simulación previa los elimina; otra ventana tiene la leyenda acción que permuta con pausa, con ella se da inicio, se detiene o reinicia una simulación; indica la trayectoria de una bola; contiene un deslizador donde se puede controlar la velocidad de caída de las bolas al moverlo de izquierda a derecha; por último tiene como opciones ver el histograma binomial teórico y la gráfica de la distribución normal, ambos construidos con los parámetros del ejercicio, lo que nos permite observar visualmente como el histograma que resulta de la simulación de la máquina de Galton se aproxima a la distribución binomial, así como la aproximación de la distribución normal a la binomial.

En la hoja GeoGebra, $n = 21$, $p = 0.5$, $q = 0.5$ y $x = 0, 1, 2, \dots, 21$, Los valores teóricos de la binomial se calculan con la fórmula $P(X = x_i) = \frac{n!}{x_i!(n-x_i)!} p^{x_i} q^{n-x_i}$. Podemos comparar valores teóricos con valores experimentales, por ejemplo:

$$P(X = 13) = \frac{21!}{13!(21-13)!} 0.5^{13} 0.5^{21-13} = 0.0970; \text{ con los valores de la hoja de trabajo}$$

$$P(X = 13) = \frac{284}{3030} = 0.0947; \text{ ó } P(X = 11) = \frac{21!}{11!(21-11)!} 0.5^{11} 0.5^{21-11} = 0.1682 \text{ contra}$$

$$P(X = 11) = \frac{196}{3030} = 0.1703$$



Figura 6. Simulación de la máquina de Galton con Geogebra

4. Aproximación de la distribución normal a la binomial

En el apartado anterior se pudo observar gráficamente como la distribución normal, con $\mu = np$ y $\sigma = \sqrt{npq}$, se aproxima a la distribución binomial. En general se sabe que cuando n es “grande” y p ó q no próximos a cero o la unidad, la aproximación es buena, más específicamente, cuando $\mu = np > 5$. Mediante la hoja de trabajo en GeoGebra que se presenta en este trabajo, es posible explorar la aproximación normal a la binomial, para valores $1 \leq n \leq 20$ y para valores $0 \leq p \leq 1$, en el entendido de que los valores de n son discretos.

En la hoja de trabajo en GeoGebra Figura 7, se pueden explorar distintas situaciones para los valores de n y p . Se puede observar como para valores n no muy grandes y p cercano a 0.5, el grafico de la binomial es muy simetrico y la normal se ajusta relativamente bien, pero a medida que p se acerca a cero o la unidad el grafico es sesgado y la aproximación no es buena. La hoja de trabajo contiene un deslizador para el valor de n y otro para p , que pueden deslizarse de izquierda a derecha ajustando a los valores desados en el rango antes dicho. Grafica un histograma para la binomial y tiene como opción una casilla para observar la aproximación normal, cuanta con otras opciones, reescalar el eje “y” así como señalar regiones críticas, finalmente tiene la opción de animación de p , animate p , que explora el gráfico para un valor de n determinado y p cambia cero a uno.

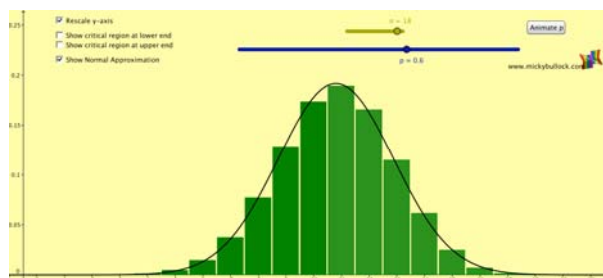


Figura 7. Aproximación de la distribución normal a la binomial

Se puede destacar que de acuerdo con Duval (1999), un estudiante comprende un concepto matemático, cuando es capaz representarlo en al menos dos registros de representación. En los ejemplos analizados hay varios registros, que son: lenguaje común, lenguaje simbólico, diagramas de Venn y tablas de contingencia.

Una de las características del GeoGebra es integrar tres diferentes registros de representación que se pueden usar simultáneamente, cuando se explora una situación de aprendizaje, estos son: algebraico y algo más, gráfico y hoja de cálculo. Se puede usar uno, o dos o los tres de manera simultánea.

Otro aspecto que es de relevancia en GeoGebra, es la capacidad que tiene para la simulación. En estadística “simulación” se puede entender como algo parecido al isomorfismo en otras ramas de las matemáticas. Consiste en poner en correspondencia dos experimentos aleatorios distintos. La condición es que a cada suceso elemental del primer experimento le corresponda uno y sólo un suceso elemental del segundo experimento, de manera que los sucesos puestos en correspondencia, en los dos experimentos, sean equiprobables. Lo importante de la simulación es que podemos operar y observar resultados del segundo experimento y utilizarlos para obtener información del primer experimento.

Por ejemplo, podemos simular el lanzamiento de una moneda con un generador de números aleatorios. Se puede generar números aleatorios entre 0 y 1, y se acuerda una regla; asignar cara si el número resulta uno, cruz si es cero; o en el nacimiento de un bebe mujer si resulta cero, varón si resulta uno. La ventaja de la simulación entonces permite condensar el experimento, además de que pueden ser desde muy sencillas como estos ejemplos, hasta procesos aleatorios complejos, como lo veremos a continuación.

Las aplicaciones anteriores vistas en este artículo, fueron recuperadas del portal de GeoGebra, www.geogebra.org, así mismo fueron adaptadas para esta presentación, con fines didácticos en la enseñanza de Probabilidad en el nivel del bachillerato. Es posible reproducir los ejemplos, instalando el software geogebra y llamar a los archivos adjuntos en el anexo.

Referencias

- Balacheff, N. y Kaput, J. (1996). Computer-based learning environments in mathematics. En A. J. Bishop et al. (Eds.) *International Handbook of Mathematics Education*. K.A.P., Netherlands.
- Consejo Académico del Bachillerato, UNAM (2001) *Núcleo de conocimientos y formación básicos que debe proporcionar el bachillerato de la UNAM; Primera Aproximación*, UNAM, México
- Duval, R. (1999). *Semiosis y pensamiento humano: Registros semióticos y aprendizajes intelectuales*. Universidad del Valle, Cali, Colombia.

GeoGebra. WWW.geogebra.org/cms/

National Council of Teachers of Mathematics. (1989). Curriculum and evaluation standards for the school mathematics. Reston, VA: NCTM.

National Council of Teachers of Mathematics. (2000). Principles and standards for school mathematics. VA: NCTM.

Noss, R. y Hoyles, R. (1996). Windows on mathematical meanings: learning cultures and computers. Kluwer Academic Publishers, Netherlands.