COMPARACIÓN DE DISTRIBUCIONES

¿Una Actividad Sencilla para los Futuros Profesores?[1]

Blanca Ruiz
ITESM, México
Pedro Arteaga y *Carmen Batanero*
Universidad de Granada, España

Resumen. En este trabajo analizamos los informes realizados por una muestra de 101 futuros profesores de educación primaria en un proyecto de análisis de datos. Estudiamos los gráficos producidos, estadísticos calculados y su interpretación en la obtención de conclusiones sobre el proyecto. Concluimos que el concepto de distribución no es comprendido en profundidad por una parte importante de los futuros profesores.

1. Introducción

La enseñanza de la estadística debe desarrollar la capacidad de leer, analizar, y hacer inferencias a partir de distribuciones de datos (Shaughnessy, 2007), pero este es un objeto matemático complejo que involucra las ideas de dato, variable estadística, valor, frecuencias, promedio, dispersión y forma. Todos estos conceptos se incluyen en el nuevo Decreto de Enseñanzas Mínimas para la Educación Primaria en España (MEC, 2006) donde la estadística se estudia desde el primer ciclo (niños de 6 y 7 años). Será importante, por tanto, asegurar la comprensión adecuada de la idea de distribución y conceptos relacionados por los profesores que tienen que enseñar este tema.

En este trabajo evaluamos la comprensión de un grupo de futuros profesores de Educación Primaria españoles sobre algunos objetos matemáticos básicos que subyacen en la idea de distribución, en particular sobre los gráficos estadísticos, medidas de posición central y dispersión. La evaluación se realiza a partir de sus producciones escritas en un proyecto abierto de análisis de datos en el que deben comparar pares de distribuciones. Se elige esta opción para la evaluación debido a que el trabajo con

[1] II Encontro de Probabilidade e Estatística na Scola. Universidade do Minho, 2009, Braga, Portugal.

proyectos es recomendado en los Decretos curriculares españoles para trabajar en estadística con los alumnos de Educación Primaria.

2. Comprensión de la idea de distribución

Uno de los tipos de análisis estadísticos más frecuentes es la comparación de dos distribuciones (bien de la misma variable estadística en dos conjuntos de datos o de dos variables estadísticas en el mismo conjunto de datos). Por ello, algunos autores han tratado de ver cómo los estudiantes razonan al realizar esta tarea.

Un ejemplo lo tenemos en los resultados obtenidos por Konold, Pollatsek, Well y Gagnon (1997) al tratar de analizar las barreras críticas en el aprendizaje de la estadística. A pesar de que la comparación de distribuciones, formalmente, se lleva a cabo a partir de las medidas de posición central y dispersión de las variables, Konold y sus colaboradores sugieren que algunos estudiantes ni siquiera usan intuitivamente la media para comparar dos conjuntos de datos. En lugar de ello se centran en las frecuencias absolutas y no en las relativas, al comparar dos conjuntos de datos, incluso cuando las muestras eran de tamaño muy diferentes. Una explicación es que algunos estudiantes no ven el dato como un valor de una variable, sino consideran sólo datos aislados. Como consecuencia no ven los resúmenes estadísticos (como media o rango) como propiedades de la distribución (Bakker y Gravemeijer, 2004).

Shaughnessy y Ciancetta (2002) estudian el razonamiento de los estudiantes al comparar distribuciones de datos. En algunos casos, sólo consideran las medidas de posición central, y son incapaces de pasar de ellas a la distribución. Al pedir a los estudiantes que escriban una distribución con un valor dado de la media o mediana, los estudiantes se limitan a repetir valores muy similares o equidistantes de la medida de posición central.

Otro requisito para comprender la distribución es la idea de variabilidad, que está siempre presente en los datos y tiene múltiples significados en estadística (Reading y Shaughnessy, 2004), entre otros los siguientes: variabilidad de resultados posibles en un experimento aleatorio; variabilidad en los datos recogidos; variabilidad en una variable aleatoria; variabilidad en las muestras o la distribución muestral. Es por ello importante que los estudiantes perciban la variabilidad y manejen modelos que permitan controlarla y predecirla.

Batanero, Estepa y Godino (1997) estudian la forma en que los estudiantes comparan dos distribuciones. Como estrategias correctas encuentran el comparar las medias o el reducir el conjunto de datos a una sola variable, restando los valores correspondientes, en el caso de muestras relacionadas y comparar luego la media con cero. También encuentran estrategias incorrectas como comparar valores aislados en las dos distribuciones.

En una serie de trabajos, Watson y colaboradores clasifican las estrategias de los estudiantes al comparar distribuciones (Watson y Moritz, 1999; Watson, 2001) de acuerdo al modelo jerárquico SOLO, definiendo niveles de comprensión. En el primer nivel de la jerarquía descrita por la autora, los estudiantes son capaces de comparar conjuntos de igual tamaño, mientras que en el segundo se comparan conjuntos de datos de diferente tamaño. Las estrategias de los estudiantes incluyen razonamiento proporcional y comparación de las gráficas y de las medidas de posición central de los dos grupos.

Shaughnessy (2007) indica que muchos de los errores descritos en las investigaciones sobre comprensión de gráficos enmascaran otros relacionados con la comprensión de la distribución. Estas investigaciones se han centrado preferentemente en dos puntos:

– Describir errores en la construcción de gráficos por parte de los estudiantes. Algunos de estos errores son: confusión de los ejes en el gráfico, representación incorrecta de los intervalos en un histograma de frecuencias, representar variables no relacionadas en una misma gráfica o usar un gráfico no adecuado para los datos representados.

– Definir niveles jerárquicos en la lectura e interpretación de gráficos. Por ejemplo Friel, Curcio y Bright (2001 diferencian entre “leer entre los datos” (lectura literal del gráfico sin interpretar la información contenida en el mismo), “leer dentro de los datos” (interpretación e integración de los datos en el gráfico), “leer más allá de los datos” (predicciones e inferencias a partir de los datos sobre informaciones que no se reflejan directamente en el gráfico) y “leer detrás de los datos” (valorar críticamente el método de recogida de datos su validez y fiabilidad).

En lo que sigue tratamos de complementar las anteriores investigaciones centrándonos en futuros profesores, un colectivo cuya comprensión de la distribución no ha sido analizada en profundidad. Evaluamos esta comprensión a partir de un proyecto abierto, tarea que no ha sido usada en las investigaciones anteriores. Más en concreto, la comprensión de la distribución se evalúa a partir de los gráficos que elaboran sobre la misma, los estadísticos (medidas de posición central y dispersión) que calculan y la interpretación de estos elementos en relación a la pregunta planteada.

3. El Estudio

La actividad de la cual se tomaron los datos era una de las prácticas de un curso de Didáctica de la Matemática para futuros profesores de Educación Primaria de la Universidad de Granada y se dedicaron a ella dos sesiones de clase, de dos horas de duración cada una. En la primera sesión se propuso a los futuros profesores la realización de un proyecto en el que ellos mismos tomaron los datos necesarios.

El proyecto se llama “Comprueba tus intuiciones sobre el azar” y forma parte de una unidad didáctica diseñada para introducir los temas sobre “tratamiento de la información, azar y probabilidad” en la formación de maestros (Godino, Batanero, Roa y Wilhelmi, 2008). Fines complementarios de la actividad fueron: a) mostrar la utilidad de la estadística para probar conjeturas; b) comprobar que las intuiciones sobre el azar a veces son engañosas. La secuencia de actividades es como sigue:

1. *Presentación del proyecto, instrucciones iniciales y discusión colectiva.* Se comenzó la sesión con una discusión sobre las intuiciones y se propuso a los futuros profesores llevar a cabo un experimento para decidir si la clase en conjunto tenía o no buenas intuiciones sobre el azar. El experimento consiste en inventar una secuencia de 20 lanzamientos de una moneda equilibrada (sin lanzarla realmente) y comparar con 20 lanzamientos reales de la moneda.

2. *Experimentos individuales y recogida de datos.* Los futuros profesores realizaron individualmente el experimento inventando una secuencia de 20 lanzamientos (secuencia simulada) y anotaron los resultados en una hoja de registro, escribiendo C para cara y + para cruz. A continuación lanzaron realmente la moneda anotando los resultados en la segunda parte de la hoja (secuencia real).

3. *Discusión y actividades de la clase.* Finalizado el experimento, se comenzó la discusión sobre cómo comparar los resultados de la clase en las secuencias reales y simuladas. Entre otras variables, algunos estudiantes sugirieron comparar el

número de caras, número de rachas y longitud de la racha más larga en las dos secuencias.

4. Al acabar la clase el profesor proporcionó a los estudiantes una hoja de datos que contenía para cada alumno los valores de estas variables (101 estudiantes y 3 pares de variables; número de caras, número de rachas y longitud de la racha más larga en las secuencias real y simulada). Pidió que analizasen los datos individualmente, y produjeran un informe escrito en que justificasen si la clase en su conjunto tenía o no buenas intuiciones sobre el azar, en base al análisis de los datos. Los estudiantes tuvieron libertad para elegir los gráficos o resúmenes estadísticos que considerasen convenientes o bien usar ordenadores.

A la semana siguiente se recogieron y analizaron las producciones de los estudiantes, centrándonos en los gráficos construidos, los resúmenes estadísticos calculados, la interpretación y conclusiones sobre la pregunta planteada.

4. Representación Gráfica de la Distribución

En primer lugar se analizaron los gráficos realizados por los estudiantes para representar la distribución de las diferentes variables estadísticas que intervienen en el proyecto.

Cuando los alumnos construyen un gráfico para todas o parte de las distribuciones contenidas en el fichero de datos, realizan una serie de acciones y emplean varios conceptos y propiedades que varían en los diferentes gráficos. En consecuencia, los gráficos producidos no son representaciones equivalentes de un concepto subyacente (la distribución de datos obtenida) sino configuraciones diferenciadas de objetos relacionados e interactuando con dicha distribución (Godino, Batanero y Font, 2007). En lo que sigue describimos los niveles de complejidad de los gráficos. De la muestra de los 101 estudiantes, 88 produjeron algún tipo de gráfico. Sólo tenemos en cuenta estos estudiantes en las tablas de frecuencia que se presentan en esta sección.

C1. Alumnos que producen una gráfica donde sólo representa los resultados de su experimento individual. Algunos alumnos producen una gráfica para representar los datos obtenidos en su experimento particular, sin considerar los datos de sus compañeros, al no haber comprendido el sentido del experimento. Estos estudiantes tratan de resolver la pregunta para su caso particular (si él mismo tiene una buena intuición), no habiendo comprendido el propósito del proyecto o no siendo capaces de realizar un análisis global de los datos. Generalmente manifiestan una intuición errónea sobre la aleatoriedad, suponiendo que una buena intuición implicaría que su secuencia simulada fuese idéntica en alguna característica a su secuencia real. En general tratan de ver si sus resultados particulares en las dos secuencias han estado igualados. Aunque la impredecibilidad es una diferencia esencial entre experimentos aleatorios y deterministas, estos estudiantes tratan de predecir el resultado o al menos predecir una parte de los resultados de la secuencia aleatoria. Este fenómeno, denominado *ilusión de control* fue observado originalmente de Langer (1975) en diferentes tipos de juegos de azar, por ejemplo la lotería, donde observó en los sujetos una creencia ilusoria sobre su capacidad para influir en el resultado.

C2. Alumnos que representan los datos de toda la clase, sin llegar a la distribución de frecuencias. Son los estudiantes que no llegan a agrupar los valores similares del número de caras, rachas o la longitud de la racha mayor obtenidos en las secuencias reales o simuladas del total de alumnos en la clase, para formar la distribución de frecuencias. En lugar de ello, representan el valor (o valores) obtenidos para cada alumno dentro del gráfico en el orden en que los datos fueron obtenidos, que es un orden arbitrario. Aunque emplean las variables estadísticas, estos estudiantes no calculan las frecuencias asociadas a cada valor. Por tanto no emplean la idea de distribución de frecuencias de la variable.

C3. Producen gráficos separados para las distribuciones de las distintas variables. Son alumnos que forman tablas de frecuencias para las variables y a partir de ellas producen los distintos gráficos o representan directamente un gráfico de cada uno de los valores diferentes de la variable con sus frecuencias (para las 6 variables). En este nivel el alumno pasa del conjunto de datos a la variable estadística y su distribución de frecuencias. El usar para cada par de variables dos gráficos dificulta a veces la comparación, sobre todo en caso de no usar la misma escala de representación en los dos gráficos.

C4. Produce para cada par de variables un gráfico conjunto de las dos distribuciones. El alumno ha llegado a formar las distribuciones de las variables y las distribuciones de cada par de variables (por ejemplo para el número de caras en las secuencias real y simulada) las representa conjuntamente en el mismo gráfico, lo cual facilitará la comparación de las distribuciones. Cada gráfico de este nivel tiene también mayor complejidad al representar conjuntamente las distribuciones de dos variables estadísticas.

En la Tabla 1 presentamos la distribución de alumnos en función del nivel de gráfico. Del total de 101 alumnos, 88 (87,1%) producen algún tipo de gráfico para analizar sus datos, incluso cuando las instrucciones de la tarea no los requerían. Este alto porcentaje indica la necesidad sentida de los estudiantes de producir un gráfico y llegar, mediante un proceso de transnumeración a un conocimiento no disponible en los datos brutos.

Tabla 1 – Clasificación de estudiantes, según nivel de complejidad semiótica y corrección de los gráficos

Nivel de complejidad	Frecuencia (%)
C1. Representa sólo sus datos	2
C2. Representa resultados individuales	15
C3. Gráficos separados para cada muestra	46
C4. Gráficos conjuntos	25
Total	88

La mayoría de los que elaboran gráficos (52,2%), los producen separados para cada variable (nivel 3), el 28,4% de los estudiantes trabajan al nivel 4 y producen un solo gráfico para cada par de variable de las 6 variables en estudio. Son pocos los estudiantes que analizan sólo sus propios datos (nivel 1) y sólo 17% estudian los valores obtenidos por cada estudiante caso a caso sin llegar a formar la distribución. En consecuencia, el concepto de distribución, al menos a nivel intuitivo, parece ser adquirido y utilizado por los estudiantes para resolver la tarea propuesta.

Errores en los gráficos

Los alumnos tienen dificultades en la construcción de los gráficos. En el ejemplo que reproducimos en la Figura 1, ejemplo perteneciente al nivel de complejidad C3, podemos observar, por un lado, que las barras del diagrama no están centradas en los valores del eje X (longitud de la racha mayor). Por otro lado, las escala del eje Y no es homogénea; finalmente faltan en el gráfico el título y no se indica la variable representada en el eje X.

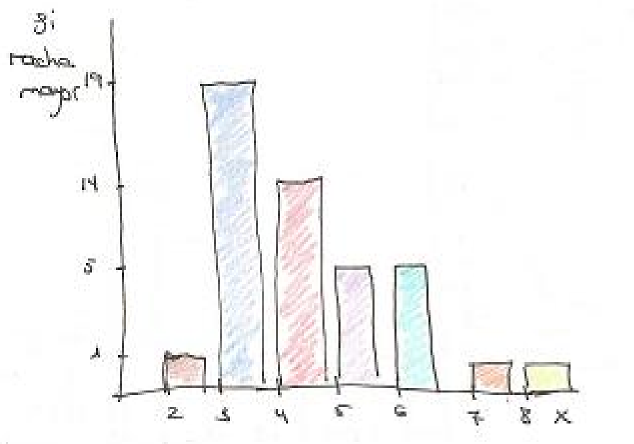

Figura 1 – Ejemplo de gráfico con errores en escalas

Algunos estudiantes representan variables no relacionadas sobre una misma gráfica; lo que indica que no comprenden el propósito de una gráfica conjunta ni discriminan las situaciones en que dos variables estadísticas son o no comparables. Un ejemplo lo tenemos en la Figura 2 en que el estudiante representa conjuntamente el número de caras, número de rachas y longitud de la racha más larga en tres polígonos de frecuencias representados sobre la misma gráfica. Por otro lado, los polígonos están incompletos al no unir los extremos con el eje de ordenadas. Unos pocos estudiantes intercambian los ejes en las gráficas, representando las frecuencias en el eje X y los valores de la variable en el eje Y, confundiendo la variable dependiente e independiente en la distribución de frecuencias de la variable estadística mostrando un error similar al descrito por Ruiz (2006) en un estudio sobre comprensión de la variable aleatoria.

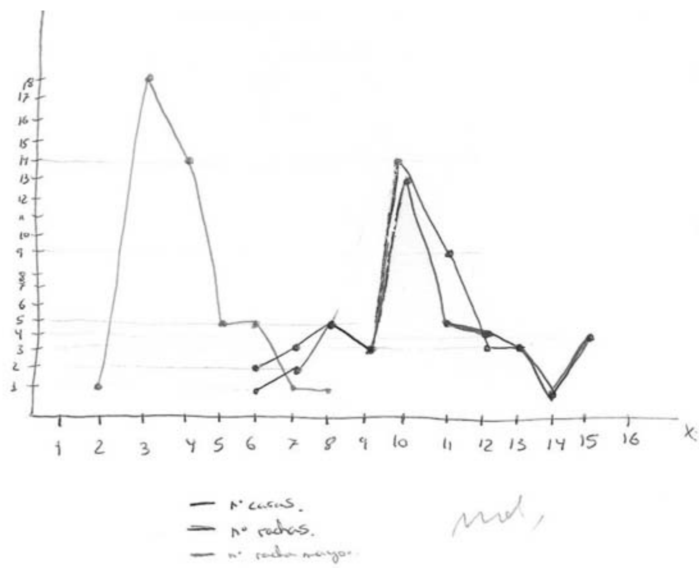

Figura 2 – Ejemplo de gráficos con variables no comparables

En la tabla 2 presentamos a los estudiantes clasificados según el nivel más alto al que llegan al construir el gráfico y según la corrección de los gráficos presentados. Vemos que la cuarta parte de los que produce gráficos, hacen errores de importancia (gráficos incorrectos) como representar variables no comparables, representar los productos de valores por frecuencia, o intercambiar frecuencias y valores de las variables en los ejes.

Otra cuarta parte tiene errores de menor importancia (gráfico parcialmente correcto): errores en las escalas, etiquetas o ejes, no usan la misma escala en los dos gráficos o bien no usan el mismo gráfico en las dos muestras dificultándose la comparación, no centran el intervalo en los histogramas, no hay coincidencia de los valores representados con la escala utilizada o no incluyen etiquetas. Por ejemplo, en la figura 3 mostramos un gráfico conjunto de las distribuciones del número de caras en las secuencias real y simulada, que es casi correcto, pero usa diferente gráfico (histograma y polígono de frecuencia) para el número de caras en las dos secuencias. Por otro lado, el polígono de frecuencias está incompleto al no unir los extremos al eje de ordenadas, error señalado en la investigación de Espinel (2007) con futuros profesores. Consideramos estos gráficos parcialmente correctos.

Tabla 2 – Clasificación de estudiantes, según nivel de complejidad y corrección de los gráficos

Nivel de complejidad semiótica	Corrección del Gráfico			Total en el nivel
	Correcto	Parcialmente Correcto	Incorrecto	
C1. Representa sólo sus datos	1		1	2
C2. Representa resultados individuales	10	1	4	15
C3. Gráficos separados para cada muestra	15	17	14	46
C4. Gráficos conjuntos	14	6	5	25
Total	40	24	24	88

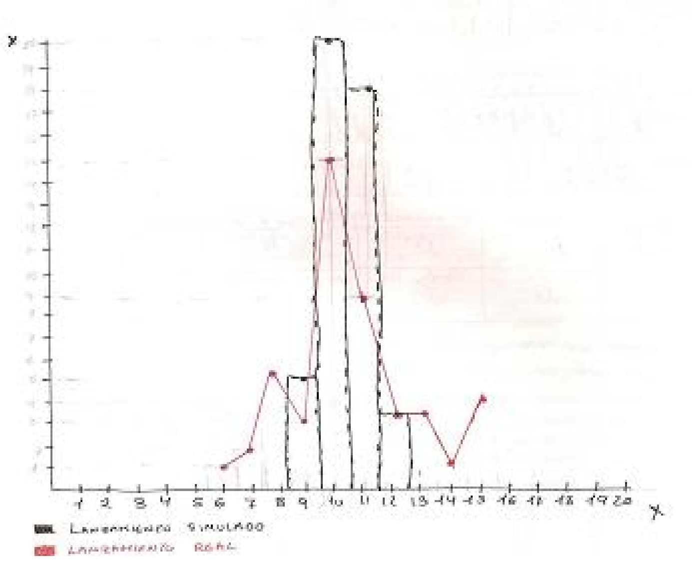

Figura 3 – Gráfico parcialmente correcto

Interpretación de los gráficos

En la Tabla 3 clasificamos los estudiantes, según nivel del gráfico e interpretación que hacen del mismo, donde una cuarta parte se limita a producir el gráfico sin llegar a interpretarlo o bien da una interpretación errónea. Aunque leen elementos aislados del gráfico (dominan la extracción de datos), no llegan al nivel de extracción de tendencias o análisis de la estructura (Bertín, 1967) en su lectura de gráficos.

Tabla 3. Clasificación de estudiantes, según nivel de complejidad semiótica e interpretación de los gráficos

Categorías de gráfico construido	Interpretación del gráfico			Total en el nivel
	Correcto	Parcialmente Correcto	Incorrecto o no interpreta	
C1. Representa sólo sus datos	1	1		2
C2. Representa resultados individuales	4	10	1	15
C3. Gráficos separados para cada muestra	15	15	16	46
C4. Gráficos conjuntos	9	11	5	25
Total	29	37	22	88

Estos errores o falta de interpretación se producen, sobre todo, en los gráficos de nivel 3 y 4 pues la menor complejidad semiótica de los gráficos de nivel 1 y 2 hace que estos sean más sencillos de interpretar para los estudiantes.

En el siguiente ejemplo de interpretación incorrecta (en un grafico de nivel 3), el estudiante muestra un lenguaje poco preciso al hablar de “frecuencia absoluta que más se repite” para referirse a la máxima frecuencia absoluta. Vemos también que el estudiante está comparando tres variables al mismo tiempo, aunque había representado cada una de las variables por separado en un diagrama de barras. Es decir, aparece un conflicto en la idea de moda, pues ella está buscando la moda (valor de más frecuencia) en el conjunto de las tres variables y no en cada variable por separado:

> *En el caso de las gráficas he optado por realizar un diagrama de barras en el cual se puede observar que una secuencia simulada la frecuencia absoluta que más se repite se encuentra en la tabla de racha mayor, siendo el valor 3 y la frecuencia 26.* (Alumno SL)

En otros casos la interpretación errónea se debe a errores conceptuales, como en el siguiente caso donde aparece un error respecto a la dependencia funcional lineal:

> *Todos los gráficos presentan una dependencia funcional lineal de – a + o viceversa, según las circunstancias.* (Alumno PF)

Un 42% de los estudiantes hacen una interpretación parcialmente correcta de los gráficos, analizando tan sólo la tendencia sin tener en cuenta la variabilidad o bien al contrario, comparando sólo los rangos de variación, sin tener en cuenta las tendencias. La mayor proporción relativa de los que hacen esta interpretación parcial se encuentra entre los estudiantes que producen un gráfico de nivel 2 que, como hemos indicado, no llegan a trabajar con la distribución de la variable estadística.

En el siguiente ejemplo la interpretación es parcialmente correcta. La alumna había hecho un gráfico de nivel C2, representando los datos alumno a alumno sin formar la distribución. Como usa un gráfico de líneas, este permite ver la mayor variabilidad en una de las dos series, porque en este caso las oscilaciones del gráfico son más amplias. Lo consideramos parcialmente correcto porque no compara la posición central.

> *En cuanto al número de caras reales que ha obtenido el grupo es mucho más variable son más homogéneos que en la simulación. Para poder realizar esta afirmación hemos usado el rango.* (Alumno JI)

El alumno AG por el contrario usa un vocabulario más impreciso; sin embargo se ven puntos de interpretación correcta del gráfico (en el diagrama de líneas al ascender la línea aumenta la frecuencia; detección de la moda).

> *El sistema de representación de datos utilizado es el diagrama de líneas ya que se pueden ver las diferencias entre la secuencia simulada y la real; si la línea asciende, la frecuencia aumenta. El punto más alto del diagrama corresponde al valor de la variable que tiene mayor frecuencia, como ocurre [pone un ejemplo].* (Alumno AG)

La mayor proporción absoluta de los que hacen esta interpretación parcial se encuentra entre los estudiantes que producen un gráfico de nivel C3 igualando en este nivel al número de interpretaciones correctas. Hay también una alta proporción de interpretaciones parcialmente correctas dentro de los que llegan a producir gráficos de nivel de complejidad C4; en algunos casos los gráficos son producidos con la ayuda de Excel, por lo que el alumno aunque llega al mayor nivel a la hora de representar los datos, no comprende bien el gráfico producido y no llega a interpretarlo en forma completa.

Una tercera parte de los estudiantes hace una interpretación completa del gráfico, detectando las tendencias y variabilidad de los datos, aumentando la proporción de los que lo hacen en los dos niveles superiores del gráfico construido:

> *Para comentar las diferencias el gráfico aporta más posibilidades, pues visualmente podemos percibir las diferencias. Las tablas necesitan de una interpretación de los datos. Centrándome en el gráfico observamos que los valores de la frecuencia simulada se concentran en torno a unos mismos intervalos 10 y 11, mientras que en la frecuencia real, los valores están más disgregados, aunque también vemos que se concentran en un núcleo 8-15, fluctuando al alza y a la baja. De esta forma en la frecuencia real comprobamos que 10 es la secuencia con mayor número de veces saliendo cara, tanto en la simulada (20) como en la real (14).... la horquilla de la secuencia real es mucho más amplia que la de la simulada.* (Alumno FL)

Encontramos también interpretaciones correctas de los gráficos en estudiantes que sólo han trabajado a nivel 1 o 2. Por ejemplo, un alumno totaliza el número de caras y número de cruces en la muestra completa de alumnos y lo representa en un sector circular, realizando el siguiente comentario:

> *Si observamos en el sector circular, nos daremos cuenta que los alumnos han puesto al azar más caras que cruces y que en la secuencia real dan unos resultados casi iguales en el porcentaje de caras y cruces.* (Alumno CG)

La interpretación de este gráfico por parte del alumno es correcta, pues en el gráfico de sectores construido hay mayor frecuencia de caras y los porcentajes están muy igualados al tratarse de una muestra bastante grande. Sin embargo, aunque la interpretación del gráfico en si es correcta, el alumno no ha entendido la finalidad del experimento, ni está considerando la variable aleatoria binomial "número de caras en 20 lanzamientos" o la correspondiente variable estadística. No diferencia los ensayos de cada estudiante y simplemente considera una única secuencia formada al yuxtaponer las secuencias individuales de todos los estudiantes de la clase.

Sin tener en cuenta el nivel C1, en el que sólo se encuentran dos alumnos, y uno de ellos interpretó bien su gráfico, con lo que el porcentaje de interpretaciones correctas fue del 50%, el nivel en el cual se dio mayor número de interpretaciones correctas fue el nivel C4. Una vez que un alumno es capaz de construir correctamente un gráfico de este nivel de complejidad, este tipo de gráficos facilita la comparación de tendencias y variabilidad en cada pareja de distribuciones. Dicho de otro modo, el mayor nivel alcanzado en la construcción de gráficos parece implicar mayor nivel en la lectura e interpretación de los mismos.

5. Momentos de la Distribución

Heitele (1975) señala tres conceptos fundamentales en relación a la variable aleatoria: su distribución, esperanza y variabilidad. Por tanto, estos tres componentes serán también fundamentales en la variable estadística, cuya generalización da lugar a la idea de variable aleatoria. Una vez analizados los gráficos producidos por los estudiantes, pasamos a analizar el uso que hacen de las medidas de posición central y dispersión.

Medidas de posición central

De un total de 101 estudiantes, 91 calculan algún promedio, lo que indica una primera aproximación a la idea de distribución en los estudiantes, pues pasan del dato aislado a un resumen estadístico del conjunto de datos. En la tabla 4 vemos que el estadístico calculado con preferencia es la media (90) o moda (80); la mediana es calculada por 72 estudiantes. En general el cálculo de los medidas de posición central es correcto, lo que indica un buen dominio de los algoritmos de cálculo, aunque aparecen algunos errores: no ordenar los datos en la mediana (17), falta de ponderación o error de cálculo en media (11) y no tener en cuenta el caso de bimodalidad para la moda (3). Estos errores coinciden con los ya señalados en estudiantes de secundaria en investigaciones previas como la de Cobo (2003).

Tabla 4 – Estadísticos calculados por los estudiantes (n=101)

	Correcto?	Incorrecto?	Total
Media	79	11	90
Mediana	55	17	72
Moda	77	3	80
Rango	54	6	60
Otra medida dispersión	41	–	41

Por otro lado, hemos encontrado algunas concepciones erróneas sobre las medidas de posición central, explicitadas en la solución del proyecto por los estudiantes. Por ejemplo, el siguiente estudiante, confunde la variable en observación (número de caras en las diferentes secuencias producidas por cada estudiante, en la que el máximo fue 12) con los sucesos del experimento (cara y cruz), al calcular la moda da la siguiente respuesta:

> *Moda = Son las caras porque su número máximo es 12, mientras que el número máximo de cruces es de 11.* (Alumno CE)

En otros casos no se tiene en cuenta la frecuencia en el cálculo de la mediana, considerando tan sólo los diferentes valores obtenidos de la variable y tomando su punto medio, lo que sería equivalente a considerar como mediana el centro del rango, error encontrado en Mayén (2006) en su estudio sobre comprensión de los medidas de posición central:

> *La mediana: ordenar de mayor a menor 2-15. (Alumno CF)*

Dispersión

El uso de medidas de dispersión es menor en los estudiantes, resultado que sugiere que no sienten la necesidad de estas medidas para comparar distribuciones. La mayoría de los que lo usan calculan el rango (60), 54 de ellos correctamente. Algunos estudiantes toman incorrectamente los mismos valores del máximo mínimo y rango en la secuencia real y simulada. Unos pocos calculan además la desviación típica (25), varianza (15) y coeficiente de variación (1). Es evidente que las medidas de dispersión son menos intuitivas para los futuros profesores, por el menor uso que se hace de las mismas, como también se confirma en la investigación de Borim y Coutinho (2008). Estas autoras indican que el significado de la varianza y la desviación típica es muy difícil de comprender por los estudiantes para profesor. En nuestra muestra fueron muy pocos los que usaron estas medidas y menos aún los que las interpretaron correctamente.

Por otro lado, se ha analizado el uso de los estadísticos por parte de los estudiantes, para comparar las distribuciones una vez los han calculado (Tabla 5), encontrado que, aunque se calculen, no siempre se emplean los estadísticos, limitándose en muchos casos a presentarlos sin ningún comentario respecto a su significado o las diferencias entre las dos secuencias.

Tabla 5 – Uso de los estadísticos en la comparación (n=101)

Tipo de comparación	Frecuencia
Comparan media	36
Comparan moda	18
Comparan mediana	12
Comparan dispersión	30
Comparan valores aislados	10
Comparan solo sus datos	11
Indican lo que esperan	12
No comparan	23

Respecto a las medidas de posición central 36 comparan las medias (32 correctamente), 18 las modas (12 correctamente) y doce las medianas (dos incorrectamente). Algunos estudiantes presentan los estadísticos pero no comparan (23), indican lo que esperan sin comparar (12), siguiendo sus creencias previas, por ejemplo, el siguiente caso, en que el alumno, aunque calculó media, mediana y moda, no las usa en la comparación:

> *En un principio es factible que el nº de caras y de cruces a priori debería ser el mismo, pero debido a que esto es un proceso aleatorio, las secuencias de caras y cruces en realidad es imprevisible.* (Alumno ES)

Otros alumnos comparan sólo sus propios datos (11), al no haber comprendido el propósito de la práctica, tratando de evaluar únicamente sus propias intuiciones. Por ello se limitan a comparar el valor de las diferentes variables en las dos secuencias, sin hacer referencia a los estadísticos; generalmente manifiestan que una buena intuición supone poder acertar el experimento, es decir, tienen una cierta ilusión de control:

> *Secuencia simulada: Longitud de la racha más larga=5, número de rachas: 0; número de caras=10. Secuencia real: Longitud de la racha más larga=3; número de rachas: 15; número de caras 11; he acertado 13 veces.* (Alumno EL)

Diez alumnos comparan valores aislados de la variable o porcentajes de valores aislados (10 estudiantes); cuatro de ellos comparan únicamente los máximos y mínimos en cada distribución. Esta estrategia de comparación de valores aislados al estudiar la diferencia de dos distribuciones ya apareció en la investigación de Estepa y Batanero (1995), quienes la explican por la existencia de una concepción local de la asociación estadística, consistente en juzgar la asociación entre dos variables considerando tan sólo una parte de los datos y no el conjunto completo.

> *El máximo que he encontrado en las dos secuencias es 15 en el número total de caras, 16 el número más alto en rachas y 8 el máximo en la racha real.* (Alumno CG)

Sólo 30 utilizan las medidas de dispersión (27 de ellos el rango y seis la desviación típica o la varianza). Un ejemplo se reproduce a continuación, en que el alumno es capaz de apreciar la diferencia de dispersión en el número de caras y su semejanza en el número de rachas:

> *En primer lugar compararé los datos del número de caras en ambas secuencias, donde destacaré las diferencias que se aprecian principalmente entre el rango, la varianza y, consecuentemente la desviación típica, de hasta seis veces más de diferencia la secuencia simulada con respecto a la real" (quiere decir lo contrario, pues la dispersión es mayor en la secuencia real). En segundo lugar destacaré los datos del número de rachas de las secuencias simuladas y la secuencia real. He de decir que son sorprendentes las similitudes del rango e incluso la varianza y la desviación típica que tan sólo llegan a diferenciarse en décimas.* (Alumno LC)

En el caso siguiente se confunde frecuencia y valor de la variable, error encontrado por Cobo (2003) en el cálculo de la media. Como vemos en el ejemplo, el alumno se refiere a la frecuencia, pues al ser sólo 20 los lanzamientos, la longitud de la racha no puede ser igual a 25.

> *En la secuencia simulada la racha mayor y en la que más ha coincidido la clase es 25.* (Alumno TF)

6. Conclusiones sobre el Problema Planteado

Para resolver el problema planteado en el proyecto, además de trabajar con las distribuciones, los estudiantes han de interpretar los resultados del trabajo matemático realizado con ellas en el contexto del problema (traducir estos resultados a lo que indican respecto de las intuiciones de los estudiantes). Es precisamente este último paso (puesta en relación del resultado con la pregunta planteada) el que ha causado más dificultad, por la falta de familiaridad de los futuros profesores con proyectos estadísticos y actividades de modelización.

En la Tabla 6 presentamos las conclusiones obtenidas en relación con las intuiciones de la clase respecto a los fenómenos aleatorios en relación al número de caras en los 20 lanzamientos. Observamos que la obtención de la conclusión es la tarea más difícil para todos los estudiantes, siendo sólo una tercera parte los que obtienen una conclusión, al menos parcial.

Tabla 6 – Clasificación de estudiantes, según conclusión obtenida sobre el número de caras

Conclusión	Frecuencia
Correcta	4
Parcialmente correcta	24
Incorrecta o no concluye	73

Solo cuatro estudiantes completan las conclusiones de que por un lado, el grupo tiene buena intuición respecto al promedio de número de caras y por otro las intuiciones sobre la variabilidad de los fenómenos aleatorios es pobre en los estudiantes. A continuación reproducimos la respuesta de uno de los estudiantes que obtiene una conclusión completa.

> *En cuanto al número de caras las intuiciones del aula fueron aproximadas a la realidad, pero no del todo, ya que la desviación típica nos indica que los datos se distancian". Anteriormente la alumna dice: "La media entre la simulada y la real se asemeja,.... la mediana y la moda dan los mismos datos.* (Alumna CG)

Veinticuatro estudiantes llegan a una conclusión parcial, debido bien a que sólo comparan las medidas de posición central sin tener en cuenta la dispersión, o al contrario.

> *La intuición de mis compañeros observando la tabla del nº de caras es buena, ya que los valores más repetidos en la secuencia simulada coinciden con los valores de la secuencia real: 10 y 11 son las más repetidas. La media de las dos secuencias es 10, por lo tanto creo que la intuición es buena.* (Alumno TG)

Hemos considerado conclusión parcial cuando aparece imprecisión de lenguaje, como en el siguiente ejemplo, donde el alumno visualiza tanto la diferencia de dispersión como los valores centrales similares, pero no llega a concluir claramente sobre la intuición de los estudiantes.

> *Respecto a las gráficas construidas sobre el número de caras he de comentar que se parece observar algunos cambios en la secuencia simulada donde encontramos cuatro posibilidades y en la real hay diez posibilidades. Con esto podríamos decir que los alumnos no han sido intuitivos, pero no creo que esto sea así, ya que si nos detenemos atentamente en las gráficas podemos ver que los valores de 9, 10, 11 y 12 en ambas han sido dados por un mayor número de alumnos que en los demás casos.* (Alumno HC)

El siguiente alumno reconoce la diferencia de dispersión en las distribuciones, pero interpreta el problema como de búsqueda de diferencia de intuiciones entre los estudiantes. Es decir, concluye que las intuiciones son similares en distintos estudiantes, pero no llega a relacionar estas intuiciones con las características del fenómeno aleatorio. Tampoco hace observaciones sobre las medidas de posición central.

> *Realmente las intuiciones de los alumnos han sido más o menos muy parecidas. No hay muchas irregularidades. Pero cuando las comparamos con sus secuencias reales, se puede ver a simple vista en el gráfico que existen grandes irregularidades. En la secuencia real el número máximo de caras es 16 y el mínimo 7, sin embargo en la simulada el máximo es 13 y el mínimo 8; su recorrido es más pequeño que el real.* (Alumno MM)

El resto no llega a una conclusión o bien hace una conclusión incorrecta. Los estudiantes no siempre conectan los resultados del trabajo matemático con la situación problemática, es decir, no ven las implicaciones de lo obtenido en el análisis estadístico sobre las intuiciones de los estudiantes.

> *Comparando los datos me he dado cuenta que son muchos los resultados entre los compañeros los que coinciden, pero aún así, sigo pensando que es mera casualidad, porque en la simulada hemos puesto lo que hemos querido.* (Alumno EL)

Otros estudiantes, aún cuando conectan el trabajo con el modelo matemático (distribución) con el problema real, fallan en la obtención de conclusiones debido a que no han comprendido la pregunta planteada y suponen que una buena intuición ha de corresponder a obtener los mismos resultados en las secuencias real y aleatoria. En la siguiente respuesta, el estudiante muestra una concepción correcta del azar (no se puede prever) y otra incorrecta al tratar de evaluar el número de coincidencias o la diferencia de valores obtenidos en cada estudiante en los experimentos, en lugar de comparar directamente las distribuciones de las variables. Es decir, este estudiante compara caso a caso no utilizando la distribución al hacer la comparación.

> *Partiendo del estudio de los gráficos, se podría decir que en valores absolutos, la previsión del grupo no ha sido demasiado desafortunada. Un juego de azar es imposible de prever con total exactitud, pero las aproximaciones sumadas a los aciertos son mayores a las previsiones muy alejadas del resultado real.* (Alumno LG)

Falta en otros la capacidad de análisis para detectar las diferencias, lo que les lleva a concluir que los resultados en las dos secuencias son similares; en el siguiente ejemplo, además de darse este caso, no se llega a relacionar el resultado con la intuición de los estudiantes.

> *Como conclusión, podemos ver que los resultados son prácticamente los mismos tanto en la real como en la simulada, de lo que podemos deducir que lo real y lo simulado es muy parecido ya que lo que te inventes puede ser prácticamente igual a lo real.* (Alumno AC)

7. Conclusiones

Al comparar las distribuciones la mayoría de los estudiantes elabora un gráfico y un porcentaje importante de estudiantes de la muestra ha calculado algún estadístico, principalmente estadísticos de posición central (media, mediana y moda) y en menor medida de dispersión.

En los gráficos construidos, la mayoría llega a formar la distribución, auque aproximadamente la mitad de los gráficos tienen algún error. Respecto a la interpretación de gráficos, nuestros resultados indican que es una habilidad altamente compleja, y confirma las dificultades descritas por Bruno y Espinel (2005) y Espinel (2007) en futuros profesores, a pesar de que han de transmitir el lenguaje gráfico a sus alumnos y utilizarlo como herramienta en su vida profesional. También amplía el trabajo de las autoras proporcionando datos sobre la capacidad de construcción de gráficos de los futuros profesores en una tarea abierta y mostrando que la mayoría de los participantes no consigue elaborar un gráfico de complejidad suficiente para permitir resolver el problema.

El cálculo de estadísticos es en general correcto, con pocas excepciones, aunque también aparecen algunos de los errores descritos en la investigación de Cobo (2003) y Mayén (2006), pero con mucha menor frecuencia que en aquellos estudios. Aunque el cálculo de las medidas de posición central fue sencillo, la mayoría de los estudiantes se limita a calcularlos, pero no los interpreta ni los usa para la comparación, aunque en caso de compararlos, la comparación es correcta. Menos aún usan la idea de dispersión, aunque el cálculo del rango lo hacen correctamente, pareciendo que aunque se comprende el procedimiento de cálculo, los alumnos no llegan a captar el significado de la dispersión ni su utilidad en la comparación de dos distribuciones.

En resumen, el concepto de distribución, esencia del razonamiento estadístico según Wild y Pfannkuch (1999) no llega a ser utilizado por una parte de los futuros profesores y el razonamiento sobre la variabilidad, que es otro de los componentes esenciales del razonamiento estadístico, es difícil para la mayoría. En consecuencia, sería necesario atender a estos problemas en la formación de los profesores de educación primaria, pues una mejora de la educación de los niños pasa por la formación del profesor.

Agradecimientos: Este trabajo forma parte del proyecto SEJ2007-60110 (MEC - Feder), beca FPU AP2007-03222.

Referencias

Bakker, A. y Gravemeijer, K. (2004). Learning to reason about distribution. En J. Garfield y D. Ben Zvi (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning and Thinking* (pp 147-168). Dordrecht: Kluwer.

Batanero, C., Estepa, A. y Godino, J. D. (1997). Evolution of Students' Understanding of Statistical Association in a Computer-Based Teaching Environment. En J. B. Garfield y G. Burrill (Eds.), *Research on the Role of Technology in Teaching and Learning Statistics. 1996 IASE Round Table Conference* (pp. 183-198). University of Minnesota: IASE.

Bertin (1967). *Semiologie graphique*. Paris: Gauthier-Villars.

Borim, C. y Coutinho, C. (2008). Reasoning about variation of a univariate distribution: a study with secondary mathematics teachers. En C. Batanero, G. Burrill, C. Reading y A. Rossman (Eds.), *Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education. Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference.* Monterrey: ICMI e IASE. CD-ROM.

Bruno, A. y Espinel, M. C. (2005). Recta numérica, escalas y gráficas estadísticas: un estudio con estudiantes para profesores. *Formación del Profesorado e Investigación en Educación Matemáticas, VII*, 57-85.

Cobo, B. (2003). *Significado de las medidas de posición central para los estudiantes de secundaria*. Tesis Doctoral. Universidad de Granada.

Espinel, C. (2007). Construcción y razonamiento de gráficos estadísticos en la formación de profesores. *Investigación en Educación Matemática, 11*, 99-119.

Estepa, A. y Batanero, C. (1995). Concepciones iniciales sobre la asociación estadística. *Enseñanza de las Ciencias,* 13(2), 155-170.

Friel, S., Curcio, F. y Bright, G. (2001). Making sense of graphs: critical factors influencing comprehension and instructional implications. *Journal for Research in mathematics Education 32*(2), 124-158.

Godino, J. D. Batanero, C. y Font, V. (2007). The onto-semiotic approach to research in mathematics education. *ZDM. The International Journal on Mathematics Education, 39* (1-2), 127-135.

Godino, J. D., Batanero, C., Roa, R. y Wilhelmi, M. R. (2008). Assessing and developing pedagogical content and statistical knowledge of primary school teachers through project work. En C. Batanero, G. Burrill, C. Reading y A. Rossman (2008). *Proceedings of the Joint ICMI /IASE StudyTeaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education. Proceedings of the ICMI Study 18 and 2008 IASE Round Table Conference*. Monterrey, México: ICMI e IASE. CD ROM.

Heitele, D. (1975). Un enfoque epistemológico sobre las ideas estocásticas fundamentales. *Educational Studies in Mathematics*, *6*, 187-205.

Konold, C., Pollatsek, A., Well, A. y Gagnon, A. (1997). Students analyzing data: Research of critical barriers. In J. B. Garfield y G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics.* Voorburg, The Netherlands: International Statistical Institute.

Langer, E.J. (1975). The illusion of control. *Journal of Personality and Social Psychology*, *32*, 311-328.

Mayen, S. (2006). *Comprensión de medidas de posición central en estudiantes mexicanos de Bachillerato*. Trabajo de Investigación Tutelada. Universidad de Granada.

MEC (2006). *Real Decreto 1513/2006, de 7 de diciembre, por el que se establecen las enseñanzas mínimas correspondientes a la Educación primaria*.

Reading, C. y Shaughnessy, J. M. (2004). Reasoning about variation. In J. Garfield y D. Ben-Zvi (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 201-226). Dordrecht, The Netherlands: Kluwer.

Ruiz, B. (2006). *Un acercamiento cognitivo y epistemológico a la didáctica del concepto de variable aleatoria*. Tesis de Maestria. CICATA. México.

Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. En F. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 957-1009). Greenwich, CT: Information Age Publishing, Inc., and NCTM.

Shaughnessy, J. M., y Ciancetta, M. (2002). Students' understanding of variability in a probability environment. In B. Phillips (Ed.), *CD of the Proceedings of the Sixth International Conference on Teaching Statistics: Developing a statistically literate society, Cape Town, South Africa.* Voorburg, The Netherlands: International Statistics Institute.

Watson, J. M. (2001). Longitudinal development of inferential reasoning by school students. *Educational Studies in Mathematics, 47,* 337-372.

Watson, J. M., & Moritz, J. B. (1999). The beginning of statistical inference: Comparing two data sets. *Educational Studies in Mathematics, 37*,145-168.