NÚMEROS
Revista de Didáctica de las Matemáticas
http://www.sinewton.org/numeros
ISSN: 1887-1984
Volumen 83, julio de 2013, páginas 65-77

Problemas de mediciones repetidas y de riesgo para desarrollar el razonamiento de estudiantes de secundaria en los temas de media y dispersión

Ernesto Sánchez (Cinvestav. México)
José Antonio Orta Amaro (Cinvestav. México)

Artículo solicitado a los autores por la revista

Resumen

En este artículo se proponen problemas para desarrollar el razonamiento de los estudiantes en los temas de *media* y *dispersión* de un conjunto de datos. Se destacan dos características de las tareas que pueden ser útiles tanto en su utilización como para elaborar otros problemas: a) comparación de grupos de datos, b) contexto (medición y riesgo). Se analizan los problemas indicando las dificultades y posibles respuestas de los estudiantes. La exposición de los problemas está precedida de un conjunto de ideas extraídas de la investigación que ayudan a esclarecer la intencionalidad, el significado y las características transferibles de los problemas propuestos.

Palabras clave Media, dispersión, comparaciones de conjuntos de datos, mediciones repetidas, riesgo.

Abstract

In this paper, some problems for developing students' reasoning about mean and spread of data are proposed. Two tasks feature which are useful for working on them and for constructing other tasks are pointed out: a) comparing groups, b) context (measures and risk). Difficulties of the tasks and some possible students' responses are analyzed. Previous to show the problems, several ideas from research for helping to clarify the intentionality, meaning and transferable features of the problems are considered.

Keywords Mean, spread, comparing groups, repeated measures, risk.

1. Introducción

La estadística tiene un papel destacado en el desarrollo de la sociedad moderna al proporcionar herramientas metodológicas generales para recoger y organizar todo tipo de datos, describir y analizar su variabilidad, determinar relaciones entre variables, diseñar en forma óptima estudios y experimentos y mejorar las predicciones y toma de decisiones en situaciones de incertidumbre. En consecuencia, una persona educada debe ser capaz de entender, y razonar con, la información estadística a la que constantemente está expuesta y, más aún, ser capaz de utilizar los instrumentos de la estadística para generar y analizar datos relevantes para su vida diaria y profesional. Hay consenso entre investigadores y educadores acerca de la necesidad de que la escuela debe favorecer el desarrollo del razonamiento estadístico de los estudiantes.

En la mayoría de países modernos, desde hace décadas, se ha incorporado la estadística en los currículos de los niveles básico, medio y universitario; sin embargo, los resultados aún son pobres. En efecto, la cultura estadística de un ciudadano promedio es inferior de lo que se asienta y pretende en los diversos programas; como lo han constatado los estudios sobre dificultades de estudiantes preuni-

MONOGRÁFICO: ESTADÍSTICA

versitarios (Garfield y Alghren, 1988; Batanero, Godino, Vallecillos, Green y Holmes,1994), falsas concepciones (Shaughnessy, 1992, 2007; Jones, Langrall y Mooney, 2007) y falsas concepciones sobre inferencia de estudiantes universitarios (Castro-Sotos, Vanhoof, Noortgate y Onghena, 2007).

Un problema común en la enseñanza de la estadística de los niveles básicos es su enfoque hacia la transmisión de procedimientos y fórmulas (elaborar gráficas y calcular resúmenes estadísticos) sin que se conviertan, y sean utilizados, como herramientas de análisis. A pesar de que la media y la desviación estándar (en general: centros y medidas de dispersión) son fundamentales en estadística, en la práctica y en problemas en los que sería pertinente considerarlos, no son utilizados por los estudiantes (Gal, 1989,1990; Watson, 1999) Un problema con la media aritmética es que la fórmula matemática es muy simple, pero su interpretación en los diferentes contextos en que se presenta, parece constituir un obstáculo de grandes dimensiones. Por su parte, percibir la dispersión es muy natural (Shaughnessy, 1997) sin embargo, sus expresiones simbólicas son algo complicadas para los estudiantes y, sobre todo, les es muy difícil interpretarla en función del contexto del que provienen los datos.

Los más recientes cambios curriculares recomiendan que los profesores busquen enfocar la enseñanza de la estadística hacia el desarrollo del razonamiento estadístico y no sólo en el aprendizaje de conocimientos, pero para hacerlo, es necesario elaborar y socializar actividades que sirvan para ese propósito. Varios autores ya han dado pasos en este sentido (Batanero y Díaz, 2011, Garfield y Ben-Zvi, 2008) y el presente trabajo pretende constituir una contribución más. Su objetivo específico es proponer y analizar 4 problemas que pueden ser útiles para la elaboración de actividades que ayuden al desarrollo del razonamiento de los temas de media y dispersión en clase de estudiantes de secundaria (14-15 años). Antes de exponer dichos problemas, se revisarán algunas ideas extraídas de la investigación que ayuden a esclarecer la intencionalidad, el significado y la trascendencia de los problemas propuestos.

2. El papel de los problemas en la planificación

A grandes rasgos, las actividades que debe realizar un profesor son: 1) planificar lecciones para cubrir los objetivos curriculares, 2) llevar a la práctica lo planificado y 3) permanentemente evaluar los avances de sus alumnos. La planificación de una lección es una tarea compleja y demandante que exige del profesor interpretar y transformar el conocimiento. Una buena planificación permite realizar las otras dos etapas con mayor eficacia. El problema de cómo pueden aprender los futuros profesores a planificar sus lecciones constituye toda una línea de investigación que no es objeto de este estudio. Rescatamos, sin embargo, una parte de la discusión que sobre este problema ofrece Jones y Smith (1997) cuando responden a la pregunta ¿cómo aprenden los futuros profesores a diseñar las lecciones de matemáticas que implementarán en clase?

Para responder a esta pregunta, Jones y Smith (1997) analizan cómo diseñan sus lecciones los profesores experimentados. Con base en informes empíricos de investigación, observa que el punto de partida del diseño para la mayoría de esos profesores, no se posiciona en los objetivos o aprendizajes esperados, sino que el elemento que dispara el diseño de la lección son las tareas y problemas con las que ya cuenta. Por lo tanto, es frecuente que el proyecto de enseñanza de un profesor esté más determinado por el catálogo de problemas que es capaz de manejar, que por los contenidos específicos y las recomendaciones didácticas del currículo. Por supuesto, ambos: objetivos curriculares y problemas, están íntimamente relacionados y el profesor lleva a cabo una negociación en la que trata de armonizarlos; pero si no cuenta con problemas apropiados es probable que no cumpla convenientemente con los objetivos o los omita en su clase. Por lo anterior, una componente importante en la formación de futuros profesores y en la actualización de profesores en servicio es la elaboración y discusión de problemas relevantes para los estudiantes.

En conclusión, las nuevas propuestas curriculares, cuyo eje es el desarrollo de competencias, necesitan ir acompañadas de problemas y tareas para elaborar actividades que contribuyan a la formación de las diversas competencias que el currículo prescribe. Los problemas que se proponen en este trabajo tienen la intención de contribuir con algunos problemas para promover el razonamiento de los estudiantes en los contenidos de media y dispersión.

3. El razonamiento estadístico

El razonamiento es un proceso mental cuya función es generar ideas (en forma de proposiciones) y apoyar su veracidad. Esos procesos pueden y suelen desembocar en discursos o representaciones simbólicas de enunciados concatenados de manera que dan lugar a una conclusión. Cuando las ideas que se generan son de tipo estadístico se habla de razonamiento estadístico. Garfield y Ben-Zvi se refieren a él de la siguiente manera:

> "El *razonamiento estadístico* es la manera en que la gente razona con las ideas estadísticas y le da sentido a la información estadística […] puede involucrar conexiones de un concepto a otro (por ejemplo centros y dispersión) o combinar ideas acerca de datos y azar. El razonamiento estadístico también significa entender y ser capaz de explicar procesos estadísticos y de interpretar los resultados estadísticos" (Garfield y Ben-Zvi, 2008, pp. 34).

Conviene destacar el aspecto de la interpretación a la que se refiere Garfield y Ben-Zvi en la anterior cita. No es suficiente que mediante un tratamiento estadístico se obtenga una gráfica, un número resumen, un intervalo de confianza o el resultado de un contraste de hipótesis, si tales resultados no se interpretan convenientemente. La interpretación concierne a la relación del resultado de un tratamiento estadístico con el contexto de la situación de donde surgieron los datos y el problema. Wild y Pfannkuch (1999) sugieren que uno de los tipos fundamentales de pensamiento estadístico es la integración de lo estadístico y lo contextual. Es por esto que al proponer un problema para desarrollar el razonamiento estadístico conviene considerar seriamente el contexto, no como un pretexto para formular un problema, sino como una preocupación genuina de entenderlo y aprender algo sobre él a través del análisis de los datos disponibles.

3.1. La media aritmética, la dispersión y sus relaciones

La definición y cálculo de la media aritmética es trivial y los estudiantes la aprenden desde la escuela primaria. No obstante, su uso e interpretación en las situaciones en las que juega un papel importante parecen estar lejos de ser simples para estudiantes de todos los niveles (Russell, 1990, Groth y Bergner, 2006). Batanero (2000) señala que la reflexión sobre las dificultades en el aprendizaje que presentan los alumnos con relación a la media aritmética, debe comenzar por un análisis *epistemológico de su significado*. Para hacerlo, sugiere partir de las situaciones-problema que dan lugar al uso de la media aritmética; identifica cuatro de ellas y un problema para cada una. Las situaciones son de:

- estimación de una cantidad desconocida, en presencia de errores de medida
- repartición equitativa de la totalidad de cantidades diferentes
- elemento representativo de un conjunto de valores dados cuya distribución es aproximadamente simétrica
- predicción del valor o valores con mayor probabilidad al tomar un elemento al azar de una distribución

Después de analizar el concepto con base en otros constructos para analizar el significado: elementos actuativos, intensivos, extensivos, ostensivos y validativos (Godino, 1996), Batanero concluye que la media aritmética y otros conceptos estadísticos, “tienen un significado complejo y por tanto será necesario un periodo dilatado de enseñanza a lo largo de la educación primaria y secundaria para lograr el progresivo acoplamiento de los significados personales que construyen los alumnos a los significados institucionales que pretendemos adquieran” (Batanero,2000, pp.10).

Por su parte, la *dispersión* es una característica de un conjunto de datos numéricos. Hay diferentes medidas de la dispersión, donde las más usadas son: rango, desviación media, desviación estándar, y cuartiles. El cálculo de algunas medidas de dispersión es algo más difícil que el de la media, sin embargo, con cualquier calculadora científica se obtienen fácilmente, una vez que se le introducen los datos. Un significado que se suele dar a las medidas de dispersión es como indicadores de qué tan separados están los datos entre sí. Basta ordenar los datos o hacer su gráfica de frecuencias para percibir intuitivamente la dispersión; en contraste, la dispersión como un número específico es para la mayoría un misterio. Las ideas de separación y la interpretación del valor numérico están en un terreno abstracto, y aunque importantes, faltaría interpretar la dispersión en los contextos en los que se presenta.

Varios investigadores (Garfield y Ben-Zvi, 2008, Konold y Pollatsek, 2004) sugieren que no es posible entender la dispersión sin verla estrechamente ligada a una medida de tendencia central, principalmente a la media aritmética. Por ejemplo, Garfield y Ben-Zvi comentan que “los libros de texto tradicionales primero introducen las medidas de tendencia central, luego introducen la dispersión y luego continúan con el siguiente tópico. Sin embargo, puede ser más útil estudiar esos tópicos juntos pues están interrelacionados” (Garfield y Ben-Zvi, 2008, pp. 88). En consecuencia, es muy importante tener situaciones y problemas para trabajar en clase y propiciar la interpretación de la media y la dispersión de manera conjunta. Las situaciones de comparación de conjuntos de datos son ideales para ese propósito.

3.2. Comparación de conjuntos de datos

Gal (1989, 1990) administró problemas simples de comparación de conjuntos de datos con niños de educación básica (8 a 12 años) y observó las estrategias que utilizaban para analizarlos y responder las preguntas. Un resultado sorprendente es que ninguna de las estrategias que desarrollaron los niños incluía el uso de la media aritmética, cuando ésta era fundamental para la solución. Posteriormente, otras investigaciones han utilizado problemas de comparación de grupos para explorar las estrategias de los estudiantes en el análisis de datos (Watson 1999,2001). Garfield y Ben-Zvi (2008) argumentan a favor del valor que tienen los problemas de comparación de conjuntos de datos y defienden que se incluya como un tópico explícito a tratar en los cursos de estadística; resumen sus razones en los siguientes cuatro puntos:

1. Comparar dos o más grupos puede estructurarse como una versión inicial e informal de inferencia estadística
2. Los problemas que involucran comparaciones de grupos son a menudo más interesantes que los que involucran a un solo grupo
3. Estudiantes de cualquier nivel requieren desarrollar estrategias para comparar grupos de datos
4. En la comparación de grupos es importante realizar representaciones gráficas y obtener resúmenes (centro y dispersión) de los datos.

Con base en estos argumentos, los problemas que se proponen en este artículo son de comparación de conjuntos de datos, pero se ha buscado formularlos en contextos en los que la interpretación de los resultados sea significativa.

4. Contextos favorables para la interpretación de la dispersión

En los apartados anteriores se ha señalado la importancia de contar con problemas y actividades para diseñar lecciones de estadística que permitan cubrir los contenidos curriculares. Dentro de los contenidos de estadística los temas de media aritmética y dispersión son cruciales. El enfoque didáctico de los actuales corrientes curriculares enfatiza en el desarrollo de habilidades de razonamiento y no solo en la adquisición de conocimientos. Con este fin, parecen convenientes las situaciones de comparación de conjuntos de datos, formuladas en contextos significativos. En lo que sigue se proponen dos contextos con potencial para propiciar la petición y búsqueda de razones para justificar una respuesta o decisión.

4.1. Contextos de mediciones repetidas

Cuando en un trabajo experimental se realizan varias mediciones de una magnitud física no se obtiene en general el mismo resultado. Surge entonces el problema de cuál es la mejor aproximación a la verdadera medida de la magnitud en cuestión. La teoría de los errores establece que la mejor aproximación es la *media aritmética* cuando las mediciones cumplen las siguientes características:

- Cada medición es independiente de cualquier otra
- Se pueden cometer con la misma frecuencia, tanto errores por defecto como errores por exceso.

La media aritmética se aproximará tanto más al valor verdadero de la magnitud cuanto mayor sea el número de medidas, ya que globalmente los errores aleatorios se compensan. Una medida o dato (D_i) se puede expresar como la suma de la medida real *T* (desconocida) del objeto más un error: $D_i = T + E_i$, la media de las mediciones será la suma del la medida real más la media de los errores: $\bar{D} = T + \bar{E}$. Como los errores son tanto positivos como negativos se espera que el promedio de los errores sea pequeño.

De acuerdo con la teoría de Gauss de los errores, cuando se supone que estos se producen por causas aleatorias, se toma como la mejor estimación del error, el llamado **error cuadrático** definido por:

$$e = \sqrt{\frac{\sum_1^n (D_i - \bar{D})^2}{n(n-1)}}$$

Esta medida es $s/\sqrt{n}$, donde *s* es la desviación estándar muestral, es decir, del conjunto dado de datos. Esta fórmula confirma y formaliza la idea intuitiva de que la precisión de un conjunto de medidas es función de la dispersión; cuando los datos son muy dispersos corresponden medidas con poca precisión, en cambio, si están muy próximos entre sí significa que las medidas son muy precisas. Las ideas anteriores ofrecen una buena oportunidad para dar una interpretación a los conceptos de media aritmética y dispersión; por ejemplo, una actividad basada en la combinación entre encontrar la mejor estimación a un dato real de un conjunto de mediciones y comparar conjuntos de datos, puede ser promisoria.

4.1.1. Un problema en el contexto de medidas

A continuación, en la Figura 1, se muestra un problema en un contexto de medidas que puede servir para elaborar una actividad que promueva el razonamiento de los estudiantes de la media y la dispersión:

Una clase de ciencias se divide en dos grupos A y B. Cada grupo está formado por 9 estudiantes. Cada uno de ellos, independientemente del otro, midió la altura de un poste y cada grupo utilizó un método diferente. Las longitudes en metros que obtuvieron se muestran en las siguientes listas:

Datos obtenidos por el Grupo A

4.6	4.2	4.7	4.6	4.5	4.3	4.6	4.7	4.3

Datos obtenidos por el Grupo B

4.7	4.4	4.0	3.9	5.2	3.9	4.7	5.0	4.7

- Cada grupo por su lado, con base en sus datos, debe proponer tan exactamente como sea posible la verdadera altura del poste ¿Qué números deben proponer?
- Se quiere evaluar entre el método para obtener las medidas que utilizó el grupo A y el método que utilizó el grupo B, ¿Cuál es más preciso? ¿Por qué?

Figura 1. Ejemplo de un problema en un contexto de mediciones

La respuesta a la primera pregunta es la media de cada conjunto de datos ($\bar{X}_A = 4.5$ y $\bar{X}_B = 4.5$). La respuesta a la segunda pregunta se puede basar en alguna medida de la dispersión; por ejemplo, los rangos son: $r_A = 0.50$,$r_B = 1.3$; se deduce que hay más precisión en las medidas del grupo A. Esta conclusión es la misma si se consideran las desviaciones medias: ($Dm_A = 0.15, Dm_B = 0.4$) las desviaciones estándar ($De_A = 0.18$,$De_B = 0.47$) o el error cuadrático ($e_A = 0.06$,$e_B = 0.16$. La gráfica también puede servir de argumento para apoyar que es más preciso el método del grupo A (ver Figura 2).

Análisis del problema

Las razones por las que la media aritmética es el mejor estimador no son evidentes y, por tanto, no la suelen dar espontáneamente los estudiantes de cualquier nivel (Gal 1989,1990). Una respuesta que prefieren es proponer la moda: 4.6 para el primer conjunto y 4.7 para el segundo. Un problema con esta propuesta de solución es que no toma en cuenta las otras medidas. Otra respuesta posible es el punto medio del recorrido, por ejemplo: $Pm_A = 4.45, Pm_B = 4.55$ y aunque ésta es una mejor propuesta que la de proponer la moda, no es tan buena como la media. La mediana también es una propuesta conveniente, sin embargo, esperar que la consideren los estudiantes es tan difícil como la propia media (en el ejemplo, casualmente las medianas coinciden con las modas).

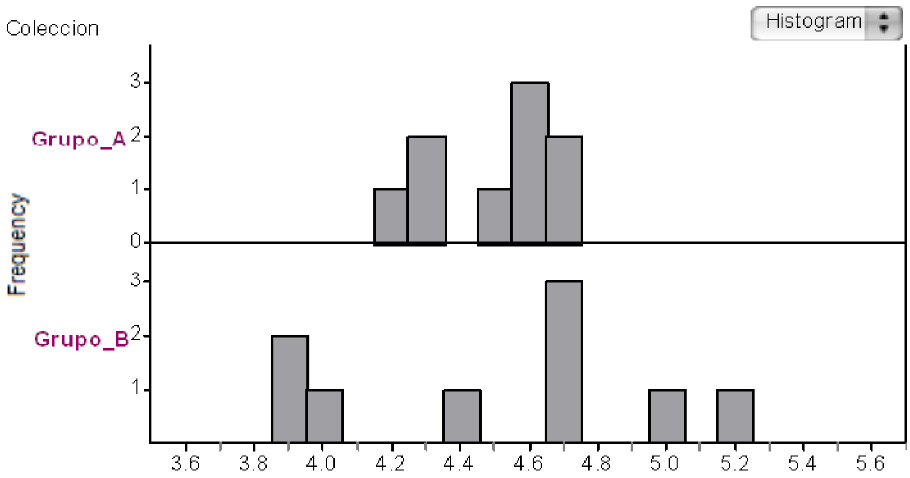

Figura 2. Gráficas del problema de mediciones

En la clase, se debe permitir que los estudiantes hagan sus propuestas y las discutan. El profesor puede plantear que la solución debe tener en cuenta todas las medidas y no sólo una o un sub-conjunto de ellas; también tiene que enfatizar en el hecho de que los errores son aleatorios y pueden ser tanto por exceso como por defecto y que la solución debiera aprovechar esta circunstancia. Estas razones pueden contribuir a entender que la media es un buen candidato. El profesor debe abstenerse de dar la solución desde un principio, pues frustraría la posibilidad de reflexión de sus estudiantes, en su lugar, tiene que desplegar recursos para motivar la discusión; cuando algún estudiante proponga la media no la debe aceptar de inmediato, sino propiciar que convenza a sus compañeros con argumentos válidos. El análisis de la precisión en este contexto no es tan difícil de percibir y de que surja naturalmente en boca de los estudiantes. La cuestión es no dejar sólo que perciban intuitivamente al grupo de datos con mayor precisión (menos dispersión) sino aprovechar dicha percepción para asociarla con una medida de dispersión.

4.2. Contextos de riesgo

A continuación se presentan algunas ideas básicas sobre el riesgo basadas en la exposición de Fischhoff y Kadvany (2011). En particular, para los fines de este artículo basta mencionar algunas ideas de tres temas que abordan estos autores: la definición de riesgo, el análisis del riesgo y la toma de decisiones en situaciones de riesgo. El riesgo se presenta cuando hay potenciales resultados no deseados que pueden traer como consecuencia pérdidas o daños. Definir el riesgo significa especificar los resultados valiosos y los no deseados en un orden que refleje el valor que se les atribuye. En muchas ocasiones, se pueden medir los resultados no deseados o valiosos, por ejemplo, las tasas de mortalidad o el producto interno bruto. Hay también eventos en los que no hay acuerdo de cómo medirlos: pobreza, bienestar o crecimiento económico. Finalmente, hay algunos en que la sola idea de medirlos es controversial, así son las amenazas a la justicia o los daños ecológicos a la naturaleza.

Una vez que se identifican los riesgos, el problema es determinar qué tan grandes o prejudiciales son y cuáles son las posibles causas que los gobiernan. El análisis del riesgo generalmente son constructos intrincados, a menudo integran los análisis de científicos de diferentes áreas del conocimiento y toman en cuenta una diversidad de fuentes de datos o información. No obstante, hay una variedad tan amplia de fenómenos, que la lógica del análisis del riesgo puede consistir en simples conteos de datos, con cuyos resultados llevar a cabo conocidos análisis estadísticos, hasta complejas simulaciones de modelos cibernéticos conteniendo información científica de áreas especializadas además de intrincados procesos estocásticos.

El análisis del riesgo ofrece información para la toma de decisiones. La teoría de la toma de decisiones en situaciones de riesgo tiene dos aspectos; por un lado, define reglas abstractas sobre lo que debería hacer la gente en situaciones de riesgo, por otro, estudia lo que hace la gente cuando se enfrenta realmente a tales situaciones: “Si la gente no sigue las reglas, quiere decir que o las personas necesitan ayuda o las reglas necesitan una revisión” (Fischhoff y Kadvany, 2011, pp.65). En los casos en que el análisis del riesgo ofrece un conjunto ordenado de posibles resultados, la regla para tomar decisiones es simple: “elija la opción cuyo resultado produzca algo con la mayor cantidad del valor que usted desea tener (dinero, descanso, acres de tierras húmedas, etc.)” (Fischhoff y Kadvany, 2011, pp. 65). La ordenación de los resultados generalmente contiene incertidumbre y está asociada a distribuciones de probabilidad. La regla entonces debe combinar el valor deseado que proporciona cada resultado con la probabilidad de que ocurra para cada posible decisión.

4.2.1. Un problema en el contexto de juegos

No iremos más allá de los comentarios anteriores, pues son suficientes para indicar la naturaleza del contexto de riesgo en el que se formula el problema que se propone (Figura 3). En el análisis del problema aclararemos el papel de la media y la dispersión para la toma de decisiones en la situación de riesgo que se propone.

En una feria, se invita a los asistentes a participar en uno de dos juegos. Juan puede participar en un juego, pero no en ambos. Para saber por cual decidirse observa, anota y ordena los resultados de dos muestras de 10 personas que han participado en cada juego. Las pérdidas (-) o premios (+) en efectivo que han obtenido las 20 personas se muestran en las siguientes listas:

Juego 1

15	-21	-4	50	-2	11	13	-25	16	-4

Juego 2

120	-120	60	-24	-21	133	-81	96	-132	18

Si tienes la posibilidad de participar en un solo juego ¿Cuál juego elegirías? Explica las razones de tu respuesta.

Figura 3. Ejemplo de un problema sobre situaciones de riesgo

Una respuesta relativamente aceptable es que *es igual jugar en cualquiera de los dos juegos*, esto con base en el argumento de que ambos conjuntos tienen la misma media ($\mu = 4.9$), es decir, a la larga en cualquiera de los juegos se ganará en promedio 4.9 unidades. No obstante, una respuesta más avanzada debe tener en cuenta la dispersión. Es notorio que en los datos del segundo juego hay mucha más dispersión que en los del primero (Figura 4). En este contexto, la dispersión se asocia al riesgo: el juego 2 es más arriesgado, pues aunque con base en los datos se puede conjeturar que es posible ganar cantidades mayores que en el juego 1 (el máximo en el juego1 es 133, mientras que en el juego 2 es 50) también es posible perder más (el mínimo en el juego 1 es -132, mientras que en el juego 2 es -25). La decisión depende de las preferencias y condiciones económicas del jugador, si es propenso al riesgo y tiene suficientes recursos para jugar, podría optar por el juego 2; si es conservador o no tiene suficientes recursos, le conviene elegir el juego 1.

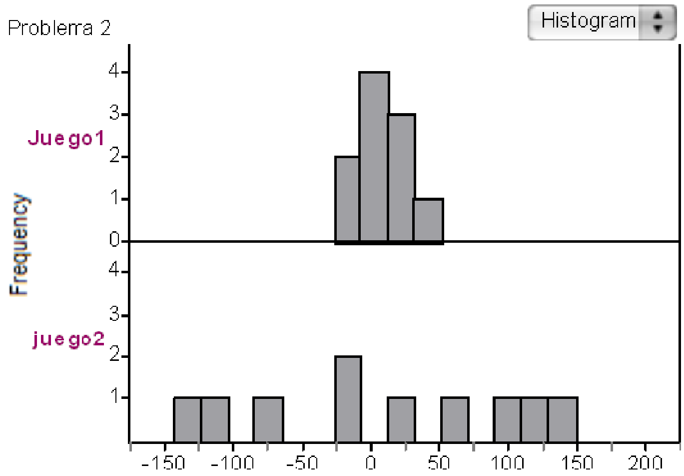

Figura 4. Gráficas del problema de riesgo

Análisis del problema

La idea del problema es propiciar que el estudiante se ubique en una situación de riesgo, de manera que se dé cuenta que la elección de un juego u otro conlleva consecuencias en términos de beneficios-pérdidas. El primer paso en el análisis es preguntarse en qué juego globalmente se gana más dinero; esta pregunta se la hacen los estudiantes de manera natural, pero la respuesta que ofrecen es muy variada. Nuevamente pensar en la ganancia promedio de cada juego (media aritmética de los datos) y compararlas, es una idea que pocos, si alguno, la consideran; la gran mayoría omiten este paso que es crucial para el análisis. Sus comparaciones suelen ser resultado de una impresión de riesgo que les producen los datos individuales de la distribución: unos comparan los máximos de las distribuciones y hacen afirmaciones como “en el juego 2 se gana más”, otros comparan los mínimos “en el juego 1 se pierde menos”; otros podrían ignorar la cantidad precisa que se gana o se pierde y sólo cuentan cuantos perdedores (cantidades negativas) y cuantos ganadores (cantidades positivas) tiene cada juego, decidiendo que da lo mismo escoger uno u otro.

Una vez que se determina con base en la ganancia promedio que los dos juegos son equivalentes, se debe formular la pregunta ¿En qué juego se corre más riesgo? Inferir que el juego con mayor dispersión es más arriesgado que el juego con menor dispersión ofrece un significado pertinente a la teoría del riesgo. El problema de valorar el riesgo consiste en ponderar entre ambos juegos: En el de mayor dispersión, se puede tener la fortuna de ganar más que en el otro juego, pero también se corre el riesgo de perder más. En el de menor dispersión, se puede ganar menos que en el primero, pero también, si se pierde se perderá menos que en el otro juego. No hay una regla general que indique cuál es la mejor decisión, pues depende de la situación económica particular del jugador así como de sus creencias y actitudes ante el riesgo.

Una dificultad que suelen encontrar los estudiantes surge de la falta de información acerca de la naturaleza del juego y sobre la manera particular en que un jugador gana determinada cantidad. Muchos estudiantes llenan este vacío imaginando situaciones: unos pueden pensar en una urna con bolas numeradas, otros en juegos de caballos, etc. Encontramos que algunos estudiantes imaginan que el primer dato del primer juego (15) corresponde a un jugador y que el primer dato del segundo juego corresponde a su oponente (120) y concluye que el jugador del segundo juego le ganó al primero (por 105). Situaciones en las que se pueden coleccionar datos ignorando el modelo subyacente son frecuentes en la investigación estadística. La idea del problema es propiciar que el estudiante realice un estricto análisis de los datos interpretando la situación a pesar de dichas restricciones. El siguiente ejemplo es también de una situación de riesgo y en él es más claro la ausencia de un modelo para explicar el comportamiento de los datos.

4.2.2. Un problema en el contexto de salud-enfermedad

Considera que debes aconsejar a una persona que padece una enfermedad grave, incurable y mortal, pero que es tratable con medicamentos que pueden extender la vida por varios años más. Es posible elegir entre tres tratamientos. Las personas tienen diferentes reacciones a las medicinas, para algunas tienen el resultado previsto, mientras que para otras pueden ser más benéficas o más prejudiciales. En las siguientes listas se muestran los años que han vivido varios pacientes que se han tratado con una de las opciones mencionadas; cada dato de las listas corresponde al tiempo que ha sobrevivido un paciente con el respectivo tratamiento. Después se muestran las gráficas correspondientes a los tratamientos.

Tiempo en años (tratamiento 1)

5.2	5.6	6.5	6.5	7.0	7.0	7.0	7.8	8.7	9.1

Tiempo en años (tratamiento 2)

6.8	6.9	6.9	7.0	7.0	7.0	7.1	7.1	7.2	7.4

Tiempo en años (tratamiento 3)

6.8	6.8	6.9	7.0	7.0	7.1	7.1	7.1	7.2	7.4

Figura 5. Ejemplo de problema es situaciones de riesgo

El primer paso en el análisis de este problema, al igual que los anteriores, es calcular el tiempo medio de vida para cada tratamiento ($\mu = 7.04$) . Con base en esta información se podría decir que cualquier tratamiento es igual, no obstante, en la medida en que lo que está en juego es de suma importancia, parece imprescindible hacer un análisis (de riesgo) más detenido. Claramente el tratamiento 1 tiene mayor dispersión (Figura 6) y, por tanto, ofrece más riesgo, en el mismo sentido del problema anterior: los datos indican que se puede vivir más tiempo que el tiempo que indican los datos en cualquiera de los otros dos tratamientos (9.1 > 7.4), pero también se puede vivir menos (5.2 < 6.8). ¿Cuál es la decisión adecuada entre elegir el primer tratamiento y cualquiera de los otros dos? Nuevamente, no hay una regla general que prescriba la mejor elección, aunque en este caso una decisión conservadora podría ser más popular que una decisión arriesgada. Decidir entre el tratamiento 2 y el 3 es más difícil, pero si el criterio de menor riesgo se ha tomado y éste se asocia con la desviación estándar muestral, los datos del tratamiento 2 indican menor dispersión ($S_1 2 = 0.17, S_1 3 = 0.18$) , aunque la diferencia es tan pequeña que se pueden considerar equivalentes.

Análisis del problema

El análisis de este problema es muy similar al anterior sobre la elección entre dos juegos. El primer paso es comparar las medias, después considerar con cuidado la dispersión. Como se ha indicado, este sencillo procedimiento no suele estar dentro de las estrategias espontáneas de los estudiantes y bien puede ocurrir, que los estudiantes hayan entendido que en el contexto de pérdida-ganancia es conveniente comparar las medias de los conjuntos de datos y no transferirlo a la situaciones de tiempo de vida en contextos de salud-enfermedad. Con relación al análisis de la dispersión, puede ocurrir que un mismo alumno sea propenso al riesgo en el caso de juegos de pérdida-ganancia de dinero, y sea adverso al riesgo en el caso de salud-enfermedad o viceversa. Lo importante de la consideración del riesgo es que lo interpreten de manera adecuada y no caigan en el error frecuente de creer que “mayor dispersión significa mayor tiempo de vida”, pues esto es falso.

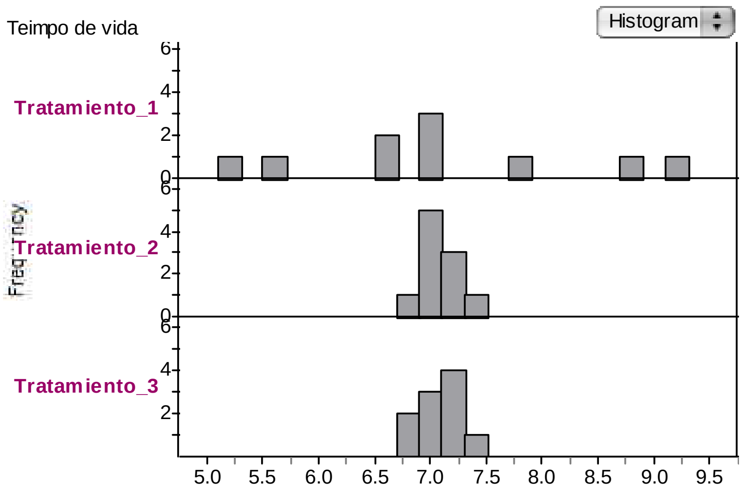

Figura 6. Gráficas de los datos del problema del tiempo de vida

5. Conclusiones

Se han propuesto tres problemas que pueden ser útiles para diseñar lecciones de estadística para desarrollar el razonamiento estadístico de los estudiantes en los temas de media y dispersión. Las razones para hacerlo son variadas: la importancia de que el profesor cuente con problemas para cubrir los contenidos curriculares, las dificultades detectadas en el tema en cuestión y la necesidad de darles un significado que sea accesible a los alumnos. Los problemas se han adaptado al formato de comparación de conjuntos de datos, pues prefiguran problemas de inferencia, atraen el interés de los estudiantes, no son triviales y su solución requiere de los instrumentos elementales de la estadística. También se han inscrito en los contextos de mediciones repetidas (ya señalado por Batanero, 2000) y de riesgo, pues ambos ofrecen oportunidades para enriquecer el significado de las nociones de media y dispersión. La intención es explorar recursos que permitan ir más allá de las interpretaciones de la media y la dispersión que se constriñen a contextos puramente matemáticos. Se han hecho algunas recomendaciones para implementarlos en clase, acordes con el espíritu actual de las recomendaciones curriculares: dejar a los estudiantes que propongan soluciones, promover la discusión y prever los errores y falsas concepciones que ya han sido detectadas en la investigación. En el futuro, investigaciones empíricas y prácticas escolares podrían proporcionar evidencias de las ventajas y desventajas del uso de este tipo de problemas en clase.

MONOGRÁFICO: ESTADÍSTICA

Bibliografía

Batanero, C. (2000). Significado y comprensión de las medidas de posición central. *UNO, 25*, 41-58.

Batanero, C. y Díaz, C. (2011) (Eds.). *Estadística con Proyectos.* Granada: Departamento de Didáctica de las Matemáticas.

Batanero, C., Godino, J. D., Vallecillos, A., Green, D. R., y Holmes, P. (1994). Errors and difficulties in understanding elementary statistical concepts. *International Journal of Mathematics Education in Science and Technology, 25*(4), 527–547.

Castro-Sotos, A. E., Vanhoof, S., Noortgate, W. y Onghena, P. (2007). Students' misconceptions of statistical inference: A review of the empirical evidence from research on statistics education. *Educational Research Review*, 2, 98–113.

Fischhoff, B. y Kadvany, J. (2011). *Risk: A very short introduction*. Oxford: Oxford University Press.

Gal, I., Rothschild, K. y Wagner, D. A. (1989) Which group is better? The development of statistical reasoning in elementary school children. *Artículo presentado en el encuentro de la Sociedad para la Investigación en el desarrollo del Niño (Society for Research in Child Development).* Kansas City.

Gal, I., Rothschild, K. y Wagner, D. A.(1990). Statistical concepts and statistical reasoning in school children: Convergence or divergence? *Artículo presentado en el encuentro de la Asociación de Investigación Educacional Americana* (*American Educational Research Association*), Boston.

Garfield, J. y Ahlgren, A. (1988). Difficulties in learning basic concepts in statistics: Implications for research. *Journal for Research in Mathematics Education*. *19*, 44-43.

Garfield, J. y Ben-Zvi, D. (2005). A framework for teaching and assessing reasoning about variability. *Statistics Education Research Journal*, 4(1), 92-99.

Garfield, J. y Ben-Zvi, D. (2008). *Developing Students' Statistical Reasoning: Connecting Research and teaching*. Springer: New York.

Godino, J.D. (1996). Mathematical concepts, their meaning, and understanding. En: L Puig y A. Gutierrez (eds.) *Proceedings of the 20th Conference of the International Group for the Psychology of Mathematics Education*, 2-417-424. Universidad de Valencia.

Groth, R. E. y Bergner, J. A. (2006). Precervice elementary teachers' conceptual and procedural knowledge of mean, median, and mode. *Mathematical Thinking and Learning, 8*, 37-63.

Jones, K. y Smith, K. (1997). Students teachers learning to plan mathematics lessons. Artículo presentado en la Conferencia Anual de 1997 de la Asociación de Profesores de Matemáticas (*1997 Annual Conference of the Association of Mathematics Education Teachers: AMET1997).* Leicester. 15-17.

Jones, G. A., Langrall, C. W. y Mooney, E. S. (2007). Research in probability: Responding to classroom realities.En F. K. Lester Jr (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* ,909-955. Reston, VA: The National Council of Teachers of Mathematics.

Konold, C. y Pollatsek, A. (2004).Conceptualizing an average as a stable feature of a noisy process. En D. Ben-Zvi y J. Garfield (Eds.). *The challenge of developing statistical literacy reasoning and thinking*, 169-199. Netherlands: Kluwer Academic Publishers.

Russell, S.J. (1990). Issues in training teachers to teach statistics in the elementary school: A world of uncertainty. En A. Hawkins (Ed.), *Training teachers to teach statistics: Proceedings of the international statistical institute roundtable conference,59-71*. Voorburg, The Netherlands: International Statistical Institute.

Shaughnessy, J. M. (1992). Research in probability and statistics: Reflections and directions. En D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (465-494). Macmillan: New York.

Shaughnessy, J. M. (1997). Missed opportunities in research on the teaching and learning of data and chance. In F. Bidulph y K. Carr (Eds.), *Proceedings of the Twentieth Annual Conference of the Mathematics Education Research Group of Australasia* , 6-22. Rotorua: University of Waikata.

Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. En F. K. Lester Jr. (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning*, 957-1009. Reston, VA, USA: The National Council of Teachers of Mathematics.

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

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

Wild, C. J. y Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, *67*(3), 223-265.

Ernesto Sánchez Sánchez. Departamento de Matemática Educativa del Cinvestav, México. Es matemático por la UNAM y Doctor en Ciencias, Especialidad en Matemática Educativa por el Cinvestav. Ha realizado estancias sabáticas de investigación en Grenoble, Francia (1997), Portland, Estados Unidos (2003) y Granada, España, (2011). Su investigación ha sido en el área de la didáctica de la probabilidad y la estadística. Nació en Oaxaca, México, 1955.

José Antonio Orta Amaro. Estudiante de doctorado del Departamento de Matemática Educativa del Cinvestav, México. Es Ingeniero Industrial de profesión, pero se ha dedicado a la docencia. Ha impartido clases de matemáticas y estadística en los niveles básico y medio superior. Está iniciando su formación como investigador en el área de didáctica de la estadística. Nació en México, 1973.