FUNDAMENTACIÓN DE UNA PROPUESTA PARA EL APRENDIZAJE DE LA PROBABILIDAD FRECUENCIAL

Yessica del Rosario May Pech, Zuleyma Sarahí Pérez Moguel

Resumen

Este trabajo presenta una propuesta didáctica de aprendizaje dirigida a estudiantes de nivel medio superior que cursan la asignatura de Probabilidad y Estadística, en la que se pretende desarrollar la noción de "Probabilidad Frecuencial" al reconocer la estabilización de las frecuencias relativas de un suceso aleatorio hacia un valor numérico fijo a medida que aumenta el número de ensayos o pruebas en un experimento, para posteriormente reconocer dicho valor numérico como la probabilidad frecuencial del suceso. A manera de introducción se propone una simulación digital de un suceso aleatorio que permita al estudiante observar y estudiar el comportamiento y regularidades de las frecuencias relativas en gráficas de polígonos de frecuencias de dicho suceso, complementando dicho estudio con actividades que favorezcan el análisis y la interpretación crítica de datos estadísticos y tablas numéricas, para tomar decisiones.

Palabras clave: Probabilidad frecuencial, estabilización de frecuencias, tendencia, aleatoriedad.

Introducción

La importancia de la probabilidad frecuencial reside en que cuantifica el grado de ocurrencia de un evento de manera empírica (a posteriori) y no deductiva, lo que permite predecir la posibilidad de ocurrencia de la mayoría de los fenómenos aleatorios que nos rodean, por medio de la enumeración estadística de la frecuencia de un evento observado, calculando así las probabilidades cuando las distintas causas o características de dicho evento no pueden determinarse a priori (Sosa, 2006).

Sin embargo, la enseñanza del concepto de probabilidad frecuencial suele limitarse al cálculo de la frecuencia relativa de un evento y se trata a este resultado como la probabilidad frecuencial, existiendo así una falta de comprensión del significado de la estabilidad de las frecuencias relativas (necesario para el concepto de probabilidad en el enfoque frecuencial), lo cual se asume como problemática en este trabajo. Por otra parte, en la realidad escolar, no se favorece la experimentación de las tendencias de las series aleatorias y no se tratan situaciones de naturaleza empírica, al utilizar ejemplos relacionados a los juegos de azar, los cuales tienen resultados predecibles, restándole importancia a la naturaleza a posteriori de la probabilidad frecuencial (Serradó, Cardeñoso y Azcárate, 2005).

Por lo mencionado anteriormente, en este trabajo se aborda el concepto fenómeno aleatorio cuando el estudiante interactúe con la simulación digital de un evento, el cual le permitirá observar los resultados que se obtienen al realizar dicho evento, y cómo estos resultados varían e influyen en la probabilidad de ocurrencia del suceso de acuerdo a cuantos más ensayos se efectúen.

Así, esta propuesta tiene como propósito determinar la probabilidad de ocurrencia de un evento de naturaleza aleatoria a posteriori, como la tendencia a la que se estabilizan las frecuencias relativas de dicho evento al realizar el experimento un número suficientemente grande de veces, a través de registros gráficos y numéricos, para la toma de decisiones.

Desarrollo conceptual y aprendizaje de la probabilidad frecuencial

El concepto de probabilidad frecuencial surge por la necesidad de medir el grado de ocurrencia de sucesos aleatorios que no pueden determinarse a priori. Estudiosos de la probabilidad como Graunt (1620-1675) y Petty (1623-1687) fueron de los primeros en intentar dar solución a situaciones no relacionadas con los juegos de azar, abordando problemas sobre demografía y percatándose que "cuantas más observaciones se hacían, más precisos eran los resultados" (Salinero, sf). Jakob Bernoulli (1654-1705) retoma estas ideas para destacar que "lo que no se puede hallar a priori, se puede obtener a posteriori" y considerando la repetición de un experimento bajo las mismas circunstancias un número grande de veces, establece su teorema conocido con el nombre de Ley de los Grandes Números o Teorema de Bernoulli que, a la vez, sirve de base a la noción intuitiva de probabilidad como medida de la frecuencia relativa (probabilidad frecuencial) (Salinero, sf; Batanero, 2005; Sosa, 2006). De manera simple, el teorema es enunciado así:

> "Si la probabilidad de que ocurra un hecho en una prueba única es **p**, y si se hacen varias pruebas independientemente y en las mismas condiciones, la *proporción* más probable de que ocurran los hechos en el número total de pruebas es también **p**; aún más, la probabilidad de que la proporción en cuestión difiera de **p** en menos que una cantidad dada, por pequeña que sea, aumenta al mismo tiempo que aumenta el número de pruebas" (Sosa, 2006, p.32).

El teorema establece una relación entre las probabilidades matemáticas y las frecuencias relativas, de los sucesos aleatorios que aparecen en diversos campos, como determinadas en una forma puramente empírica (Sosa, 2006). Otros seguidores de esta visión fueron Ellis, Bolzano, John Venn, Cournet, Richard Von Mises, Reichenbach, R. Fisher, A. Wald y E. Ternier, quienes admitían el axioma conocido con el nombre de "límite de Venn", el cual enunciaba lo siguiente: "Si un suceso ocurre en gran número de veces, entonces la probabilidad de ocurrencia del suceso es el límite cuando el número de pruebas tiende a infinito del cociente entre el número de veces que se presenta el suceso y el número total de pruebas". Es decir, se postula la existencia de dicho límite (Franquet, 2008).

Para el desarrollo de esta propuesta, se adopta la visión de Von Mises (1920) descrita en Franquet (2008), quien define el concepto de probabilidad para eventos que resultan de considerar experimentos aleatorios que tienen una "regularidad estadística": al realizar un experimento se observa la aparición o no de un suceso **S**, y ante una sucesión indefinida de experimentos se tiene, sucesivamente, un nuevo valor de la frecuencia relativa **v/n** (donde **v** es la frecuencia absoluta del suceso **S** y **n** el número total de pruebas); pese al comportamiento caprichoso de los resultados individuales, la frecuencia relativa presenta una regularidad insistente al realizar el experimento un gran número de veces, tendiendo a aproximarse a cierto número fijo **P(S)**, es decir, un número hipotético hacia el cual tiende la frecuencia relativa al estabilizarse y el cual es la probabilidad del evento **S** (Sosa, 2006; Franquet, 2008).

De acuerdo a Kolmogorov (1979), este número **P(S)** representa una ley probabilística útil para explicar la ocurrencia o no de ese suceso **S** en particular, donde no se conocen a priori todas las relaciones/parámetros que intervienen, y permite tomar decisiones a partir de dicho resultado. Cabe mencionar que estas leyes probabilísticas sólo aparecen, justamente, en procesos masivos, de aquí la importancia de la Ley de los Grandes Números. Esta ley probabilista, Kolmogorov (1979) la explica de la siguiente manera:

> "La afirmación de que un suceso **A** ocurre en las condiciones **S** con una determinada probabilidad $\boldsymbol{P(A/S) = p}$ equivale a decir que en una serie suficientemente larga de ensayos (es decir, realizaciones del sistema de condiciones **S**), las frecuencias $\boldsymbol{v_r = \mu_r / n_r}$ de ocurrencia del suceso **A** (donde $\boldsymbol{n_r}$ es el número de ensayos realizados en la r-esima serie, y $\boldsymbol{\mu_r}$ el número de ellos en que ocurre **A**) son aproximadamente idénticas unas a otras y están próximas a **p**. La hipótesis de la existencia de una constante $\boldsymbol{p = P(A/S)}$ (determinada objetivamente por la relación entre el sistema de condiciones **S** y el suceso **A**) tal que las frecuencias **v** se aproximen más (hablando en términos generales) a **p** a medida que aumenta el número de ensayos, está justificada en la práctica para una amplia clase de sucesos. Los sucesos de este tipo se denominan normalmente aleatorios o estocásticos" (Kolmogorov, 1979, p.271).

En la actualidad, la falta de comprensión del significado de la estabilidad de las frecuencias relativas, de acuerdo a Kahneman, Slovic y Tversky (1982) citados en Serradó, Cardeñoso y Azcárate (2005) se ve reflejado al considerar que la estabilidad de las frecuencias se cumple en un número limitado de pruebas y tomar decisiones a partir de ello. Esto se ve asociado a la persistencia de sesgos y heurísticas como son:

- La *insensibilidad del tamaño de la muestra*, la cual conduce a pensar que cualquier muestra tomada de una población, aunque se trate de una muestra de pequeño tamaño, debe reproducir todas las características de la población, como si se creyese en una "ley de los pequeños números" (Barragués y Guisasola, 2007);
- El *heurístico de la representatividad*, que consiste en evaluar la probabilidad de un suceso de acuerdo a la representatividad del mismo respecto a la población de la que proviene, es decir, se asignan probabilidades altas a aquellos sucesos que parecen ser prototípicos de una población y bajas probabilidades a los que parecen no serlo, produciéndose una confianza indebida en las pequeñas muestras (Serrano, Batanero, Ortíz y Cañizares, 1998; Guisasola y Barragués, 2002);
- La *falacia del jugador*, donde se tiende a pensar que si en una serie de experimentos se produce una racha de un mismo resultado, se espera entonces que aumente la probabilidad del resultado contrario en el próximo experimento, considerando que la estabilidad se puede dar en series limitadas de números (Serrano et al, 1998; Serradó et al, 2005) y el

- *Outcome approach*, que lleva a los sujetos a considerar que cada una de las repeticiones de un experimento está aislada y que no tienen por qué guardar relación con las anteriores o posteriores (Serrano, Batanero y Ortíz, 1996).

Todo esto se refuerza, en la realidad escolar, al introducir ejemplos que utilizan un número reducido de experimentos o bien limitarse al cálculo de la frecuencia relativa de un evento y tratar a este resultado como la probabilidad frecuencial, sin favorecer la experimentación de las tendencias de las series aleatorias.

Es por ello que en este trabajo se abordará el concepto de “Probabilidad Frecuencial” de modo que se analicen las regularidades que se presentan al repetir un experimento (Álvarez, 2011) y se trabaje el vínculo construido entre la frecuencia relativa y la probabilidad, tomando consciencia de la estabilización de las frecuencias relativas a medida que aumenta el número de pruebas para introducir gradualmente la aproximación frecuencial de probabilidad, favoreciendo el análisis e interpretación de datos críticamente, que permitan tomar decisiones y opiniones fundamentadas sobre la base de estudios estadísticos (Barragués y Guisasola, 2009).

Método

Para el desarrollo del concepto “probabilidad frecuencial”, en esta propuesta se resalta la importancia del tamaño de la muestra y la tendencia de las frecuencias relativas de un evento, introduciendo el uso de una applet (tomada y modificada del repositorio oficial de construcciones y recursos relacionados con el software GeoGebra: http://www.geogebratube.org/student/m16503), atendiendo a la sugerencia de Batanero (2006), de manera que permita experimentar las muestras de tamaño creciente, para posteriormente enfocarnos en el uso de gráficas de polígonos de frecuencias y tablas numéricas.

Para el logro de este propósito, se plantearon tareas de aprendizaje divididas en dos momentos:

El primer momento, se inicia con el análisis de los registros gráficos, en los cuales el estudiante deberá interpretar el comportamiento tendencial de las frecuencias relativas para tomar decisiones. Posteriormente el estudiante debe determinar el valor al cual tienden dichas frecuencias relativas a partir del análisis de tablas numéricas. En una tercera parte del primer momento, se pretende que el estudiante reconozca a la estabilización de la tendencia de las frecuencias relativas como la probabilidad frecuencial de un suceso aleatorio y a partir de esta probabilidad tomar una decisión. Para ello, se dispondrá de una computadora y una hoja de actividad por cada bina, en la que se les presenta una situación sobre “tiro con arco” y la cual podrán simular mediante un applet, de manera que el estudiante observe la aleatoriedad del evento estudiado (acertar a la diana) y cómo se registran los resultados de este evento en un polígono de frecuencias al realizar los ensayos. Los estudiantes pueden correr el applet cuantas veces sea necesario para asociarlo e interpretar las representaciones gráficas y numéricas que se le presentan a continuación en la hoja de actividades.

El siguiente applet Geog_simulacion, simula los tiros a la diana de un arquero cualquiera y registra los aciertos obtenidos en cierta cantidad de tiros:

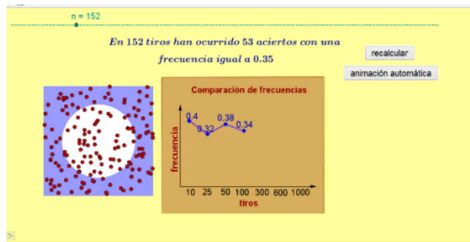


Figura 1. Captura del applet Geog_simulación

Los valores en la gráfica representan la frecuencia relativa de la cantidad de aciertos a la diana en cierta cantidad de tiros.

a) En la Figura 2 y Figura 3 siguientes, se representan los polígonos de frecuencias de los aciertos a la diana de los arqueros A y B, en una serie de ensayos de 25 tiros por día durante 12 días.

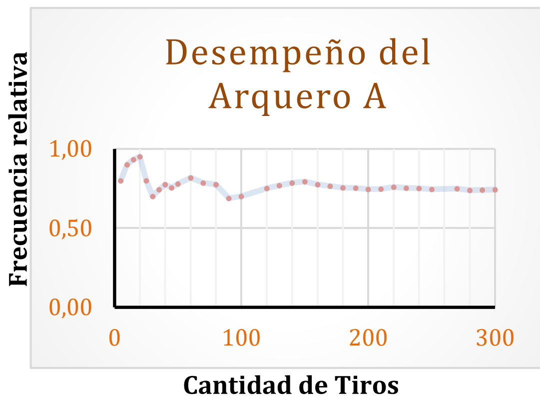


Figura 2. Polígono de frecuencias de los aciertos del Arquero A

Figura 3. Polígono de frecuencias de los aciertos del Arquero B

- Al realizar 20 tiros, el Arquero A tuvo aproximadamente 15% más de aciertos que el Arquero B. Describe si se mantiene esta relación entre los 100 y 150 tiros.
- Al finalizar los 300 tiros, ¿quién de los dos arqueros fue el más certero? Argumente su respuesta.

b) Los resultados de los aciertos totales obtenidos por los arqueros después de 36 días, se anotaron en la Tabla1 siguiente:

	Cantidad de Tiros	20	50	100	150	200	300	350	400	450	500	600	700	800	850	900
Arquero A	**Aciertos**	19	39	70	119	140	223	255	284	312	338	410	470	535	565	600
	Frecuencia relativa (aciertos/tiros)	.95	.78	.70	.79	.75	.74	.73	.71	.69	.68	.68	.67	.67	.66	.67
Arquero B	**Aciertos**	16	39	70	112	124	198	240	283	330	378	460	541	615	655	700
	Frecuencia relativa	.80	.78	.70	.75	.62	.66	.69	.71	.73	.76	.77	.77	.77	.77	.78

Tabla 1. Frecuencias relativas de los aciertos de los arqueros A y B

- A partir de los datos anteriores, se determinaron tres valores que representen la medida de acertar a la diana.

Arquero A: 0.67 0.75 0.95

Arquero B: 0.66 0.77 0.80

¿Cuál valor representa mejor esa medida? Justifique su respuesta.

c) Considere la siguiente conjetura:

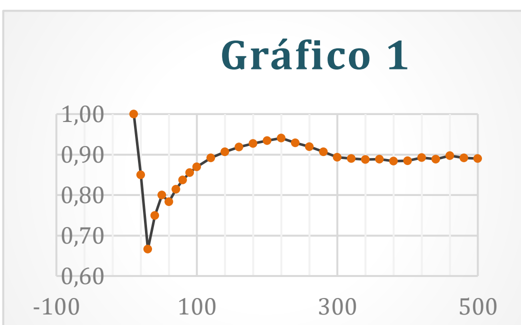
"Para cada arquero, existe un número al que se aproximan las frecuencias relativas de los aciertos a la diana al incrementarse el número de tiros. Cuantos más tiros se realicen, mejor será la estabilidad de las frecuencias hacia dicho número. Este número es la probabilidad frecuencial de acertar a la diana".

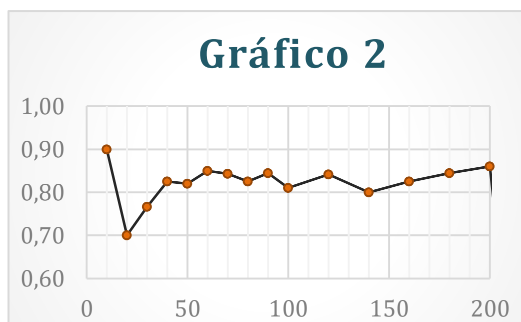
- De acuerdo a la conjetura anterior y lo estudiado en los incisos a y b, ¿qué arquero tiene mayor probabilidad de acertar a la diana en sus tiros? ¿A cuál seleccionaría como el más certero? Argumente su respuesta.

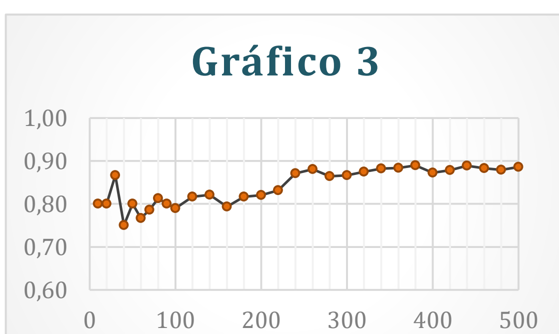
Finalmente, en el segundo momento y a manera de valoración de la probabilidad frecuencial, se pretende que el estudiante relacione a ésta como la estabilización de las frecuencias relativas para representar el comportamiento de las mismas y asociarla a la ocurrencia de un evento, realizados **n** ensayos.

Un arquero tiene una probabilidad de .89 de acertar a la diana.

a) De las tres gráficas que se te presentan a continuación, ¿cuál representa mejor los registros de los aciertos a la diana dados por esta probabilidad? Justifica tu respuesta.







b) Si el total de tiros realizados fue 500, ¿aproximadamente cuántos aciertos tuvo el arquero?

Dichas tareas se realizarán en binas, esperando generar un ambiente de discusión y análisis por parte de los estudiantes. Con un tiempo estimado para su resolución de una hora y veinte minutos, se propone realizarla en un centro de cómputo. El profesor actuará como instructor en las actividades y moderador al momento de formalizar los resultados.

Reflexiones finales

Consideramos pertinente la idea propuesta para el aprendizaje del concepto de probabilidad frecuencial, ya que rompe los esquemas tradicionales y erróneos, favoreciendo la experimentación del suceso aleatorio y permitiendo mirar a la probabilidad de ocurrencia de un evento como la estabilización de las tendencias de las frecuencias relativas, realizados un número grande de ensayos, y apoyados de los registros gráfico y numérico. La importancia de esta propuesta está en que el estudiante se ve impulsado a estudiar las diferentes representaciones gráficas y numéricas, compararlas y asociarlas entre sí para dar una explicación a lo que constata, obtener conclusiones, determinar sus resultados y tomar decisiones, permitiendo al estudiante construir su propio aprendizaje. Además, el estudiante puede observar y descartar algunas ideas erróneas (heurísticas) acerca de lo que sucede en un evento al realizar un número reducido de ensayos contra lo que sucede al realizarlo un número suficientemente grande, para determinar la probabilidad en eventos de naturaleza empírica.

Referencias bibliográficas

Álvarez, J. (2011). *Simulaciones y problemas probabilísticos con GeoGebra*. Recuperado el 20 de mayo de 2015 de http://thales.cica.es/sites/thales.cica.es.geogebra/files/II_Jornadas_GeoGebra/material/talleres/probabilidad/taller_probabilidad.pdf

Barragués, J. y Guisasola, J. (2007). Simulación por ordenador de experimentos aleatorios en la enseñanza de la probabilidad. *Sigma No. 31*, 207-223.

Barragués, J. y Guisasola, J. (2009). Una propuesta para la enseñanza de la probabilidad en la universidad basada en la investigación didáctica. *Educación Matemática, 21 (3),* 127-162.

Batanero, C. (2005). Significados de la probabilidad en la educación secundaria. *Revista Latinoamericana de Investigación en Matemática Educativa, 8 (3)*, 247-263

Batanero, C. (2006). Razonamiento probabilístico en la vida cotidiana: un desafío educativo. En P. Flores y J. Lupiáñez (Eds.). *Investigación en el aula de matemáticas. Estadística y azar*. Granada: Sociedad de Educación Matemática Thales.

Franquet, J. (2008). *El estudio operativo de la psicología. Una aproximación matemática*. España: UNED-Tortosa.

Guisasola, J. y Barragués, J. (2002). Heurísticas y sesgos de los estudiantes de primer ciclo de universidad en la resolución de problemas de probabilidad. *Enseñanza de la ciencias, 20 (2)*, 285-302

Kolmogorov, A. (1979). La Teoría de Probabilidades. En A. Aleksandrov, A. Kolmogorov y M. Laurentiev (Eds.). *La matemática: su contenido, métodos y significado* (pp. 269-309), tomo II. España: Alianza Universidad.

Salinero, P. (sf). *Historia de la Probabilidad*. Universidad Autónoma de Madrid. (Asignatura: Historia de la Matemática)

Serradó, A., Cardeñoso, J. y Azcárate, P. (2005). Obstáculos en el aprendizaje del conocimiento probabilístico: su incidencia desde los libros de texto. *Statistics Education Research Journal, 4(2)*, 59-81

Serrano, L., Batanero, C. y Ortíz, J. (1996). Interpretación de enunciados de probabilidad en términos frecuenciales por alumnos de bachillerato. *Suma, 22*, 43-50.

Serrano, L., Batanero, C., Ortíz, J. y Cañizares, M. (1998). Heurísticas y sesgos en el razonamiento probabilístico de los estudiantes de secundaria. *Educación Matemática, 10 (1)*, 7-25

Sosa, L. (2006). *Desarrollo Conceptual de la Probabilidad*. Tesis de Maestría no publicada, Universidad Autónoma de Morelos.

Autores

Yessica del Rosario May Pech; UADY. México; mpyes@hotmail.com
Zuleyma Sarahí Pérez Moguel; UADY. México; zuleyma.2703@gmail.com