

CONCEPCIONES DE LOS ALUMNOS ACERCA DE LA PROBABILIDAD

María Inés Rodríguez, Héctor L. Agnelli

Universidad Nacional de Río Cuarto

mrodriguez@exa.unrc.edu.ar

Campo de investigación: Pensamiento relacionado con
probabilidad y estadística

Argentina

Nivel: Superior

Resumen. Con la finalidad de analizar las concepciones y formas de razonar de los alumnos acerca de la probabilidad, se realizó una encuesta en un curso introductorio de Estadística destinado a estudiantes de ciencia biológicas y del profesorado de matemática. Las preguntas se orientaron a indagar su comportamiento ante la concepción de la probabilidad desde un punto de vista clásico, frecuencial o subjetivo. Estos conocimientos previos con que el alumno llega al curso de estadística pueden convertirse en obstáculos para la enseñanza de la probabilidad y dificultar, luego, el aprendizaje de los conceptos propios de la inferencia estadística. En el desarrollo del presente trabajo se describen los enfoques probabilísticos denominados clásico, frecuencial y subjetivo, así como sus implicaciones para la enseñanza de la estadística. Se describe la encuesta y se muestran los resultados obtenidos, los que ponen en evidencia la necesidad de orientar esfuerzos para clarificar las distintas interpretaciones de la probabilidad, más que en los aspectos algorítmicos del tema.

Palabras clave: probabilidad, interpretaciones: clásica, frecuencial, subjetiva

Introducción

Como señala Shaughnessy (2002), “nuestros estudiantes no son páginas en blanco, esperando que la teoría normativa de la probabilidad descienda de nuestra boca. Los estudiantes ya tienen sus propias heurísticas, sesgos y creencias acerca de la probabilidad y estadística”. Por lo tanto a la vez que esta situación plantea un problema para la enseñanza y el aprendizaje de los conceptos de probabilidad y estadística, también nos señala la importancia de conocer cuáles pueden ser algunas de estas concepciones previas.

Para aportar al conocimiento de esta situación, en particular sobre las concepciones y formas de razonamiento acerca de la probabilidad con que llegan los alumnos a un curso inicial de estadística, se realizó una encuesta a estudiantes de ciencias biológicas y con fines comparativos se administró la misma encuesta a estudiantes del profesorado de matemáticas antes de tomar su primer curso de probabilidades. Las preguntas formuladas están orientadas a indagar su comportamiento ante problemas cuya resolución implica concebir la probabilidad desde un punto de vista clásico, frecuencial o subjetivo. Una encuesta similar a la presentada fue utilizada por Albert (2003).

Antecedentes

En los cursos introductorios, la enseñanza de la estadística puede desarrollarse sobre la base de tres concepciones fundamentales de la probabilidad: clásica, frecuencial y Bayesiana. Batanero y Díaz (2007), sugieren no limitarse a una sola ya que ellas están ligadas dialécticamente. Las diferencias fundamentales entre estas tres perspectivas radica en la manera de asignar probabilidades y en la interpretación de los valores de probabilidad obtenidos después de realizar cálculos. Es importante distinguir entre asignar y calcular probabilidades. Para el cálculo de probabilidades se aplican las propiedades derivadas de la construcción axiomática de la probabilidad, pero estos cálculos dependen de asignaciones iniciales de probabilidad o de la adopción de ciertos modelos distribucionales, o de ambos. Así, por ejemplo, cuando una característica observable en una situación experimental se modela como una variable aleatoria que sigue una distribución binomial, la probabilidad de un evento se calcula usando la función de densidad correspondiente, pero el valor de la probabilidad del éxito (p) debe ser asignado inicialmente de alguna manera.

La interpretación clásica que se atribuye a Laplace, aunque ya aparece en los trabajos de Pascal, Bernoulli, Huygens, y Leibniz, asigna igual probabilidad a todos los resultados posibles. De esta forma, la probabilidad de un evento es el cociente entre el número de maneras en que puede ocurrir y el número de resultados posibles. La aplicación de esta interpretación queda reducida a espacios muestrales finitos y con resultados equiprobables.

La interpretación frecuencial asume que el experimento es repetido muchas veces bajo condiciones similares y la probabilidad de un evento es estimada por la frecuencia relativa de su aparición en el conjunto de resultados experimentales. Así expresada esta interpretación tiene cierta similitud con la interpretación clásica, ya que otorga igual peso a cada miembro de un conjunto de eventos y, simplemente, calcula la proporción de los favorables en el total de los resultados producidos. La diferencia esencial es que el enfoque frecuencial tiene en cuenta los resultados producidos por el experimento, mientras que el enfoque clásico tiene en cuenta los resultados posibles de un experimento. Para poder aplicar la interpretación frecuencial se debe asumir que la situación aleatoria se repite bajo condiciones similares, es decir no abarca los sucesos aleatorios de ocurrencia única en el tiempo, Batanero (2005). Si bien esta interpretación extiende la interpretación clásica a situaciones en las que los resultados no son equiprobables, su

definición genera problemas. Ejemplo, si una moneda se tira una vez, la frecuencia relativa de cara es 0 ó 1, cualquiera que sea su sesgo. Si se tira dos veces, las frecuencias relativas podrían ser 0, $\frac{1}{2}$, 1. Es decir, la “separación” entre los valores de las frecuencias relativas estará dada por $1/n$. Surge así, el problema de cuál es el número n adecuado para la asignación de probabilidades. Exponentes frequentistas han sido Venn (1876) y von Mises (1957).

La interpretación subjetiva considera la probabilidad como una medida numérica de la creencia que tiene una persona acerca de la ocurrencia de un evento. La persona asigna una probabilidad de manera que refleje su creencia acerca de la verdad o falsedad del evento. Este es el enfoque más general, que se aplica a eventos que pueden no ser equiprobables y a los que no pueden repetirse bajo las mismas condiciones. De esta manera, la probabilidad está referida a la incertidumbre y no únicamente a la repetición de experimentos. La incertidumbre significa, en muchas situaciones, conocimiento incompleto y no ausencia absoluta de conocimiento. Esta interpretación es personal, pues diferentes personas pueden tener distintas opiniones y en consecuencia, asignar al mismo evento diferentes probabilidades. También la probabilidad depende de la información que posea la persona al momento de emitir su juicio; si dispone de nueva información, su asignación de probabilidad puede variar. La formalización de la probabilidad subjetiva, fue realizada por De Finetti (1972), para quien la incertidumbre debe ser expresada por una distribución de probabilidad, siendo ésta la característica esencial de la metodología estadística Bayesiana.

La probabilidad y la inferencia estadística

Moore (1997), en un trabajo en el que analiza la conveniencia de enseñar inferencia estadística clásica o Bayesiana, debate en el que obviamente subyace la cuestión de si asignar más o menos importancia a la interpretación de probabilidad frecuencial o a la subjetiva, expresa, “Como es usual, en los primeros cursos la barrera principal es la probabilidad. Tanto la estadística clásica como la Bayesiana están basadas en el concepto de probabilidad y en ambos casos este concepto puede ser presentado con mayor o menor grado de formalidad. Creo que en cualquier nivel de formalismo el razonamiento Bayesiano requiere una más compleja noción de probabilidad y una maquinaria probabilística mayor que la necesaria para la inferencia estándar.” Para sustentar esta opinión argumenta, “Las dificultades acerca de las ideas probabilistas han sido documentadas

tanto por los psicólogos, quienes investigan cómo las personas piensan las chances, como por los investigadores en educación, quienes estudian los efectos de nuestra intervención (enseñanza) sobre el pensamiento de los estudiantes. El mejor trabajo psicológico conocido es el de Tversky y Kahneman (1982), quienes muestran que: los juicios intuitivos acerca de las probabilidades marginales, conjuntas y condicionales no son probablemente coherentes, esto es, pueden no satisfacer las restricciones de la teoría de probabilidad.” También Falk y Konold (1992), señalaron que la intuición de las personas al aprender de su experiencia y revisar sus creencias, es consistente con el análisis Bayesiano. Por su parte, Lécoutre (2006), cuestiona seriamente el énfasis que desde la educación matemática se da a la interpretación frecuencial ignorando la concepción subjetivista, abogando por no mostrar a las mismas como opuestas, sino como complementarias.

Metodología

Para recopilar datos sobre la forma de razonar de los estudiantes, se formularon nueve preguntas, tres para cada tipo de interpretación de la probabilidad, que fueron contestadas en el aula el primer día de clase. Se solicitó que las respuestas fuesen razonadas, expresadas numéricamente y además justificadas brevemente, estipulándose una duración de 45 minutos para concretar las mismas.

Si bien en los planes de estudio de la escuela media está previsto el desarrollo de temas vinculados con la probabilidad y la estadística, estos forman parte de la asignatura matemática y no necesariamente, por diversas razones, son cubiertos en todas las escuelas. Esta situación es expuesta por los mismos profesores cuando realizamos actividades académicas destinadas a darle continuidad a su formación. En este grupo, menos del 8% de los alumnos confirmó haber recibido algunas nociones de probabilidad en el secundario. Por otra parte, indagando en los textos de estudio más utilizados en este nivel educativo, como por ejemplo, Kaczor (2002) y Camuyrano (2003), hemos comprobado que los mismos desarrollan el enfoque probabilístico clásico y frecuencial, pero no el subjetivo.

Cuestionario

Interpretación clásica.

Pregunta 1. Supongamos que Usted elige de la caja de la Figura 1, una bolilla al azar ¿Cuál es la probabilidad de sacar una bolilla numerada con 5 o mayor que 5?

Pregunta 2. Suponga que Usted elige de la caja de la Figura 2 un objeto al azar. ¿Cuál es la probabilidad de elegir un rectángulo?

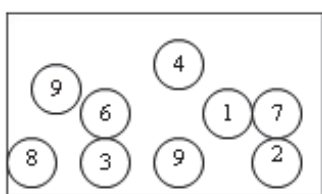


Figura 1

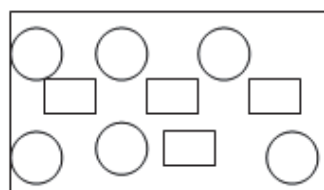


Figura 2

Pregunta 3. Supongamos que Usted tiene un juego como el que muestra la Figura 3. La aguja se impulsa y gira libremente alrededor del punto O hasta que se detiene. ¿Cuál es la probabilidad de que al detenerse lo haga dentro del área sombreada?

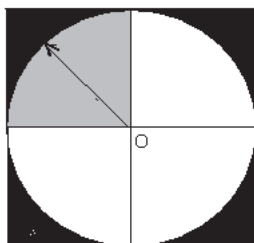


Figura 3

Estas tres preguntas pertenecen al tipo clásico. La diferencia entre la primera y la segunda está dada por la forma de los objetos. Por lo tanto esta situación será útil durante el desarrollo del curso para motivar la necesidad de utilizar dispositivos, ajenos a los objetos y al individuo, para generar extracciones aleatorias. La pregunta número tres, si bien se puede responder teniendo en cuenta que el sector elegido es uno de los cuatro sectores posibles en los que puede detenerse la aguja, la situación planteada también será útil para mostrar, más adelante, que para calcular algunas probabilidades es necesario medir en lugar de contar.

Interpretación frecuencial

Pregunta 4.- Suponga que Usted tira una moneda 20 veces y sale 19 veces cara y una cruz. Si Usted tira la misma moneda una vez más ¿Cuál es la probabilidad de que salga cara?

Pregunta 5.- Si esta mañana se produce en el hospital local un nacimiento. ¿Cuál es la probabilidad de que sea varón?

Pregunta 6.- Suponga que elige al azar un estudiante de primer año de su carrera. ¿Cuál es la probabilidad de que sea mujer?

La respuesta a la pregunta 4 podría deberse a la desatención total a la información suministrada, o bien teniéndola en cuenta. Al ignorar la información, la respuesta, asumiendo que la moneda es equilibrada (dato no suministrado), la probabilidad de cara sería $\frac{1}{2}$. Y si se asume la información, se podría asignar una probabilidad basada en la frecuencia relativa $\frac{19}{20} = 0.95$. Este problema es interesante para discutir los tamaños de muestras necesarios para estimar probabilidades a partir de frecuencias relativas, ya que si el número de repeticiones no es elevado, y no se presta atención a las rachas y fluctuaciones del proceso estocástico, se puede reforzar un sesgo, ya caracterizado en la literatura de investigación, conocido como creencia en la ley de los pequeños números (Tversky y Kahneman, 1982, p. 23-31). En la pregunta 5, a diferencia de la anterior, se pide la asignación de probabilidad sin proporcionar información numérica en el enunciado. Una posible respuesta puede derivar de la referencia a la composición poblacional de niños y niñas de la ciudad y dar una estimación de probabilidad basada en la frecuencia relativa; otra posible respuesta es asignar probabilidad a partir de que sólo existen dos resultados posibles, varón o mujer. Esta última situación, da lugar al así denominado por Lécoute (1992) sesgo de equiprobabilidad. Éste consiste en asignar la misma probabilidad a todos los resultados asociados a cualquier experimento aleatorio, incluso para aquéllos en que no es aplicable el principio de indiferencia o donde no hay una simetría física. Aunque para la pregunta 6 tampoco se proporcionan datos, sí se conoce que las mujeres representan el 80% aproximadamente de los estudiantes de las carreras incluidas en esta encuesta.

Interpretación Subjetiva.

Pregunta 7.- ¿Cuál es la probabilidad de que Boca gane su próximo partido contra Gimnasia y Esgrima de La Plata el próximo fin de semana?

Pregunta 8.- ¿Cuál es la probabilidad de que Ud se gradúe en la UNRC en 4 años o menos?

Pregunta 9.- ¿Cuál es la probabilidad de que Usted se case antes de los 25 años?

Para la pregunta 7, se podría esperar que, si se dispone de información acerca de cómo se venían desempeñado los equipos, sería razonable asignar probabilidades teniendo en cuenta esa información. Dadas las características del curso y del evento, se esperaría que los alumnos contaran con poca información y esto podría originarles dificultades en expresar la probabilidad a partir de sus creencias. La pregunta 8 se podría responder mediante una interpretación frecuencial basada en la proporción de estudiantes que se gradúa en los tiempos establecidos y asumir que el tiempo que requiere la graduación no está influenciado por las condiciones personales. Alternativamente, se puede responder la pregunta desde la situación personal de cada individuo. La pregunta 9 es la que mejor configura la característica de que el evento es no repetible en condiciones similares en el tiempo y su concreción está determinada por el comportamiento personal.

Resultados

Interpretación clásica. Las respuestas a las tres primeras preguntas relativas a la interpretación clásica la respondieron correctamente el 87% de los estudiantes. De cualquier manera, de éstos el 28% expresaron la probabilidad como un número entre 0 y 1; el otro 72% expresó la probabilidad como un porcentaje. Todos los que contestaron correctamente llegaron a una expresión numérica considerando la cantidad de resultados posibles finita (cuatro cuadrantes en el caso del círculo) y estos resultados equiprobables. Por otra parte, es notable que varios alumnos distinguieron entre sacar la bolilla numerada con el 5 (no existente en la caja), evento al que asignaron probabilidad nula, y las bolillas numeradas con valores mayores que 5, evento al cual asignaron probabilidad 50%.

Interpretación frecuencial. Pregunta 4: Un 43% de los alumnos optó por asignar probabilidad 50%. El resto hizo una asignación de probabilidad basada en la información disponible pero, de ellos, sólo el 14% calculó correctamente la frecuencia relativa. Pregunta 5: La mayor parte de los alumnos dio como respuesta 50%. Pregunta 6: La mayoría de las respuestas tuvieron en cuenta la interpretación frecuencial y la asignación de probabilidades se expresó por valores tales como 80%, 90%, 99%, como atribuibles a probabilidad “alta”. Una cantidad importante de respuestas asignaron probabilidad $\frac{1}{2}$ al dicotomizar la situación en: “mujer” o “varón”. Entre éstas, parece oportuno citar la respuesta: “si bien la mayoría son mujeres la probabilidad es 50%”, ya que si bien el estudiante posee información él no utiliza para asignar probabilidades.

Interpretación Subjetiva. Pregunta 7: El 70% de los estudiantes contestaron a la pregunta asignando probabilidad $\frac{1}{2}$ ó $\frac{1}{3}$. Para el primer caso consideraron los resultados posibles ganar o perder y, para el segundo, ganar, empatar o perder. Pero en ambas situaciones consideraron a los resultados equiprobables. El 20% dio una respuesta que refleja sus creencias y el resto no contestó. Pregunta 8: La mayoría (67%) contestó sobre la base de su opinión y asignó probabilidades “altas” o “bajas”- pero no valores numéricos- al evento y también hubo casos con respuesta 100% o 0%. En algunas respuestas con probabilidades altas se justificaban esos valores a partir de las expectativas de concretar la graduación. También el 12% de los alumnos dicotomizaron la situación en: “me gradúo” o “no me gradúo” considerando nuevamente la equiprobabilidad, respondieron $\frac{1}{2}$. Pregunta 9: Las respuestas a esta pregunta presentaron las diferencias más notables entre los alumnos de matemáticas y los de biología. En el primer grupo, las respuestas fueron “objetivas”, en el sentido de emplear algún mecanismo de cálculo para generar la respuesta, e iguales a $\frac{1}{2}$, respondiendo al esquema “me caso” o “no me caso”, una respuesta llamativa fue $\frac{2}{25}$. En el grupo de biología, las respuestas fueron marcadamente subjetivas. Hubo respuestas atribuibles a la interpretación frecuencial: “la probabilidad es baja porque la mayor parte de la gente se casa después de los 25 años”.

Conclusiones

En general los estudiantes no distinguen las tres interpretaciones de probabilidad. Si bien con relación a las preguntas vinculadas con la interpretación clásica las respuestas fueron correctas, es notorio que se asume de manera automática la equiprobabilidad de los resultados y esto conduce a asignaciones de probabilidad incorrectas cuando se pasa a situaciones donde no está presente la simetría.

Si bien no se presentan problemas en la asignación de probabilidades bajo la interpretación clásica, parece conveniente utilizar más tiempo, en reflexionar acerca de que no necesariamente los resultados de los experimentos dicotómicos son equiprobables; que la probabilidad puede ser expresada como un número; y que implementar extracciones equiprobables requiere en algunos casos usar mecanismos aleatorios ajenos al sujeto, como por ejemplo el uso de tablas de números aleatorios o rutinas aleatorias en un computador.

Es marcado el déficit en la asignación de probabilidades desde el punto de vista frecuencial. Es conveniente, por lo tanto, enfrentar al alumno a diversas situaciones en las que tenga que estimar probabilidades a partir del cálculo de frecuencias relativas. Esta interpretación es la más común en los cursos de inferencia estadística, ya que constituye la herramienta necesaria para interpretar las distribuciones muestrales, intervalos de confianza y valores de p en las pruebas de significación.

El creciente desarrollo teórico y el auge en las aplicaciones de la metodología Bayesiana, basada en la interpretación subjetiva, hacen necesario prepararse para enseñar la misma en los cursos introductorios. Dado que los resultados pusieron en evidencia las dificultades que tienen los alumnos en traducir sus apreciaciones a valores numéricos de probabilidad, sería oportuno ejercitar la asignación de probabilidades subjetivas mediante experimentos de calibración como los descritos en el libro de Berry (1996).

Referencias bibliográficas

Albert, J. (2003). College Students' Conceptions of probability. *The American Statistician* 57, (1), 37- 45.

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

Batanero, C. y Díaz, C. (2007). Probabilidad, grado de creencia y proceso de aprendizaje. *XIII Jornadas Nacionales de Enseñanza y Aprendizaje de las Matemáticas*. Granada, Federación Española de Profesores de Enseñanza de las Matemáticas.

Berry, D.A. (1996). *Statistics: A Bayesian Perspective*. Belmont: Duxbury Press.

Camuyrano, M. B., Net, G. y Aragón, M. (2003). *Modelos matemáticos para interpretar la realidad*. Buenos Aires: Estrada.

De Finetti, B. (1972). *Probability, Induction and Statistics*. Chichester: Wiley.

Falk, R. y Konold, C. (1992). The Psychology of Learning Probability. En F. Gordon y S. Gordon (Eds) *Statistics for the Twenty-First Century* 26, 151-164.

Kaczor, P.; Schaposchnik, E. F.; Cicala, R. y Diaz, B. (2002). *Matemática I*. Buenos Aires: Santillana

Lêcoutre, M. (1992). Cognitive models and problem spaces in “purely random” situations. *Educational Studies in Mathematics*, 23, 557-568.

Lêcoutre, B. (2006). People intuitions about randomness and probability. *Statistical Education Research Journal* 5(1), 20-35.

Moore, D. (1997). Bayes for Beginners? Some reasons to hesitate. *TAS* 51 (3), 254- 261

Shaughnessy, M. (2002). *Investigación en Probabilidad y Estadística: Reflexiones y Orientaciones*. Departamento de Matemática Educativa. Cinvestav. México.

Tversky, A.; Kahneman, D.; Slovic, P. (eds.) (1982). *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Press.

Venn, J. (1962). *The logic of the chance*. New York: Macmillan

Von Mises, R. (1957). *Probability, Statistics and Truth*. New York: Macmillan.