Experimentos aleatorios y sus representaciones

Hugo Alvarado y Lidia Retamal [1]

Resumen

La enseñanza de la probabilidad y estadística en la educación secundaria con apoyo informáticon, proporciona una mayor potencia en el cálculo de probabilidades, las representaciones gráficas y la introducción de ideas de inferencia como las distribuciones en el muestreo. El diseño del Taller considera la generación de resultados aleatorios y situaciones de aprendizaje con ficheros de datos. Las actividades proporcionan aspectos interpretativos y conceptuales sobre la organización de datos, la distribución de frecuencias, parámetro y estadístico y las conexiones con otros conceptos tales como las variables aleatorias y su distribución. En consecuencia, se evalúa la correspondencia entre experimentos aleatorios sencillos y su aplicabilidad a situaciones reales mediante los dispositivos de la planilla Excel y applet del modelo binomial.

Palabras clave: Experimentos aleatorios, dispositivos didácticos, probabilidad binomial.

Introducción

Los procesos de instrucción en estadística están en una etapa de cambio en los métodos de enseñanza que se utilizan, teniendo en cuenta que los alumnos usan comúnmente el lenguaje de la tecnología. Esto conduce al profesor de matemática a modificar su actuar en el aula, teniendo presente que algunas veces estos recursos informáticos aumentan los problemas educativos en vez de resolverlos. Surgen preguntas básicas en la exploración de experimentos aleatorios acerca de qué dispositivos didácticos

[1] Universidad Católica de la Santísima Concepción- Chile

(dados, fichas, Excel, software, plataforma virtual, Internet, libros de textos, etc.) son los adecuados para la enseñanza de la estadística en los distintos niveles educativos y qué efecto tiene la informática a favor de un aprendizaje significativo. DelMas, Garfield y Chance (1999) señalan que la tecnología por sí sola no es suficiente para la comprensión de propiedades de la distribución en el muestreo, sino que las actividades y la enseñanza de tipo constructivista tienen un rol importante. Moore (1997) recomienda una renovación tanto en los métodos como en el contenido. Es conveniente que los estudiantes tengan nociones de probabilidad si pretendemos que entiendan el significado de la inferencia estadística. El estudio de las frecuencias relativas en un experimento aleatorio proporciona las bases sobre las cuales se elabora el concepto de probabilidad. Por lo general se presenta la probabilidad en el contexto de juegos de azar con resultados equiprobables. Muchas situaciones probabilísticas conllevan una serie de experimentos aleatorios donde la simulación permite obtener una estimación de la solución. Batanero (2003) señala que en los estudiantes de secundaria la simulación contribuye a conocer las diferencias entre la probabilidad experimental y la teórica.

En este trabajo se pretende plantear actividades que permitan el tratamiento estadístico de un experimento aleatorio desde varios acercamientos de representaciones progresivas para familiarizar a los participantes con la noción de modelo de probabilidad. La primera sesión considera actividades de experimentos aleatorios sencillos con dispositivos manipulables y con uso de Excel. Se desarrolla el caso del experimento de simulación de lanzamientos de dos dados de seis caras. La segunda sesión contempla reforzar algunas actividades planteadas, contextualizándola a situaciones reales con argumentos de visualización y cálculo tanto de la planilla Excel como el miniprograma applet, escrito en Java, del modelo de probabilidad binomial. Finalmente, se propone analizar una base de datos sobre consumo de luz y se implementa un cuestionario en línea, con ítems conceptuales de selección múltiple y de cálculo de probabilidad, a desarrollar por los participantes con apoyo de los recursos tecnológicos.

- **Tipos de representaciones**

Las actividades que se plantean en el Taller se sustenta en las configuraciones didácticas del enfoque onto-semiótico de la cognición e instrucción matemática (Godino, Batanero y Font, 2007; Godino, Contreras y Font, 2006). En un sentido amplio, la noción cognitiva de competencia entendida como "saber hacer" implica los diversos elementos del significado para el conocimiento matemático, que se describen en: problemas, definiciones, propiedades, algoritmos, lenguaje y argumentación. Estas entidades primarias están relacionadas entre sí formando *configuraciones* que se definen como redes de objetos intervinientes y emergentes de los sistemas de prácticas. En este trabajo, las actividades propuestas conducen a la apropiación progresiva mediante tres tipos de representaciones para argumentar los experimentos aleatorios. Se entiende por recursos didácticos a todos aquellos instrumentos que facilitan la comunicación hacia la exploración, aplicación y argumentación de conceptos importantes en el aula de estadística.

- *Representación manipulativa* el alumno trabaja con dispositivos manipulativos (dados, fichas,...), papel-lápiz o calculadora, sin utilizar notación o cálculo algebraico. Aparecen la noción de experimento aleatorio y estadístico, y se argumenta mediante ejemplos y contraejemplos. Los procedimientos son empíricos y gráficos y el lenguaje se reduce a expresiones verbales y gráficas en papel y lápiz.
- La *representación algebraica* se caracteriza por el lenguaje simbólico y la demostración deductiva, así como el recurso a elementos de álgebra. Los procedimientos serían analíticos.
- La *representación computacional* amplía la variedad de gráficas dinámicas. Además, del lenguaje icónico, incorpora como procedimiento la simulación y permite trabajar con las variables estadísticas y aleatorias simultáneamente. No posibilita el lenguaje algebraico ni la demostración deductiva. El argumento preferible es inductivo, estudio de ejemplos, contraejemplos y la generalización. Aparecen algunos conceptos como el experimento de probabilidad binomial.

- **Recursos en internet**

Actualmente existe bastante información disponible en internet sobre la enseñanza de la probabilidad y estadística; sin embargo, se carece de metodologías para evaluar instrumentos de aprendizaje. El papel del profesor de matemática es fundamental en la iniciación de experimentos aleatorios mediante la informática educativa en estadística, que por su carácter metodológico e instrumental la estadística contribuye en el diseño de experimentos. Una descripción de los recursos que ofrece la web es la siguiente:

a) *Software libre*: ver http://statpages.org/javasta2.html. Los estudiantes universitarios en la UCSC han usado el programa @risk, de licencia limitada, a diferencia del programa estadístico R que no tiene costo.

b) *Java Applet*: Este recurso se ha diseñado e implementado en nuestro curso mediante proyectos de docencia e investigación en las distribuciones de probabilidades discretas con el fin de ilustrar, simular y explorar conceptos estocásticos en un ambiente dinámico y gráfico.
Ver http://fyl.cl/fondecyt/Binomial_aprox/Inicio.html. En el caso continuo los estudiantes están experimentando con Geogebra las distribuciones de probabilidad normal, gamma y beta.
Ver http://fyl.cl/fondecyt/Gamma_aprox/Inicio.html.

c) *Materiales educativos en línea*: se pueden encontrar libros en línea de interés que incorporan representaciones interactivas y que se puede manipular para explorar los conceptos estadísticos, por ejemplo The World of Visual Statistics; Introductory Statistics: Concepts, Models and Applications; HyperStat; The Little Handbook of Statistical Practice; Electronic Statistics Textbook.

d) *Revistas electrónicas y boletines de información*: De gran interés son los resultados de investigación, siendo las de mayor divulgación: Journal of Statistics Education (JSE) publicada por la American Statistical Association (ASA); Statistics Education Research Journal (SERJ) principal revista electrónica de responsabilidad de la Internacional

Association for Statistics Education (IASE); Teaching Statistics; Statistics Teacher Network boletín de la Nacional Council Teachers of Mathematics (NCTM).

De la experimentación a la generalización

El Taller considera la participación activa de los estudiantes comenzando con experimentos manipulables sencillos de juego de dados, cuyos resultados de los experimentos se tabulan y se representan gráficamente en Excel y luego se formaliza con el cálculo de probabilidades de forma algebraica y con un applet, disponible en la plataforma virtual de aprendizaje EV@ que ofrece la Dirección de Docencia de la UCSC. Las actividades propuestas en las dos sesiones serán desarrolladas en el curso "Coloquio matemática" disponible en la plataforma EV@: http://uvirtual.ucsc.cl/

Figura 1. Diseño del Taller en la Plataforma virtual EV@

Las sesiones del Taller está planificada sobre la base de la experiencia de experimentos de enseñanza de la estadística para estudiantes universitarios en la Universidad Católica de la Santísima Concepción (UCSC), mediante la utilización del programa @risk (Retamal, Alvarado y Rebolledo, 2007), la planilla Excel en actividades de simulación, análisis de las dificultades de comprensión de distribuciones de probabilidad aproximadas (Alvarado y Batanero, 2007) y el análisis de libros de texto de estadística (Alvarado, 2004).

- **Selección de conceptos a tratar**

A continuación se mencionan los conceptos básicos para el estudio de las nociones de probabilidad que se tratarán en este taller. Contextos sobre predecir el consumo de luz y contestar un cuestionario sorpresa pueden servir para apreciar las características de los fenómenos aleatorios.

1. *Variable estadística*. En un estudio estadístico recogemos datos de los individuos de una población. Si las características o variables que queremos estudiar son cualitativas, los datos obtenidos clasifican a la población en categorías y se llaman categóricas o cualitativas. Así por ejemplo, el nivel educacional (básica, media, superior) y lugar de vacaciones donde le gustaría tener "más espacio" (en el avión, en la habitación del hotel, otro) son variables cualitativas. Una característica la llamamos variable estadística porque los valores que toma pueden variar en los distintos individuos de la población.

Cuando, al estudiar alguna característica, obtenemos datos numéricos, se trata de una *variable numérica*. Por ejemplo, son variables numéricas número de hijos de las familias chilenas, altura de los alumnos de una clase. En el primer caso se trata de una variable discreta porque la variable no puede tomar ningún valor intermedio, mientras que en el segundo caso se trataría de una variable continua. Cabe señalar, que en general se trabaja con resultados discretizados, ya que los resultados que se recogen en cualquier proceso de medición siempre son aproximados, y muchas veces es conveniente redondear para simplificar los cálculos. Hay que considerar que cuando una variable numérica puede tomar muchos valores, por ejemplo, tiempo que se demora un alumno el ir de la casa al colegio, resulta apropiado agrupar los datos y representarlos mediante un histograma.

2. *Parámetros estadísticos*. Un solo parámetro estadístico da poca información sobre la distribución y es adecuado calcular el

máximo posible, por ejemplo la media aritmética, mediana, rango, cuartiles y desviación estándar.

Las tablas de frecuencias y sus correspondientes gráficos son instrumentos útiles para visualizar cómo se distribuyen los datos organizados. Nuestro interés en este trabajo es tratar las variables numéricas, cuya descripción puede resumirse en algunos parámetros. A objeto de hacer comparaciones entre distintas poblaciones es conveniente disponer de parámetros teniendo presente que se pierde información al resumir lo datos. Los parámetros de interés son la media aritmética, el rango y la desviación estándar.

La *media* representa el centro de gravedad de un conjunto de datos. Si se representa los datos gráficamente, con su correspondiente frecuencia sobre una barra horizontal, la media correspondería al punto respecto del cual la barra mantendría el equilibrio. Sin embargo, se aporta más información cuando este parámetro va acompañado con la dispersión de los datos. El *rango* de la distribución es la longitud del intervalo donde se distribuyen los datos. Es fácil de calcular e interpretar, aunque es sensible a la existencia de valores extremos. La *desviación estándar* es la más usada y mide la dispersión de las observaciones respecto a la media aritmética.

3. *Experimento aleatorio*. Hay que diferenciar dos tipos de experimentos. En un experimento determinista, a partir de las mismas condiciones iniciales se obtienen se obtienen los mismos resultados, mientras que en un ***experimento aleatorio*** se conocen todos los posibles resultados pero no se puede determinar cuál va a ser el resultado que se va a obtener. Un *suceso* es cada uno de los posibles resultados de un experimento aleatorio; será elemental cuando no puede descomponerse en otros más simples, en caso contrario se denominan sucesos compuestos.

Es importante, antes de entrar en la iniciación al cálculo de probabilidades, describir de forma adecuada en qué consiste el experimento aleatorio y luego identificar todos los posibles resultados que constituyen el espacio muestral. Esto facilitará la simulación de fenómenos aleatorios realistas con diversos materiales. La idea de aleatoriedad en sus comienzos se entendía como aquello cuyas causas son desconocidas y el "azar" estaría personificado como las causa de los fenómenos aleatorios (Batanero, 2001).

4. *Simulación*. Llamamos simulación a la sustitución de un experimento aleatorio por otro equivalente con el cual se experimenta para obtener estimaciones de probabilidades de sucesos asociados al primer experimento. La estimación de probabilidad que se obtiene con el experimento simulado es tan válida como si tratase del experimento real.

5. *Variable aleatoria*. La variable estadística representa únicamente los *n* resultados de *n* realizaciones de un experimento aleatorio. Si el experimento se repite infinidad de resultados posibles da origen a la noción de variable aleatoria asociada al experimento. En general, cada resultado de un experimento puede ser asociado con un número que es especificado por una regla de asociación. Por ejemplo, el número de artículos defectuosos de una partida de 100 en una semana. Esta regla de asociación es llamada una variable aleatoria; una *variable* ya que valores numéricos diferentes son posibles y *aleatoria*, pues el valor observado depende de cuales de los resultados posibles del experimento ocurre. Una variable aleatoria es *discreta* si su recorrido es un conjunto discreto de números (finito o numerable), es decir, sus elementos pueden ser listados de manera que hay un primer elemento, un segundo elemento, ..,etc, en la lista (asumen sólo ciertos valores, comúnmente número enteros y resulta principalmente del conteo).

6. *Función de probabilidad*. Sea X una variable aleatoria discreta con recorrido R_X. Sea p una función que asigna a cada x $\in R_X$ un número $p(x) = P(X = x)$ llamado probabilidad de *x*. La función de probabilidad p de la v.a. X debe satisfacer las siguientes propiedades:

a) $p(x) > 0, \ \forall \ x \in R_X, \ \forall \ x \in \mathbb{R}$ b) $\sum p(x) = 1 \ \ x \in R_X$

La función p es expresada usualmente como los pares (x, p (x)) $\forall x \in R_X$, llamada *distribución de probabilidades*, que es una lista de todos los resultados posibles de un experimento y de la probabilidad relacionada con cada resultado.

7. *Experimento de probabilidad binomial*. Es el experimento integrado por ensayos repetidos, que posee las siguientes propiedades:

a) Hay n ensayos repetidos independientes e idénticos

b) Cada ensayo tiene dos resultados posibles (éxito o fracaso)

c) Prob(éxito) = p, Prob(fracaso) = 1- p = q

d) La variable aleatoria binomial X es el conteo del número de ensayos con éxito que ocurren, donde x puede adoptar cualquier valor entero de cero a n.

La propiedad de ensayos independientes significa que el resultado de un ensayo no afecta la probabilidad de éxito de cualquier otro ensayo en el experimento. En otras palabras, la probabilidad de "éxito" permanece constante a lo largo de todo el experimento. Las propiedades c) y d) están relacionada con la notación algebraica de cada ensayo de todo el experimento. Es importante que x y p estén asociadas con "éxito".

Lo fundamental para trabajar con cualquier experimento de probabilidad es su distribución de probabilidad, pues permite encontrar las probabilidades para cada valor posible de x. En particular, se define la *función de probabilidad binomial*:

Para un experimento binomial, sean p la probabilidad de un "éxito" y q la probabilidad de un "fracaso" en un solo ensayo; entonces P (X = x), la probabilidad de que haya exactamente x éxitos en n ensayos, es

$$P(X=x)=\binom{n}{x} \mathrm{p}^{x}\ q^{n-x} \quad \text{para} \quad x = 0,1,2,3,\ldots, n.$$

8. *Esperanza matemática*. Así como en estadística descriptiva se calcula la media de un conjunto de datos, también se puede determinar la media de una distribución de probabilidad. La

esperanza matemática de una variable aleatoria discreta es la media ponderada de todos los posibles resultados en los cuales los pesos son las probabilidades respectivas de tales resultados, y se refiere al centro de gravedad de una distribución de masa a través de una línea. Tanto el valor esperado como la varianza de una variable son llamados parámetros de una distribución y sirven para especificar completamente la distribución de probabilidad. En términos simbólicos el *valor esperado* de *X* se define por:

$$E(X) = \mu = \sum_{x \in R_X} x \cdot p(x)$$

La varianza de una distribución de probabilidad es el promedio de las desviaciones al cuadrado con respecto de la media, y se escribe como:

$$V(x) = E[(x-\mu)^2] = \sum_{x \in R_X} (x-\mu)^2 \cdot p(x) = E\left(X^2\right) - \mu^2$$

- **Secuencia de las sesiones**

<u>Sesión 1</u>.

La primera sesión comienza comunicando el objetivo del taller, los tipos de representaciones que conforman una experimento aleatorio y se comentan los conceptos de variable estadística, parámetro estadístico experimento aleatorio y simulación. Posteriormente se plantean algunos experimentos aleatorios analizando en caso del juego de dados mediante dispositivos manipulables y con uso de Excel.

El objetivo es apoyar a los profesores a comprender el significado de experimento aleatorio, observar las frecuencias relativas, observar cómo varían y cómo se estabilizan. En los experimentos de dados u otros hay que tener presente que intuitivamente se asignan probabilidades considerando la simetría. Por ello, es necesario en la experimentación que los participantes tiren dados y anoten los resultados obtenidos. La observación de una larga secuencia de resultados, obtenidos a partir de su propia experimentación. Las actividades planificadas son las siguientes:

Actividad 1.1. Se ha llevado a cabo una encuesta a un grupo de alumnos. Clasifica las siguientes características, según la escala de medida y tipo de variable: Peso, religión, número de hermanos, orden de nacimiento respecto a sus hermanos, si tiene o no licencia de conducir, tiempo que tarda en completar la encuesta, deporte preferido, número de año de escolaridad, tiempo para completar una tarea, número de sillas de una clase.

Actividad 1.2. Indicar qué tienen en común los siguientes experimentos. En algunos de ellos describe un suceso simple u otro compuesto.

a) Se lanza un dado y se observa el número que aparece en la cara superior.

b) Se lanza una moneda cuatro veces y se cuenta el número total de caras obtenidas.

c) Se lanza una moneda cuatro veces y se observa la sucesión de caras y sellos obtenidos.

d) Se fabrican artículos en una línea de producción y se cuenta el número de artículos defectuosos producidos en un periodo de 24 horas.

e) Se fabrica una ampolleta. Luego se prueba su duración en una casa y se anota el tiempo transcurrido (en horas) hasta que se quema.

f) En un lote de 10 artículos hay 3 defectuosos. Se elige un artículo después de otro (sin sustituir el artículo elegido) hasta que se obtiene el último artículo defectuoso. Se cuenta el número total de artículos sacados del lote.

g) Se fabrican artículos hasta producir 10 no defectuosos. Se cuenta el número total de artículos manufacturados.

h) De una urna que contiene sólo bolitas negras, se escoge una bola y se anota su color.

Actividad 1.3. Juego de dados. Observar las caras que muestran dos dados al ser lanzados al aire. Este experimento tiene 36 eventos elementales. Considere los eventos:

A_1 : “La suma de los dos números es divisible por 4”

$A_1 = \{(i, j) \in \Omega \, / \, i + j = 4n, n = 1,2,3\}$ = {(1,3), (2,2), (2,6), (3,1), (3,5), (4,4), (5,3), (6,2), (6,6)}

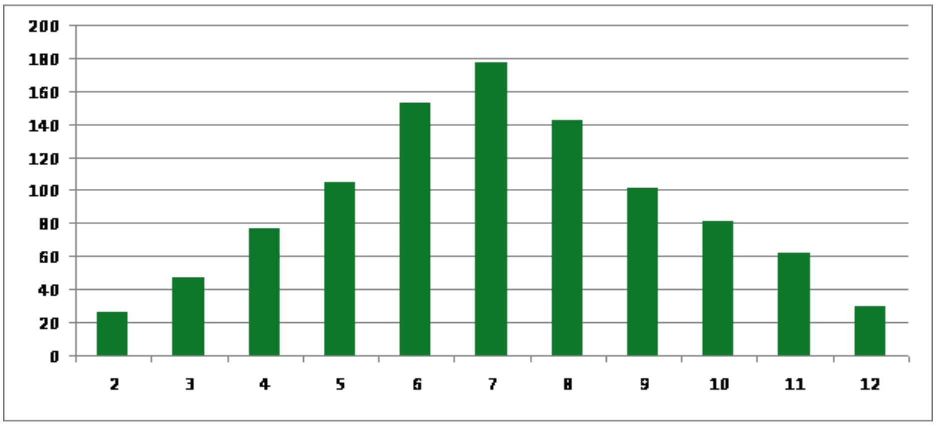
A_2 : “Los dos dados muestran el mismo número”

$A_2 = \{(i, j) \in \Omega \, / \, i = j \}$ = {(1,1), (2,2), (3,3), (4,4), (5,5), (6,6)}

La probabilidad de que la suma de los dos números de los dados sea divisible por 4 es P(A_1) = 1/4 y la probabilidad de que los dos dados muestran el mismo número es P(A_2) = 1/6

Actividad 1.4. Generar números aleatorios con la planilla Excel. Del experimento del lanzamiento de dos dados, simular la suma en 1000 lanzamientos. Las representaciones que se discutirán en el laboratorio serán de la forma siguiente:

D_1+D_2	2	3	4	5	6	7	8	9	10	11	12
Frecuencia	26	47	77	105	153	177	142	101	81	62	29



Actividad 1.5. Simulación con dados. Suponga que gana si obtiene un “6” al lanzar un dado legal. Simule las posibilidades de ganar en los siguientes casos de varios lanzamientos del dado *(n, m)* = (4,5), (10,10) y (30,10). Observe que *n* es el número de ensayos y *m* el número de réplicas.

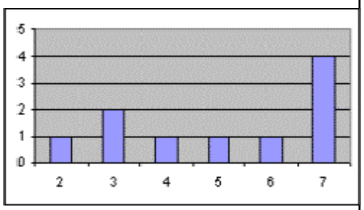
Tabla de experimento con dados para (30,10)	Tabla de distribución de frecuencias

Tabla de experimento con dados para (30,10)

X1	0	0	0	1	0	0	0	0	0	0	X16	0	0	0	0	1	0	0	1	0	0
X2	0	0	0	0	1	0	0	0	0	0	X17	0	0	0	0	1	0	1	0	0	1
X3	0	0	1	0	0	0	0	0	0	0	X18	0	0	0	0	0	0	0	0	0	0
X4	0	0	0	0	0	0	0	0	0	0	X19	0	1	0	1	0	0	0	0	1	0
X5	0	1	0	1	0	1	0	0	0	0	X20	1	0	0	0	0	0	1	0	0	0
X6	0	0	0	0	0	0	0	0	0	0	X21	0	1	0	0	0	0	1	0	0	0
X7	1	0	0	1	0	0	0	0	0	0	X22	0	0	0	0	0	0	1	0	0	0
X8	0	0	0	0	0	1	0	1	0	0	X23	1	0	0	0	0	1	0	0	0	0
X9	1	0	1	0	0	0	0	0	1	0	X24	0	0	0	0	1	0	0	0	0	0
X10	0	0	0	0	0	1	0	1	0	0	X25	0	0	0	0	0	1	0	0	0	0
X11	1	1	0	1	0	0	0	0	0	1	X26	0	0	0	0	0	0	0	0	0	0
X12	0	0	0	1	0	0	0	0	0	0	X27	0	0	0	0	1	0	0	0	1	0
X13	0	0	0	0	0	1	1	0	0	0	X28	0	0	1	0	0	0	0	0	0	0
X14	0	1	0	1	0	0	0	0	0	0	X29	1	0	0	0	0	0	1	0	1	0
X15	0	0	0	0	1	1	0	0	0	0	X30	1	0	0	0	1	0	0	0	0	0
											∑	7	5	3	7	7	7	6	3	4	2

Tabla de distribución de frecuencias

obt. el "6"	*Ni*	*Ni*	*fi*	*Fi*
2	1	1	0,1	0,1
3	2	3	0,2	0,3
4	1	4	0,1	0,4
5	1	5	0,1	0,5
6	1	6	0,1	0,6
7	4	10	0,4	1,0



El procedimiento común de análisis es comenzar introduciendo una notación simbólica para designar la variable y modelar el experimento por medio de una sucesión de variables aleatorias identificando el parámetro *p*:

Se define X_i: *Obtener un 6 al lanzar un dado. Considere* $X_i = 1$, *si se obtiene el número "6" y 0, en otro caso*

El parámetro p = P(E) = P(obtener el "6") = 1/6

Los estudiantes pueden simular los experimentos mediante una representación matricial, y llegar a definir la variable suma de los resultados. Construir una distribución de frecuencias con su representación gráfica. En la segunda sesión se modelará este experimento por medio de la probabilidad binomial; identificando los parámetros *n* y *p*, comparar y utilizar los conceptos de parámetro, estimador, media y varianza (teórica y experimental) como se muestra en la siguiente tabla:

(n , m) m réplicas de n ensayos	Parámetros y estimadores	Simulación
(4,5)	Media teórica	0,667
	Media experimental	1
(4,5)	Varianza teórica	0,55
	Varianza experimental	1
(10,10)	Media teórica	1,667
	Media experimental	1,8
(10,10)	Varianza teórica	1,389
	Varianza experimental	1,734
(30,10)	Media teórica	5
	Media experimental	5,1
(30,10)	Varianza teórica	4,167
	Varianza experimental	3,878

Se proponen algunos ejercicios a desarrollar al final del taller.

Ejercicio 1.1. Lanzamientos de dos dados. Imagina que estás jugando dados con un amigo. Tu compañero indica que hay tres posibilidades diferentes al lanzar dos dados:

I) que los números sean pares;

II) que los números sean impares;

III) haya un par y un impar.

a) El afirma que los tres casos son igual de probables. ¿Tú que opinas?

b) Otro compañero sugiere que realices el siguiente experimento para resolver la discusión. Trata de adivinar cuantas veces, aproximadamente, saldrá el 3 y cuantas el 5 si lanzas un dado 20 veces. Escribe este número en la columna "número esperado de veces"

Resultado	Recuento	Frec. absoluta	Frec. relativa	Número esperado de veces
a)				
b)				
c)				
Total		20	1	20

c) Lanza el dado 20 veces y anota los resultados en la tabla.

d) A continuación se mostrará en la pizarra los resultados de toda la clase. Compara estos resultados con los vuestros y con la estimación realizada. ¿Cuál de los sucesos es más probable?

Ejercicio 1.2. Alonso y Sofía han inventado un juego de dados con la siguiente regla: Lanzan dos dados sucesivamente y calculan la diferencia de puntos entre el mayor y el menor. Si resulta una diferencia de 0,1 ó 2 entonces Alonso gana una ficha. Si resultan 3,4 ó 5 es Sofía quien gana una ficha. Comienzan con un total de 20 fichas y el juego termina cuando no quedan más. ¿Te parece que este juego es equitativo? Si tuvieras que jugar, ¿Cuál jugador preferirías ser? (Godino y Batanero, 1996).

<u>Sesión 2</u>.

La segunda sesión tiene por objeto intentar generalizar los experimentos por medio de modelos de probabilidad. Se introduce los conceptos de variable aleatoria, función de probabilidad, experimento de probabilidad binomial y la esperanza matemática. A continuación, con uso de la planilla Excel y el applet del modelo binomial, se argumentan de forma visual y con cálculo de probabilidades los experimentos aleatorios del juego de dados. Se espera, a través de las actividades graduando su nivel, contar con una panorámica de representar los experimentos, anotando los resultados con objetos tangibles, solución de forma algebraica y de visualización computacional.

Actividad 2.1. Identifique las propiedades comunes de los siguientes experimentos.

a) Lanzar una moneda 50 veces y llevar un registro de las caras que se obtienen.

b) Si usted fuera un inspector en una línea de producción de una fábrica de televisores, estaría preocupado por identificar el número de televisores defectuosos.

c) Lanzar tres monedas. Calcular la probabilidad de obtener al menos dos caras.

d) Supongamos que el 10% de lámparas de una fábrica es defectuosa. Las lámparas se venden en cajas de 4 unidades. Simular el experimento consistente en abrir una caja y contar el número de defectos.

e) Un comerciante garantiza que ninguno de sus paquetes con 12 huevos contiene más de uno en mal estado. Si un paquete contiene más de un huevo en mal estado, él sustituirá toda la docena y permitirá que el cliente se lleve los huevos originales. Si la probabilidad de que un huevo individual esté en mal estado es de 0.05, ¿cuál es la probabilidad de que el comerciante tenga que reemplazar un paquete de huevos?

Actividad 2.2. En el caso de experimentos de probabilidad binomial podemos utilizar el microprograma applet disponible en la plataforma, para representar gráficamente y calcular probabilidades binomiales para distintos valores de sus parámetros *n* y *p.* ¿Qué observas de las gráficas para los siguientes valores de los parámetros? Valores para *p = 0,3 con n = 4, 8, 24 y para p = 0,1 con n = 4, 8 y 50.* Envía tus comentarios al Foro de EV@.

En esta Actividad, los alumnos pueden interactuar con el modelo de probabilidad binomial, cualquiera sean los parámetros *n* y *p*, por medio de un applet. Esperamos que estudien la forma de la función de probabilidad para distintos valores de sus parámetros.

Actividad 2.3. Usar el Applet disponible en la plataforma EV@. Un sistema está formado por 100 componentes cada una de las cuales tiene una confiabilidad igual a 0,95. (Es decir, la probabilidad de que la componente funcione correctamente durante un tiempo específico es igual a 0,95). Si esas componentes funcionan independientemente una de otra, y si el sistema completo funciona correctamente cuando al menos funcionan 80 componentes, se pretende estudiar acerca de la confiabilidad del sistema. ¿Cómo desarrollarías esta actividad con los conocimientos previos del Taller? ¿Qué aspectos destacarías en este problema?

Actividad 2.4. Experimento de probabilidad binomial. Un profesor aplica a sus alumnos un cuestionario sorpresa con cinco *ítems* de opción múltiple. Uno de los alumnos no ha estudiado el material del cuestionario y, por tanto, decide contestar los cinco ítems al azar, adivinando las respuestas sin leer las preguntas ni las opciones de respuestas.

Hoja de respuestas del cuestionario. Instrucciones: encierre en un círculo la mejor respuesta de cada ítem

1.	A	B	C
2.	A	B	C
3.	A	B	C
4.	A	B	C
5.	A	B	C

Encierre en un círculo sus respuestas antes de continuar. Antes de ver las respuestas correctas del cuestionario y encontrar cómo le fue a este alumno, se considerarán algunos hechos que podrían suceder si un cuestionario se contesta de esta forma.

a. ¿Cuántos de los cinco ítems es probable contestar correctamente?

b. ¿Cuán probable es contestar de manera adecuada más de la mitad de los ítems?

c. ¿Cuál es la probabilidad de elegir las respuestas acertadas de los cinco ítems?

d. ¿Cuál es la probabilidad de elegir las respuestas incorrectas de los cinco ítems?

e. Si todo un grupo contesta al azar el cuestionario, ¿cuál cree que será el número promedio de respuestas correctas del grupo?

Para responder estas interrogantes, de forma algebraica, se obtiene el espacio muestral Ω de 32 formas. Cada ítem fue contestado correcta(C) o incorrectamente (I).

Ω = {CCCCC, CCCCI, CCCIC, CCICC, CICCC, ICCCC, CCCII, CICIC, ICICC, CCIIC, CIICC, IICCC, ICCCI, ICCIC, CICCI, CCICI, CCIII, CIIIC, IIICC, ICICI, IICIC, CIICI, IICCI, ICCII, CICII, ICIIC, CIIII, ICIII, IICII, IIICI, IIIIC, IIIII}

Considere la variable aleatoria, X el "número de respuestas correctas" en el cuestionario cuando éste fue contestado al azar por alguna persona. La variable X puede adoptar cualquiera de los valores: 0, 1, 2, 3, 4 y 5 para cada cuestionario. Debido a que cada ítem individual tiene una sola respuesta correcta de las tres posibles, la probabilidad de elegir la respuesta correcta de un ítem individual es 1/3. La probabilidad de elegir una respuesta incorrecta en cada ítem es 2/3. Las probabilidades de cada valor X es:

La probabilidad de que cero ítem hayan sido contestados correctamente y que cinco hayan sido contestados incorrectamente es:

$$P(X = 0) = \frac{2}{3} \times \frac{2}{3} \times \frac{2}{3} \times \frac{2}{3} \times \frac{2}{3} = \left(\frac{2}{3}\right)^5 = \frac{32}{243} = 0.13169$$

La respuesta a cada ítem individual es un evento separado e independiente, por lo que es posible multiplicar las probabilidades.

La probabilidad de que un ítem haya sido contestado correctamente y los otros cuatro hayan sido contestados incorrectamente es:

$$P(X = 1) = 5 \times \frac{1}{3} \times \frac{2}{3} \times \frac{2}{3} \times \frac{2}{3} \times \frac{2}{3} = 5\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)^4 = \frac{80}{243}$$

$$= 0.32922.$$

La probabilidad de que dos ítems hayan sido contestados correctamente y los otros tres hayan sido contestados incorrectamente es:

$$P(X = 2) = 10 \times \frac{1}{3} \times \frac{1}{3} \times \frac{2}{3} \times \frac{2}{3} \times \frac{2}{3} = 10\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^3 = \frac{80}{243}$$

$$= 0.32922.$$

La probabilidad de que tres ítems hayan sido contestados correctamente y los otros dos hayan sido contestados incorrectamente es:

$$P(X = 3) = 10 \times \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} \times \frac{2}{3} \times \frac{2}{3} = 10\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^2 = \frac{40}{243}$$

$$= 0.16461.$$

La probabilidad de que cuatro ítems hayan sido contestados correctamente y uno tres hayan sido contestados incorrectamente es:

$$P(X = 4) = 5 \times \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} \times \frac{2}{3} = 5\left(\frac{1}{3}\right)^4\left(\frac{2}{3}\right)^1 = \frac{10}{243}$$

$$= 0.04115$$

Finalmente, la probabilidad de que los cinco ítems haya sido contestados correctamente es:

$$P(X = 5) = \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} = \left(\frac{1}{3}\right)^5 = \frac{1}{243} = 0.00412$$

Así, se obtiene la distribución de probabilidad para el cuestionario de cinco ítems. Las respuestas a las interrogantes son:

a. La ocurrencia más probable es obtener una o dos respuestas correctas.

b. Tener más de la mitad de las respuestas correctas se representa con 2 ó 3, su probabilidad total es 0.49383. Es decir, el cuestionario se aprueba sólo el 49% de las veces, si se contesta al azar.

c. Todas las respuestas son correctas $P(X = 5)$ sólo el 0.4% de las veces.

d. Las cinco respuestas son incorrectas $P(X = 0)$ sólo el 13% de las veces.

e. Puede esperarse que el promedio del grupo sea 1/3 de 5, o 1.666, respuestas correctas.

X	Elementos de Ω	Probabilidad de X	$X \cdot P(X)$
0	IIIII	0.13169	0
1	CIIII, ICIII, IICII, IIICI, IIIIC	0.32922	0.32922
2	CCIII, CIIIC, IIICC, ICICI, IICIC, CIICI, IICCI, ICCII, CICII, ICIIC	0.32922	0.65844
3	CCCII, CICIC, ICICC, CCIIC, CIICC, IICCC, ICCCI, ICCIC, CICCI, CCICI	0.16461	0.49383
4	CCCCI, CCCIC, CCICC, CICCC, ICCCC	0.04115	0.1646
5	CCCCC	0.00412	0.0206
sumas	32 elementos	1	1.66669

Muchos experimentos están compuestos de ensayos repetidos cuyos resultados pueden clasificarse en alguna de las siguientes categorías: éxito o fracaso. Estos experimentos se denominan experimentos de probabilidad binomial. Así, X se distribuye binomial con n = 5 y p = 1/3. La probabilidad de que tres ítems hayan sido contestados correctamente y los otros dos hayan sido contestados incorrectamente es:

$$P(X = 3) = \binom{5}{3}\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^2 = 10 \times \frac{4}{243} = 0.16461.$$

La esperanza de la variable aleatoria X está dada por

$$E(X) = n\,p = 5 \times \frac{1}{3} = 1.666$$

Actividad 2.5. Con apoyo de Excel, simular la aplicación del cuestionario anterior a 20 alumnos, de manera independiente.

a) Elabore una distribución de frecuencias de la variable “número de respuestas correctas en cada cuestionario” y grafique.

b) Estimar el número de respuestas correctas por alumno.

c) La próxima semana los 20 alumnos deben desarrollar otro cuestionario compuesto de cinco preguntas de verdadero y falso. Estime el número de respuestas correctas por alumno y compare.

d) Compare los resultados de cálculo de probabilidades obtenidos en la Actividad 2.4 mediante la planilla Excel y el Applet del modelo binomial.

Conclusiones

Enseñar conceptos y propiedades básicas de probabilidad y estadística de una manera práctica y con apoyo de recursos informáticos es aún un desafío con algunas dificultades como el tiempo requerido para el profesorado. Es deseable tener una visión amplia de lo que significa la estadística para que una práctica realista y significativa se implemente en el aula. El objetivo del taller, por un lado es la de animar a profesores en la dinámica de conducir actividades de experimentos aleatorios hacia su generalización con uso de dispositivos didácticos. Por otro lado, es la búsqueda y comunicación de los temas principales de la estadística y probabilidad en los diferentes niveles educativos.

Agradecimientos

Este trabajo es financiado en parte por el proyecto Fondecyt Nº11080071.

Referencias

Alvarado H. (2004). *Significados del teorema central del límite y sus campos de problemas en los textos de estadística para ingenieros.* Memoria de Tercer Ciclo, Universidad de Granada, España.

Alvarado, H. y Batanero, C. (2007). Dificultades de comprensión de la aproximación normal a la distribución binomial. *Números, 67*.

Batanero, C. (2001). *Didáctica de la Estadística*. Grupo de Investigación en Educación Estadística, Universidad de Granada.

Batanero, C. (2003). La simulación como instrumento de modelización en probabilidad. *Educación y Pedagogía, 35*, 37-64.

delMas, R. C., Garfield, J. B. y Chance, B. L. (1999). A model of classroom research in action: developing simulation activities to improve students´ statistical reasoning.

Journal of Statistic Education, *7*, 3. On line: http://www.amstat.org/publications/jse.

Godino, J. D., Batanero, C. y Cañizares, Mª. J. (1996). *Azar y Probabilidad*. Madrid: Síntesis.

Godino, J. D., Batanero, C. y Fonts, V. (2007). The onto-semiotic approach to research in mathematics education. *Zentralblatt für Didaktik der Mathematik*, *39*(1-2), 127-135.

Godino, J. D., Contreras, A. y Font, V. (2006). Análisis de procesos de instrucción basado en el enfoque ontológico-semiótico de la cognición matemática. *Recherches en Didactiques des Mathematiques, 26* (1), 39-88.

Moore, D. (1997). New pedagogy and new content: The case of statistics. *International Statistical Review*, *65* (2), 123-155.

Retamal, L., Alvarado, H. y Rebolledo, R. (2007). Understanding of sample distributions for a course on statistics for engineers. *Ingeniare*, Vol. 15, pp. 6-17.