

# **Desarrollo de ambientes de aprendizaje para promover el razonamiento estadístico: El caso de los intervalos de confianza**

Santiago Inzunza Cazares<sup>1</sup>  
Miguel Contreras Montoya<sup>2</sup>

## **Resumen**

En el presente artículo se presenta una propuesta alternativa de ambiente de aprendizaje basado en el uso de dos herramientas de software (Fathom y Excel) para la enseñanza y aprendizaje de la estimación de parámetros por intervalos de confianza, un tema ineludible en la mayoría de los cursos de estadística a nivel universitario. La propuesta se puso a prueba con un grupo de 17 estudiantes universitarios del área de estudios internacionales. Los resultados muestran que un elevado porcentaje de estudiantes lograron desarrollar un razonamiento adecuado sobre el tema, de acuerdo con un cuestionario administrado al final de las actividades y con las entrevistas realizadas a 2 alumnos.

## **INTRODUCCION**

La tecnología computacional ha mostrado un enorme potencial para ayudar a los estudiantes a comprender conceptos difíciles en probabilidad y estadística (Ben-Zvi, 2000; Mills 2004; Chance y Rossman 2006). Por ejemplo, por medio de simulación un estudiante puede explorar y comprender conceptos y principios (por ejemplo: las distribuciones de probabilidad, muestreo aleatorio y distribuciones de estadísticos) que de otro modo serían mucho más abstractos, contribuyendo con ello a mejorar la experiencia e intuición probabilística. Por su parte, en estadística, la computadora puede ser de gran ayuda en la automatización de cálculos laboriosos (por ejemplo: cálculo de medidas descriptivas como la desviación estándar y el coeficiente de correlación), en la exploración de datos y en la construcción de gráficas.

Teniendo en cuenta lo anterior, en el presente trabajo de investigación nos hemos propuesto diseñar y poner a prueba un ambiente de aprendizaje para la enseñanza de la estimación por intervalos de confianza, un concepto de inferencia estadística complejo para los estudiantes universitarios. Este tema involucra otros conceptos de suma importancia como son el error estándar, margen de error, confiabilidad, variabilidad y tamaño de muestra, cuya relación e impacto en la amplitud y confiabilidad de una estimación no

---

<sup>1</sup> Universidad Autónoma de Sinaloa, México, [sinzunza@uas.uasnet.mx](mailto:sinzunza@uas.uasnet.mx)

<sup>2</sup> Universidad Autónoma de Sinaloa, México, [info.contreras@gmail.com](mailto:info.contreras@gmail.com)

resulta fácil de comprender desde un enfoque tradicional basado solamente en el uso de fórmulas y tablas de probabilidad. Consideramos que la tecnología puede ayudar a visualizar y comprender las complejas relaciones que existen entre estos conceptos, sobre todo por la capacidad dinámica y de múltiples representaciones que tiene el software Fathom (Finzer et al., 2002) y la capacidad de Excel para generar una multitud de casos en forma recursiva.

## **ANTECEDENTES**

Un intervalo de confianza es un rango de valores calculado a partir de los datos de una muestra entre los cuales se estima que podría estar el valor de un parámetro de la población. Dado que la población no fue estudiada en su totalidad, toda estimación está sujeta a cierto nivel de confiabilidad, la cual indica el porcentaje de muestras que al ser tomadas en condiciones idénticas, el intervalo calculado estaría incluyendo el verdadero valor del parámetro. La idea matemática que da sustento a lo anterior y que permite calcular los límites del intervalo de confianza y su confiabilidad, consiste en que los valores muestrales que se usan para realizar la estimación tienen una distribución de probabilidad bien definida. De acuerdo con el teorema del límite central, dicha distribución es aproximadamente normal siempre que el tamaño de muestra sea mayor a 30, -lo cual es bastante frecuente en muchas aplicaciones de la estadística-.

Entonces, seleccionada una muestra aleatoria de una población, se calcula el valor muestral de interés el cual constituye una estimación puntual del valor del parámetro que se desea conocer; posteriormente y dado que se conoce su distribución de probabilidad, es posible calcular el margen de error, el cual se suma y se resta a la estimación puntual para formar el intervalo.

Los desarrollos que se utilizan para la formulación anterior en los cursos universitarios y en la mayoría de los libros de estadística, son expresados a través de un lenguaje matemático y teoría de la probabilidad que con frecuencia está fuera del alcance de muchos estudiantes, sobre todo aquellos que son de áreas no matemáticas. Sin embargo, a partir del desarrollo de la tecnología computacional aplicada a la educación estadística experimentado en los últimos años se sugiere con frecuencia la utilización de simulación computacional como alternativa para abordar la problemática del aprendizaje de la

inferencia estadística (Gordon y Gordon, 1992; Scheaffer, 1992, Meletiou\_Mavrotheris, 2004). Se señalan diversas ventajas de la simulación respecto al enfoque tradicional de enseñanza, como es el hecho de permitir un acercamiento empírico mediante la selección repetida de muestras de una misma población, calculando el estadístico en cada una de las muestras y acumulándolos para formar la distribución muestral, que es la base para los métodos de inferencia estadística. Este proceso está más relacionado conceptualmente con el proceso real de inferencia y requiere de pocos antecedentes matemáticos por parte de los estudiantes.

Para fijar ideas sobre los elementos que intervienen en un intervalo de confianza, consideremos el caso de uno de los estadísticos más comunes como es el caso de la proporción, y un tamaño de muestra mayor a 30 para asegurar que su distribución muestral es normal o aproximadamente normal.

Intervalo de confianza para la proporción: 
$$\hat{p} \pm Z_{\alpha} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Donde

- $\hat{p}$  es el estimador puntual obtenido de la muestra para estimar la proporción poblacional.

- $n$  es el tamaño de la muestra.

- $$\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

se denomina error estándar de la proporción.

- $Z_{\alpha}$  es un valor crítico que se toma distribución normal y está en función del coeficiente de confianza.

- $$Z_{\alpha} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

se denomina margen de error de la proporción muestral.

De esta manera, la comprensión del intervalo de confianza requiere de significados de otros objetos matemáticos previos (tanto conceptos como procedimientos), como población y muestra, estadístico y parámetro, error estándar, y cálculo del mismo para diversos

estadísticos, distribución muestral, valor crítico o uso de las tablas de diferentes distribuciones (Olivo y Batanero, 2007).

Con base en el análisis y reflexiones anteriores en el presente trabajo nos hemos planteado la siguiente pregunta de investigación: ¿En qué medida un ambiente computacional basado en el uso de simulación ayuda a que los estudiantes desarrollen un razonamiento adecuado y a que comprendan la relación que existe entre los diversos factores que intervienen en un intervalo de confianza?

## **PERSPECTIVA TEÓRICA**

En la presente investigación hemos adoptado el enfoque en el cual la computadora es vista como una herramienta cognitiva para el aprendizaje de las matemáticas en el sentido definido por Pea (1987) y un modelo de implementación para desarrollar ambientes de razonamiento estadístico propuesto por (Garfield y Ben-Zvi, 2008).

Para Pea (1987, p. 91) “una herramienta cognitiva es cualquier medio que ayuda a trascender las limitaciones de la mente, en el pensamiento, en el aprendizaje y las actividades de resolución de problemas”. Particularmente en el caso de las computadoras, constituyen una extraordinaria y potente herramienta cognitiva para aprender a pensar matemáticamente, con ellas se pueden operar no solo números, sino también símbolos, y permiten almacenar y manipular símbolos dinámicamente y permiten interacciones con los usuarios en tiempo real.

En cuanto al modelo de implementación, el presente trabajo retoma principios teóricos para crear Ambientes de Aprendizaje para el Razonamiento Estadístico (AARE) definidos por Garfield & Ben-Zvi (2008) y Cobb & McClain (2004). Estos principios se basan en las implicaciones de un enfoque constructivista y el uso de tecnología para una buena práctica de enseñanza; su propósito es estimular a los estudiantes a construir su conocimiento mediante actividades que les proporcionen oportunidades de pensar, razonar y reflexionar en su aprendizaje, además de la discusión y reflexión con sus compañeros.

Garfield & Ben-Zvi (2008), definen un Ambiente de Aprendizaje para el Razonamiento Estadístico (AARE) como un efectiva y positiva clase de estadística donde los estudiantes desarrollan una profunda y significativa comprensión, y la habilidad para pensar y razonar estadísticamente; “enfaticamos que es más que un libro de texto, actividades o trabajos que

damos a los alumnos. Es la combinación de materiales de texto, actividades en clase y cultura, discusión, tecnología, métodos de enseñanza y evaluación” (p. 48). El modelo AARE está basado en seis principios del diseño instruccional descritos por Cobb y McClain (2004), los cuales se describen a continuación:

1. Se enfoca en el desarrollo de las ideas estadísticas centrales en lugar de un conjunto de herramientas, técnicas y procedimientos de presentación.
2. Usa datos reales y motivadores para interesar a los estudiantes a probar conjeturas.
3. Usa actividades en clase para apoyar el desarrollo del razonamiento de los estudiantes.
4. Integra el uso de herramientas tecnológicas adecuadas que permitan a los estudiantes probar sus conjeturas, explorar y analizar datos, y desarrollar su razonamiento estadístico.
5. Promueve un discurso en clase que incluye argumentos estadísticos e intercambios sustentados que se enfoquen en ideas estadísticas significativas.
6. Usa el diagnóstico para aprender lo que los estudiantes saben y para monitorear el desarrollo de su aprendizaje estadístico para evaluar los planes de instrucción y su avance.

## **METODOLOGÍA**

El estudio se realizó con un grupo de 17 estudiantes universitarios de la Licenciatura en Estudios Internacionales de la Universidad Autónoma de Sinaloa mientras tomaban el curso de Probabilidad durante el primer semestre del ciclo escolar 2008-2009. Antes de iniciar con las actividades se había abordado el tema de intervalos de confianza y se habían desarrollado las expresiones para el intervalo de una media y una proporción. Se diseñaron dos actividades basadas en datos reales que obtuvimos de la compañía encuestadora Consulta Mitofsky ([www.consulta.com.mx](http://www.consulta.com.mx)). La primera sobre una encuesta que se realizó para conocer la proporción de mexicanos que tienen familiares en los Estados Unidos; la segunda relacionada la proporción de mexicanos que considera que la inseguridad es el principal problema del país (México); ambas actividades tienen contexto de interés para los estudiantes de estudios internacionales.

En la primera sesión se hizo énfasis en que los estudiantes construyeran distribuciones muestrales para diferentes tamaños de muestra (10, 20 y 50) con el propósito de que exploraran e identificaran visualmente y en términos de la desviación estándar, que a mayor tamaño de muestra disminuye la variabilidad, y por tanto se generan intervalos más estrechos pero con la misma confiabilidad.

Otra parte importante del uso de Fathom en que se puso en juego en ambas actividades (sesión 2 y 3) consistió en buscar la comprensión de la confiabilidad de un intervalo de confianza. El software permite simular un gran número de intervalos e identificar si el parámetro de interés se encuentra dentro o fuera del intervalo; el porcentaje de intervalos que capturan al parámetro debe ser igual o aproximado a la confiabilidad especificada. Para ello los estudiantes hicieron uso de la expresión del intervalo de confianza para una proporción que había desarrollado en clases anteriores y la cual involucra el margen de error. Un ejemplo de ello se muestra en la figura 1.

Measures from Sample of POBLACION					
	P_INSEGU	MARGEN_ERROR	LIMITE_INFERIOR	LIMITE_SUPERIOR	CAPTURA_p
191	0.19	0.0243151	0.165685	0.214315	CAE ADENTRO
192	0.202	0.0248847	0.177115	0.226885	CAE ADENTRO
193	0.215	0.025463	0.189537	0.240463	CAE ADENTRO
194	0.204	0.0249763	0.179024	0.228976	CAE ADENTRO
195	0.2	0.0247923	0.175208	0.224792	CAE ADENTRO
196	0.228	0.0260035	0.201996	0.254004	CAE FUERA
197	0.218	0.025591	0.192409	0.243591	CAE ADENTRO
198	0.198	0.0246988	0.173301	0.222699	CAE ADENTRO
199	0.198	0.0246988	0.173301	0.222699	CAE ADENTRO
200	0.197	0.0246517	0.172348	0.221652	CAE ADENTRO

Fig. 1: Tabla con los resultados de la simulación

En la cuarta sesión y en el marco del problema de la actividad 1 (mexicanos con familia en los EU), Excel fue utilizado para introducir una fórmula en función de los parámetros que inciden en un intervalo de confianza (el valor del estadístico  $p$ , la confiabilidad  $Z$ , y el tamaño de la muestra  $n$ ). Manteniendo fijos los dos primeros, se hizo variar el tamaño de la muestra. El propósito era que los estudiantes visualizaran el efecto del tamaño de la muestra y la confiabilidad en el margen de error y la amplitud de los intervalos, viendo los resultados generados y apoyándose mediante alguna gráfica como se muestra en la figura 2.

## RESULTADOS Y DISCUSIÓN

### ACTIVIDAD 1:

Una encuesta nacional realizada por CONSULTA MITOFSKY durante el primer trimestre de 2008 señala que 4 de cada 10 mexicanos tenemos un pariente trabajando o viviendo en los Estados Unidos.

1. Abre el programa Fathom y define la variable FAMILA EN EU.
2. Genera una instrucción que reproduzca la proporción de mexicanos que tienen familia en EU; es decir  $p=0.40$ , para muestras de 10, 20 y 50.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	p	0.4						nivel de confianza 90%					1.65	
2	1-p	0.6						nivel de confianza 95%					1.96	
3														
4														
5														
6				Confianza 90%	Confianza 95%	Confianza 90%	Confianza 95%	Confianza 90%		Confianza 95%				
7	n	desviación estándar	margen de error	margen de error	ancho del intervalo	ancho del intervalo	Limite inferior	Limite superior	Limite inferior	Limite superior	Limite inferior	Limite superior		
8	10	0.154919334	0.255616901	0.303641894	0.51233802	0.607283789	0.144383099	0.655616901	0.096358106	0.703641894				
9	20	0.109544512	0.180748444	0.214707243	0.361496888	0.429414485	0.219251556	0.580748444	0.185292757	0.614707243				
10	50	0.069282032	0.114315353	0.135792783	0.228630707	0.271585567	0.285684647	0.514315353	0.264207217	0.535792783				
11	100	0.048989795	0.080833162	0.096019998	0.161666323	0.192039996	0.319166838	0.480833162	0.303980002	0.496019998				
12	200	0.034641016	0.057157677	0.067896392	0.114315353	0.135792783	0.342842323	0.457157677	0.332103608	0.467896392				
13	300	0.028284271	0.046669048	0.055437172	0.093338095	0.110874343	0.353330952	0.446669048	0.344562828	0.455437172				
14	400	0.024494897	0.040416581	0.048009999	0.080833162	0.096019998	0.359583419	0.440416581	0.351990001	0.448009999				
15	500	0.021908902	0.036149689	0.042941449	0.072299378	0.085882897	0.363850311	0.436149689	0.357058551	0.442941449				
16	600	0.02	0.033	0.0392	0.066	0.0784	0.367	0.433	0.3608	0.4392				
17	700	0.018516402	0.030552063	0.036292148	0.061104127	0.072584296	0.369447937	0.430552063	0.363707852	0.436292148				
18	800	0.017320508	0.028578838	0.033948196	0.057157677	0.067896392	0.371421162	0.428578838	0.366051804	0.433948196				
19	900	0.016329932	0.026944387	0.032006666	0.053888774	0.064013332	0.373055613	0.426944387	0.367993334	0.432006666				
20	1000	0.015491933	0.02556169	0.030364189	0.05112338	0.060728379	0.37443831	0.42556169	0.369635811	0.430364189				
21	1500	0.012649111	0.020871033	0.024792257	0.041742065	0.049584514	0.379128967	0.420871033	0.375207743	0.424792257				
22														

Fig. 2: Cálculo de Intervalos de Confianza para diversos tamaños de muestra y nivel de Confiabilidad.

El propósito de esta primer actividad era que los alumnos identificaran visualmente y en términos de la desviación estándar, que a mayor tamaño de muestra disminuye la variabilidad y por tanto se generan intervalos más estrechos pero con la misma confiabilidad. Se les solicitó que construyeran gráficas con los resultados de la simulación y calcularan media y desviación estándar de cada distribución. Un ejemplo de los argumentos expresados por los estudiantes se muestra a continuación:

Cuando tomamos la muestra de 10 personas, la desviación estándar se presentaba distante de la media, ya que la desviación era de 0.15. En la segunda gráfica tomamos una muestra de 20 personas y los datos se van agrupar mas en el centro ya que la desviación disminuyó a 0.1. Para terminar con la muestra de 50 personas, los datos están más centrados, pues la desviación estándar disminuyó a 0,06. Como conclusión

determinamos que entre mayor es la muestra. El margen de error y la desviación estándar es menor y la realidad esta mostrada con mayor precisión (Silvia).

De los resultados anteriores podemos observar que los estudiantes han logrado apreciar el efecto del tamaño de muestra en propiedades importantes de las distribuciones muestrales, las cuales son la base para un acercamiento empírico a los intervalos de confianza. Las características de herramienta cognitiva de Fathom, como es el hecho de disponer de varias representaciones gráficas y cálculos de medidas descriptivas de manera simultánea para diferentes tamaños de muestra, les facilitó a los estudiantes la identificación del patrón de las distribuciones.

## ACTIVIDAD 2

En febrero de 2008 Consulta Mitofsky realizó una encuesta con una muestra aleatoria de 1000 mexicanos y les preguntó sobre qué problema consideraban principal para el país y el 20% contestó que la inseguridad. Simula el problema anterior (200 veces) y determina el margen de error e intervalo de confianza del 95% para la proporción de mexicanos que piensa que la inseguridad es el principal problema del país. Algunos resultados se muestran a continuación:

Al realizar la muestra con 200 personas, el 94% cayó dentro del intervalo y el 19.6% fue el resultado obtenido de las personas que creen que el problema más grande del país es la inseguridad, muy cercano al 20% que esperábamos (Luis)

La confiabilidad de la encuesta se muestra debido a que la mayoría cae dentro del intervalo; por su parte, las que cayeron fuera se encuentra en los límites (Gabriela).

Finalmente, en la siguiente tabla se muestran los resultados del cuestionario en términos del número y porcentaje de respuestas correctas por cada ítem obtenido por los estudiantes al final de las actividades.

Número de ítem	Respuestas correctas
1	12 (71%)
2	13 (76%)

3	9 (53%)
4	10 (59%)
5	10 (59%)
6	10 (59%)
7	12 (71%)

Los resultados muestran que muchos estudiantes lograron desarrollar un razonamiento adecuado sobre los conceptos que se involucran en los intervalos de confianza; sin embargo es importante señalar las principales dificultades que tuvieron muchos otros estudiantes para comprenderlos. Por ejemplo, en el ítem 3 que involucra la definición de intervalo de confianza, 8 de los 17 estudiantes consideraron que un intervalo de confianza especifica un rango de valores dentro de los cuales cae el parámetro con seguridad, cuando en realidad especifica un intervalo de posibles valores para el parámetro, y un porcentaje de intervalos que cubrirán aproximadamente dicho valor para el mismo tamaño de muestra. En el ítem 4, que requería identificar el efecto del tamaño de muestra en la precisión en un intervalo, 6 de los 17 estudiantes consideraron que ambos intervalos tienen la misma precisión, lo que muestra además que confundieron la precisión con la confiabilidad. En los ítems 5 y 6, la principal dificultad consistió en que muchos estudiantes no tienen claro el efecto del nivel de confiabilidad en el ancho de un intervalo.

En suma, el efecto del tamaño de muestra y la confiabilidad en el ancho de un intervalo de confianza, la confusión entre precisión y confiabilidad y la idea que un intervalo de confianza especifica un intervalo de valores que captura con seguridad a un parámetro, constituyeron las principales dificultades para los estudiantes; no obstante que las actividades estaban diseñadas para mostrar la relación entre estos conceptos; lo que demuestra la complejidad del concepto de intervalos de confianza, misma que con frecuencia es subestimada cuando se aborda en un ambiente de lápiz y papel centrado en el uso de fórmulas y procedimientos. Una explicación para ello podría ser el poco tiempo que se dedicó al abordaje de estos conceptos, pues fueron solo dos actividades donde se trabajó con ellos. La complejidad de estas relaciones ha quedado de manifiesto en otras investigaciones y se debe profundizar más en ellas en su enseñanza.

## CONCLUSIONES

La puesta a prueba del ambiente computacional para la enseñanza de la estimación de parámetros a través de intervalos de confianza -no obstante que consistió solo de un par de actividades-, muestra que los estudiantes pueden construir un razonamiento adecuado sobre conceptos estadísticos difíciles, sin necesidad de recurrir a conocimientos matemáticos avanzados como suele darse en el enfoque tradicional. El poder de simulación de Fathom y sus multiplicidad y flexibilidad de representaciones permitieron a los estudiantes explorar fácilmente la relación entre el tamaño de muestra, el margen de error, la amplitud del intervalo y la confiabilidad, y con ello construir un razonamiento adecuado sobre la relación entre ellos. Asimismo, Excel permitió calcular los intervalos de confianza para una gran cantidad de tamaños de muestra, con lo cual los estudiantes pudieron identificar algunos patrones de comportamiento en los conceptos involucrados. En un ambiente de lápiz y papel, esta exploración resulta muy difícil de darse, por lo que generalmente se recurre al cálculo aislado de los intervalos de confianza, y en ausencia de otros tipos de representaciones diferentes a las simbólicas.

## BIBLIOGRAFIA

- Ben-Zvi, D. (2000). "Toward Understanding the Role of Technological Tools in Statistical Learning". *Mathematical Thinking and Learning*, 2(2), Mahwah NJ: Lawrence Erlbaum Associates, Inc. pp. 127-155.
- Chance, B. y Rossman, A. (2006). "Using simulation to teach and learn statistics". En A. Rossman y B. Chance (eds.), *Proceedings of the Seventh International Conference on Teaching Statistics*. [CD-ROM]. Voorburg, The Netherlands: International Statistical Institute.
- Cobb, P. y McClain, K. (2004). "Principles of Instructional Design for Supporting the Development of Students' Statistical Reasoning". En D. Ben-Zvi y J. Garfield (eds.) *The Challenge of Developing Statistical Literacy, Reasoning and Thinking*. Springer Verlag, 375-395.
- Finzer, W., Erickson, T. y Binker, J. (2002). "Fathom Dynamic Statistics Software". Key Curriculum Press Technologies.