

**SEMINARIO - TALLER DE INTEGRACIÓN II
TRABAJO FINAL**

**LA IMPORTANCIA DE COMPRENDER LOS INTERVALOS
DE CONFIANZA**

Retegui, Darío Rubí

Profesorado Universitario en Matemática

Tutoras:

Dra. Camos, Cristina

Universidad Abierta Interamericana

Lic. Casetta, Inés

Universidad Abierta Interamericana

12 de Marzo de 2020



Universidad Abierta Interamericana

Facultad de Tecnología Informática

Índice

Índice general	1
Introducción	2
1. Objetos matemáticos intervinientes en el trabajo sobre	
Pruebas de Acceso a la Universidad (P.A.U.)	4
1.1. Estadística	4
1.2. Estadística Descriptiva	4
1.3. Estadística Inferencial	4
1.4. Población	4
1.5. Muestra	5
1.6. Variables y datos	5
1.7. Medidas Descriptivas	5
1.8. Teoría de la Probabilidad	5
1.9. Importancia de la muestra – Construcción del muestreo	6
1.10. La variable y su Distribución de Probabilidad	7
1.11. Construcción de la Distribución Binomial	7
1.12. Ley de los Grandes Números	11
1.13. Teorema Central del Límite	12
1.14. Intervalos de Confianza	15
2. El Enfoque Ontosemiótico como marco teórico en el estudio empírico de las P.A.U.	16
3. Investigaciones previas y metodologías utilizadas en el estudio empírico	19
4. Análisis sobre los problemas propuestos de intervalos de confianza y los contenidos evaluados en cada uno de ellos	22
5. Acerca de la importancia de la Estadística y Probabilidad en las Pruebas de Acceso a la Universidad (P.A.U.)	26
Conclusión	29
Referencias	31

Introducción

Los cambios vertiginosos y permanentes que se presentan en el mundo actual, nos llevan a recurrir a la Estadística y la Probabilidad con la finalidad de tomar decisiones lo más precisas posibles, ya sea en situaciones cotidianas o en el abordaje de diversas áreas, constituyéndose por su multidimensionalidad como una ciencia transversal.

El desafío de la Estadística es encontrar modelos estadísticos y probabilísticos que permitan incluir bases de datos muy grandes, permitiendo su análisis e interpretación de forma rápida y eficiente para la toma de decisiones.

Atendiendo a los conceptos mencionados, y por el futuro que presentan la Estadística y la Probabilidad recurrí a elegir un trabajo sobre el “Estudio Empírico de los problemas sobre intervalos de confianza en las Pruebas de Acceso a la Universidad (P.A.U.)” (López Martín *et al.*, 2016) realizado en España con alumnos que finalizan el bachillerato y deben obligatoriamente aprobar la prueba de selectividad para ingresar a la universidad.

Los autores son profesores, licenciados, doctores y colaboradores del grupo de investigación sobre Educación Estadística de la Universidad de Granada (UGR) España coordinado por Carmen Batanero.

En dicho artículo, se analiza el contenido de los ejercicios sobre intervalos de confianza en las (P.A.U.) propuestos en Andalucía en la especialidad de Bachillerato de Ciencias Sociales durante los años 2003 a 2014.

En la Sección 1 del presente trabajo se presentan múltiples y diversos objetos matemáticos sobre Estadística y Probabilidad que surgen del artículo analizado (López Martín *et al.*, 2016), de manera que alumnos, estudiantes avanzados y graduados recientes tengan disponibles conceptos anteriores que faciliten la comprensión del objeto matemático: intervalos de confianza.

En la Sección 2 se desarrolla brevemente el marco teórico utilizado en el Estudio Empírico sobre las P.A.U. bajo el análisis del Enfoque Ontosemiótico del Conocimiento y la Instrucción

Matemática desarrollado por Juan Díaz Godino y colaboradores, desde finales del siglo XX.

En la sección 3 se muestran algunas investigaciones previas con sus conclusiones sobre las pruebas de selectividad llevadas a cabo en diferentes lugares de España, consideradas por los autores en su trabajo original.

Además se presenta la metodología utilizada en el estudio empírico, la muestra formada por los problemas de intervalos de confianza en sus dos categorías: tipo B1 sobre la construcción e interpretación de los intervalos de confianza, y tipo B2 sobre la relación entre confianza, error de estimación y tamaño de la muestra.

En la Sección 4 se analizan dos problemas que aparecen en las pruebas de selectividad de los años 2006 y 2007, según la categoría tipo B1 - intervalos de confianza, y tipo B2 - relación confianza, error de estimación y tamaño muestral, de acuerdo a conceptos, propiedades y procedimientos evaluados.

A continuación de las secciones estudiadas, se agregan consideraciones acerca de la importancia y la multidimensionalidad de la Estadística y la Probabilidad en el mundo actual, y aún más en las Pruebas de Acceso a la Universidad de España para aquel estudiante que finaliza el Bachillerato.

1. Objetos matemáticos intervinientes en el trabajo sobre

Pruebas de Acceso a la Universidad (P.A.U.)

El abordaje de la definición de los objetos matemáticos es complejo por las múltiples miradas de los autores, intentaremos sistematizarlas y mostrar sus enlaces.

1.1. Estadística

Es la ciencia que podemos considerar como la aplicación del método científico para analizar datos y tomar decisiones. Desde un punto de vista histórico nace por ejemplo para registrar nacimientos y muertes, registrar propiedades, etc. Aparecen en el siglo XVII con Pascal y Chevalier de Mere estudios que sientan las bases de la Teoría de Probabilidades. Entonces podemos hablar de Estadística Descriptiva y Estadística Inferencial.

1.2. Estadística Descriptiva

Se encarga, mediante métodos de: recoger, ordenar y presentar un conjunto de datos para describir diversas características de ese conjunto de datos.

1.3. Estadística Inferencial

Podemos definirla como los métodos que se aplican para estimar una característica de una población o tomar decisiones respecto a esta población, habiendo obtenido un muestreo de la población estudiada.

1.4. Población

Es la totalidad o universo de cosas o artículos que consideramos para realizar un estudio.

1.5. Muestra

Es cualquier subconjunto de la Población que estudiamos.

1.6. Variables y datos

Al seleccionar una muestra para realizar un estudio, la misma es aleatoria, es decir no elijo uno a uno los elementos, por lo tanto, estudiaremos variables aleatorias de interés que producen dos tipos de datos: cuantitativos, que pueden medirse o cuantificarse, y se clasifican en continuos y discretos; o bien datos cualitativos que no tienen una medición, sino que representan una cualidad o un atributo.

1.7. Medidas Descriptivas

Para poder caracterizar los datos que obtuvimos de un determinado estudio se utilizan distintos tipos de medidas. Se clasifican en: Medidas de Posición, Medidas de Dispersión y Medidas de Forma. Dada que en este artículo se nombra la Media, mencionaremos algunas características de esta medida de posición.

Encontrar la media o media aritmética, es hallar el promedio de todos los datos obtenidos. Es un “estadístico” ya que se calcula sobre una muestra. También se menciona la desviación típica y la varianza que son medidas de dispersión.

La desviación estándar o típica es una medida que se utiliza para cuantificar la variación o dispersión de un conjunto de datos numéricos, por ejemplo, respecto de la media. La varianza es la desviación típica al cuadrado.

1.8. Teoría de la Probabilidad

El significado del término probabilidad no es único, “Existen actualmente muchas explicaciones útiles de las diversas ideas de probabilidad” (Hacking, 1995).

Consideramos la Probabilidad desde una perspectiva frecuencial y la definimos: la probabilidad de un suceso A asociado con un experimento aleatorio (ζ) corresponde al límite de la frecuencia relativa, cuando el número de experiencias tiende a infinito.

$$P(A) = \lim_{n \rightarrow \infty} \frac{n_A}{n}$$

donde n es la cantidad de repeticiones, n_A es la cantidad de veces que ocurre A en las n repeticiones, $P(A)$ es la probabilidad de suceso A .

El valor de esta perspectiva es que proporciona un modelo matemático adecuado a la descripción e interpretación de cierta clase de fenómenos observados y estudia sus consecuencias lógicas.

1.9. Importancia de la muestra – Construcción del muestreo

Debemos ser críticos y cuidadosos al momento de inferir y concluir acerca de una población, ya que las muestras se caracterizan por su variabilidad, es decir que una muestra puede ser distinta a otra.

Entre la muestra con la que se trabaja (n) y la población de interés (N), aparece la denominada población de muestreo. El muestreo es el método utilizado para seleccionar a los componentes de la muestra del total de la población.

El tipo de muestreo puede ser probabilístico o no probabilístico, el primero es el más recomendable en investigaciones cuantitativas, ya que todos los elementos de la población tienen la misma probabilidad de ser seleccionados para formar una muestra aleatoria, y además nos permitirá calcular el error de la muestra.

Recordemos que un experimento aleatorio es aquel en el que no es posible predecir su resultado. Por ejemplo si tiramos un dado no cargado, podremos obtener 1, 2, 3, 4, 5 o 6, sin saber de antemano el resultado.

Una muestra aleatoria es una colección de variables aleatorias que son independientes entre

sí y tienen la misma distribución que la variable aleatoria original.

Según Batanero y Romero (1995), un objeto es o no considerado como miembro aleatorio de una clase, dependiendo, en esta interpretación, de nuestro conocimiento sobre el mismo. Lo que puede ser aleatorio para una persona, puede no serlo para otra. La aleatoriedad no es una propiedad física “objetiva”, sino que tiene carácter subjetivo.

1.10. La variable y su Distribución de Probabilidad

Una variable aleatoria dijimos anteriormente que produce datos. Estos datos pueden o no ser números, pero a los fines de estudiarlos los transformamos siempre a valores numéricos. Estos valores corresponden al resultado de un experimento aleatorio.

Con el objeto de formalizar la definición de variable aleatoria, diremos que es una función que asocia números a los resultados de un determinado suceso incluido en un espacio muestral.

Los fenómenos aleatorios están representados por variables aleatorias que están relacionadas, por lo dicho anteriormente con sucesos de un espacio muestral, y por lo tanto con las probabilidades de estos sucesos. Si apareamos cada valor producido por una variable aleatoria con su probabilidad, creamos una distribución de probabilidad.

Por ejemplo: si un experimento consiste en “lanzar dos veces una moneda”, y se define la variable aleatoria como: “asignar a cada suceso simple el número que corresponde al número de caras”, se obtienen para los diferentes valores de la variable discreta 0, 1 y 2 la correspondiente probabilidad de $\frac{1}{4}$, $\frac{2}{4}$ y $\frac{1}{4}$ respectivamente. La suma de las probabilidades calculadas es 1.

1.11. Veamos ahora como construimos la Distribución de Probabilidad

Binomial, paso a paso:

Supongamos que se tiene una caja con 3 bolitas verdes (V) y 3 bolitas rojas (R), y se extraen 3 bolitas con reposición. Queremos calcular la probabilidad de obtener 2 bolitas verdes y una roja en las 3 extracciones.

Vamos a llamar “ p ” a la probabilidad de obtener una bolita verde, pero ocurre que puedo no obtener bolita verde, entonces llamaré “ $1 - p$ ” a la probabilidad de no obtener bolita verde. Para simplificar la notación a “ $1 - p$ ” lo llamamos “ q ”.

Ocurre que en este problema si no obtengo una bolita verde, necesariamente obtengo una roja, entonces la probabilidad de no verde es lo mismo que la probabilidad de si obtener roja. Luego lo escribimos en lenguaje matemático o simbólico:

Tomamos $P(V) = p$ en cada extracción, y la $P(R) = 1 - p = q$. Siendo p los éxitos y q los fracasos.

Si se quiere calcular la probabilidad de obtener 2 bolitas verdes y una roja en las 3 extracciones.

$$P(2) = \binom{3}{2} p^2 (1 - p)^1$$

$$P(2 \text{ bolitas verdes}) = P(2 \text{ bolitas verdes y 1 roja}) = P(VVR \text{ o } VRV \text{ o } RVV) =$$

Por regla de adición.

$$= P(VVR) + P(VRV) + P(RVV)$$

Por regla de multiplicación de eventos independientes.

$$\begin{aligned} &= P(V)P(V)P(R) + P(V)P(R)P(V) + P(R)P(V)P(V) \\ &= pp(1 - p) + p(1 - p)p + (1 - p)pp \\ &= p^2(1 - p) + p^2(1 - p) + p^2(1 - p) \\ &= 3p^2(1 - p)^1 \\ &= \binom{3}{2} p^2 (1 - p)^1 \end{aligned}$$

El coeficiente combinatorio $\binom{3}{2} = 3$ indica los diferentes órdenes posibles para reglas especiales

de conteo por permutaciones.

$$P(x=2) = \binom{3}{2} \left(\frac{3}{6}\right)^2 \left(\frac{3}{6}\right)^{(3-2)}$$

$$P(x=2) = 0,375$$

Si se calcula la probabilidad de obtener 1 bolita verde y 2 rojas en las 3 extracciones.

$$P(1) = \binom{3}{1} p^1 (1-p)^2$$

$$P(1 \text{ bolita verde}) = P(1 \text{ bolita verde y 2 rojas}) = P(VRR) \text{ o } P(RVR) \text{ o } P(RRV) =$$

Por regla de adición.

$$= P(VRR) + P(RVR) + P(RRV)$$

Por regla de multiplicación de eventos independientes.

$$\begin{aligned} &= P(V)P(R)P(R) + P(R)P(V)P(R) + P(R)P(R)P(V) \\ &= p(1-p)(1-p) + (1-p)p(1-p) + (1-p)(1-p)p \\ &= 3p^1(1-p)^2 \\ &= \binom{3}{1} p^1 (1-p)^2 \end{aligned}$$

El coeficiente combinatorio $\binom{3}{1} = 3$ indica los diferentes órdenes posibles para reglas especiales de conteo por permutaciones.

$$P(x=1) = \binom{3}{1} \left(\frac{3}{6}\right)^1 \left(\frac{3}{6}\right)^{(3-1)}$$

$$P(x=1) = 0,375$$

Luego se calcula la probabilidad de obtener 3 bolitas verde y ninguna roja en las 3 extracciones.

$$P(3) = \binom{3}{3} p^3 (1-p)^0$$

$$P(3 \text{ bolitas verdes}) = P(VVV)$$

Por regla de multiplicación de eventos independientes.

$$= P(V)P(R)P(R) + P(R)P(V)P(R) + P(R)P(R)P(V)$$

$$= P(V)P(V)P(V)$$

$$= ppp$$

$$= p^3 (1-p)^{(3-3)}$$

$$= \binom{3}{3} p^3 (1-p)^0$$

El coeficiente combinatorio $\binom{3}{3} = 1$ indica los diferentes órdenes posibles para reglas especiales de conteo por permutaciones.

$$P(x=3) = \binom{3}{3} \left(\frac{3}{6}\right)^3 \left(\frac{3}{6}\right)^{(3-3)}$$

$$P(x=3) = 0,125$$

Por último se calcula la probabilidad de obtener ninguna bolitas verde y 3 rojas en las 3 extracciones.

$$P(0) = \binom{3}{3} p^0 (1-p)^3$$

$$P(3 \text{ bolitas rojas}) = P(RRR)$$

Por regla de multiplicación de eventos independientes.

$$\begin{aligned} &= P(R)P(R)P(R) \\ &= (1-p)(1-p)(1-p) \\ &= 1p^0(1-p)^3 \\ &= \binom{3}{0}p^0(1-p)^3 \end{aligned}$$

El coeficiente combinatorio $\binom{3}{0} = 1$ indica los diferentes órdenes posibles para reglas especiales de conteo por permutaciones.

$$\begin{aligned} P(x=0) &= \binom{3}{0} \left(\frac{3}{6}\right)^0 \left(\frac{3}{6}\right)^3 \\ P(x=0) &= 0,125 \end{aligned}$$

Generalizamos obteniendo la Función de la Distribución Binomial — $B(n, p)$

$$P(x=k) = \binom{n}{k} p^k q^{(n-k)}$$

La Distribución Binomial es una distribución de probabilidad de una variable aleatoria discreta que describe el número de éxitos “ p ” al realizar un número “ n ” de experimentos independientes entre si, acerca de una variable aleatoria.

1.12. Ley de los Grandes Números

El teorema fundamental de la Teoría de Probabilidad es la Ley de los Grandes Números y explica por qué el promedio de una muestra tomada al azar de una población de gran tamaño, tenderá a estar cerca de la media de la población de interés.

Sean $X_1, X_2, \dots, X_n$ variables aleatorias independientes, con la misma distribución (misma media μ , varianza σ^2 y desvío poblacional σ). Entonces, para $S_n = X_1, X_2, \dots, X_n$ y cualquier

real $\varepsilon > 0$:

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) = 0$$
$$\lim_{n \rightarrow \infty} P\left(\left|\frac{S_n}{n} - \mu\right| < \varepsilon\right) = 1$$

donde X_i son las variables aleatorias independientes idénticamente distribuidas, μ la media poblacional, σ^2 la varianza poblacional, σ el desvío poblacional, S_n la sumatoria de las variables y n el número de ensayos.

Si repitiéramos muchas veces un mismo experimento (tendiendo a infinitas veces), la frecuencia de que suceda un cierto evento o suceso, tiende a coincidir con la probabilidad de que ocurra ese evento.

En muchos casos no podemos calcular la probabilidad por la Regla de Laplace (casos favorables/casos totales), y entonces se recurre a la Ley de los Grandes Números donde la probabilidad de un suceso es un número entre 0 y 1 alrededor del cual se estabiliza la frecuencia relativa de dicho suceso cuando se repite el experimento muchas veces en las mismas condiciones.

Por ejemplo, si tiramos un dado 10 veces y queremos obtener el número 1 puede ser que su frecuencia relativa sea 0,30. Si lanzamos el mismo dado 100 veces se obtendrá 0,18 por ejemplo, y si al experimento lo realizamos un número muy grande de veces llegaremos aproximadamente a 0,16, que es la probabilidad de obtener el número 1 cuando se tira un dado de 6 caras.

Se puede observar que la probabilidad experimental tiende a la probabilidad teórica a medida que aumenta el número de repeticiones del experimento, siendo la Ley de los Grandes Números la expresión matemática que formaliza lo desarrollado empíricamente

1.13. Teorema Central del Límite

De las distintas distribuciones de probabilidad que existen, la distribución Normal explica el mundo que nos rodea, siendo convincente en diversos campos como la Biología, la Zoología, la

Física, la Economía, etc.. El Teorema Central del Límite nos enseña el porqué de tal hipótesis.

Sean $X_1, X_2, \dots, X_n$ variables aleatorias independientes idénticamente distribuidas, cada una de ellas con función de distribución F , y supongamos que $E(X_k) = \mu$ y $\text{var}(X_k) = \sigma^2$ para cualquier elemento del conjunto. Si designamos a la suma normalizada de n términos con el símbolo:

$$S_n^* = \frac{X_1 + X_2 + \dots + X_n - n\mu}{\sigma\sqrt{n}}$$

entonces la sucesión de sumas normalizadas converge en ley a la variable aleatoria Normal tipificada $Z - N(0, 1)$, es decir:

$$S_n^* \rightarrow Z$$

donde X_i son las variables aleatorias independientes idénticamente distribuidas, μ la media poblacional, σ^2 la varianza poblacional, σ el desvío poblacional, S_n la sumatoria de las variables y n el número de ensayos.

El teorema anterior tiene dos corolarios importantes:

1. Si consideramos la suma ordinaria de las n variables aleatorias, es decir, $S_n = X_1 + X_2 + \dots + X_n$, entonces la sucesión de sumas ordinarias converge en ley a una Normal de media $n\mu$ y varianza $n\sigma^2$.
2. Si consideramos el promedio de las n variables aleatorias, es decir, $n^{-1}S_n$, entonces la sucesión de promedios converge en ley a una Normal de media μ y varianza $n^{-1}\sigma^2$.

Si se toma un conjunto de variables aleatorias independientes, y se obtiene la media y la varianza para ese conjunto de variables se podrá observar que tienden a una distribución Normal de media $n\mu$ y varianza $n\sigma^2$, siendo n lo suficientemente grande.

Ejemplo:

$$X_1 = 10; X_2 = 20; X_3 = 30$$

$$\mu = 20$$

$$\sigma^2 = \frac{200}{3} = 66,67$$

$$S_n = 60 = n\mu$$

$$S(x - \mu)^2 = n\sigma^2$$

Según Mendez (1991) existen cuatro propiedades básicas que deben tener presentes los estudiantes en la comprensión del teorema, ya que no son capaces de dar una explicación intuitiva de su significado o aplicarlo a casos reales.

- a) La media de la distribución muestral es igual a la media de la población, e igual a la media de una muestra cuando su tamaño tiende a infinito.
- b) La varianza de la distribución muestral es menor que la de la población cuando ($N > 1$)
- c) La forma de la distribución muestral tiende a ser acampanada a medida que se incrementa el tamaño muestral, y aproximadamente normal, independiente de la forma de la distribución en la población.
- d) La forma de la distribución muestral crece en altura y decrece en dispersión a medida que aumenta el tamaño muestral.

Si de una población se toman varias muestras ($n > 30$) y se calcula la media muestral de cada una de ellas, estaremos ante una distribución de medias. Esta distribución tenderá a una distribución Normal sin importar la distribución de la población. Mientras más muestras se seleccionen menor será la variación entre el estimador ($\bar{x}$) y el parámetro poblacional (μ).

Un estimador es un estadístico (una función de la muestra) utilizado para estimar un parámetro desconocido de la población.

Este teorema nos ayuda a entender el desvío que tiene la media muestral respecto de la media poblacional, dando lugar a otros objetos matemáticos como los intervalos de confianza, test de hipótesis y estimación del tamaño de muestra.

Por ejemplo, el 95 % de los datos normalmente distribuidos están a ± 2 desvíos estándar de la media. Por eso se tiene la probabilidad del 95 % que la media de la muestra ($\bar{x}$) esté a 2 errores

estándar del parámetro poblacional (μ).

Ejemplo 1: Si la variable es el “lanzamiento de una moneda” se puede observar que se ajusta a una distribución Binomial, pero si lanzamos 75 veces la moneda (cada una independiente entre sí) se distribuye según una distribución Normal.

Ejemplo 2: Si el peso de una notebook sigue una distribución normal con una media de 2 Kg. y un desvío estándar de 0,2 Kg. Determine la probabilidad de que una computadora elegida al azar entre 100 de ellas, tenga un peso mayor de 2,10 kg..

$$N(2; 0,2)$$

$$Z = \frac{2,10 - 2}{0,2}$$

$$Z = 0,5$$

$$P(x > 2,10) = 0,3085 \text{ valor según tabla de distribución de probabilidad Normal.}$$

En la inferencia estadística existe la posibilidad de encontrar un límite inferior y superior a los parámetros de interés en las poblaciones, obteniendo intervalos de confianza para ellos.

1.14. Intervalos de Confianza

Los intervalos de confianza permiten estudiar la precisión de la estimación de estudio y la relevancia de los resultados, mediante pares de números entre los cuales se estima que estará cierto valor desconocido, con una determinada probabilidad de acierto.

Del análisis de una muestra se proporciona “alguna” información de la población, aumentando el conocimiento de la misma, es por ello que entran en juego: la representatividad y la variabilidad muestral.

Otro de los contenidos de la estadística inferencial es el “test de hipótesis” o “contraste de hipótesis”, que abarca varios procedimientos estadísticos con un denominador común la hipótesis nula, hipótesis alternativa, nivel de significación, etc..

“La aplicación correcta de esos procedimientos precisa muchos tipos de elecciones, incluyendo: el tamaño de la muestra, el nivel de significación (α) y el estadístico apropiado” (Batanero, 2001).

2. El Enfoque Ontosemiótico como marco teórico en el estudio empírico de las P.A.U.

Dado que el estudio analizado utiliza como marco teórico el Enfoque Ontosemiótico, a continuación se desarrollan algunos conceptos de esta teoría con el fin de facilitar la comprensión del presente trabajo. Consideraremos su impronta, sus lineamientos, los roles del alumno, del docente y de las instituciones educativas en los procesos de enseñanza y de aprendizaje.

El Enfoque Ontosemiótico del Conocimiento y la Instrucción Matemática desarrollado por Juan Díaz Godino a finales del siglo XX, conlleva en sus fundamentos iniciales la formulación de una ontología de los objetos matemáticos que tiene en cuenta el triple aspecto de la matemática: como actividad de resolución de problemas, socialmente compartida; como lenguaje simbólico; y como sistema conceptual lógicamente organizado, pero teniendo en cuenta además la dimensión cognitiva individual (Godino y Batanero, 2007).

El análisis ontológico se caracteriza por la naturaleza de ser, la existencia y la definición de los objetos matemáticos, ya que difícilmente se podrían estudiar los procesos de enseñanza y de aprendizaje de objetos difusos o indefinidos.

Esta corriente didáctica articula diversas dimensiones configuradas por significados sistémicos, funciones semióticas, configuraciones didácticas, y todos aquellos elementos suficientes y necesarios que conforman cualquier proceso educativo.

En cuanto a los significados sistémicos, de acuerdo a la concepción y al involucramiento que las instituciones tengan sobre el objeto de estudio, el estudiante se apropiará de dichos significados durante el proceso de enseñanza-aprendizaje.

La enseñanza implica la participación del estudiante en las instituciones que soportan los significados institucionales, y el aprendizaje supone la apropiación de dichos significados por parte de estudiante.

En este sentido, “las instituciones se conciben como “comunidades de prácticas”, es por ello que incluye grupos étnicos, las culturas y contextos sociales, dando lugar a la relatividad socioepistémica de los sistemas de prácticas, de los objetos emergentes y los significados” (Godino y Batanero, 2007).

Las situaciones problemáticas, tipificadas en el E.O.S. como objeto matemático primario, presentadas a los estudiantes en las P.A.U. aportarán a constituir sistemas conceptuales más complejos que propiciarán el aprendizaje de los diferentes objetos matemáticos, y por tanto incidirán en el significado que los estudiantes les asignarán.

El marco teórico del artículo en análisis, interpreta el significado de un objeto matemático desde dos perspectivas, por una parte considera el significado que la institución comparte y concreta para los estudiantes, y por otra parte el significado que adquiere para una persona, que puede ser diferente al aceptado dentro de la institución.

“Dentro de la institución se diferencian, el significado global que es el más amplio (el significado de la inferencia en la propia Estadística); el significado pretendido del objeto dentro de un nivel educativo fijado por la institución (marcado por las orientaciones curriculares); el significado implementado en la enseñanza (lo que el profesor y los estudiantes trabajan en el aula); y el significado evaluado que se refiere al contenido relacionado con el objeto matemático en las pruebas de evaluación” (López Martín *et al.*, 2016).

En el estudio empírico se utiliza la noción de “significado institucional”, entendido como sistema de prácticas, para diseñar, implementar y analizar el objeto matemático intervalos de confianza en las pruebas de selectividad.

Los autores de este trabajo se centran en determinar el significado institucional evaluado del objeto matemático, intervalo de confianza en las P.A.U., con la finalidad de recibir respuestas de esas pruebas de selectividad, y desde allí determinar futuras planificaciones del tema en estudio para llevar al aula.

Se trata de que el significado evaluado tenga una estrecha relación con el significado institucional pretendido, construido por el docente a partir de un conjunto de variables como

los objetivos a alcanzar, las propuestas didácticas, la utilización de recursos, etc., todo ello a partir siempre del significado institucional de referencia.

Es difícil elaborar planificaciones de intervalos de confianza, en base a los resultados de las pruebas de acceso a la universidad dado que muchas veces son escasos y deficientes, lo que lleva a una revisión profunda de los objetos matemáticos básicos de estadística y probabilidad en el bachillerato.

Una buena manera de asegurar que los alumnos adquieren un buen significado personal de un objeto matemático, consiste en conseguir un buen significado personal de los objetos previos sobre estadística y probabilidad.

Es necesaria e indispensable la relación entre los significados personales y los significados institucionales en un sistema de prácticas que realiza un estudiante o compartidas en una institución (comunidad) para resolver situaciones problemáticas en el proceso de enseñanza-aprendizaje.

3. Investigaciones previas y metodologías utilizadas en el estudio empírico

El estudio empírico analizado está basado en investigaciones previas realizadas por varios autores que han estudiado las pruebas de selectividad en España referentes a la asignatura Matemática o Matemática Aplicada a las Ciencias Sociales, tomando muestras en diferentes provincias españolas.

Se han encontrado resultados muy bajos por clasificar a las pruebas de acceso a la Universidad de España como muy difíciles. En el País Vasco inesperadamente hubo buenos resultados en alumnos que utilizaban una metodología tradicional con poca presencia de tecnología.

Se realizó el estudio en las P.A.U. de Matemática en la Universidad de Barcelona en el año 2012, y se observó que los estudiantes resuelven los problemas en forma mecánica, sin precisión en los planteamientos y con poca creatividad en la solución de los problemas (Mallart Solaz, 2014).

También se concluyó que el tiempo destinado a la enseñanza de la estadística metodológica o inferencial en el bachillerato es escaso, como para dominar y manejar los contenidos que se presentan en las situaciones problemáticas de las pruebas de evaluación.

La metodología utilizada en el estudio es de tipo cualitativa, basada en el análisis de contenido, centrándose en el análisis de distintos significados y una adecuada aplicación del contenido. Para realizar este análisis se construyó una tabla cuyos contenidos serán explicitados en la siguiente Sección.

Según la investigación, se clasificaron los problemas, teniendo en cuenta el tipo de situaciones problemáticas que se le pide al estudiante que resuelva en las P.A.U., en dos categorías de actividades de acuerdo a los contenidos matemáticos evaluados por la universidad (conceptos, propiedades y procedimientos), B1-Cálculo e interpretación de Intervalos de

Confianza y B2-Relación entre confianza, error de estimación y tamaño muestral en un intervalo de confianza.

De los 144 problemas seleccionados por sorteo y analizados durante el período 2003-2014, son 128 ítems que corresponden a intervalos de confianza (43 relacionados con tipo B1 y 85 tipo B2), esto se debe a que los problemas tenían varios apartados, donde se trabajan ambos tipos de problemas. Todo indica que el estudio se ha realizado a lo largo del tiempo, siempre poniendo la atención en el análisis de los diversos significados y la correcta aplicación de los contenidos.

Los contenidos matemáticos evaluados en los dos campos de problemas (tipo B1 y tipo B2) que se proponen en las pruebas P.A.U., son conceptos y propiedades: la población y la muestra, la distribución muestral, la relación entre estadísticos en la población y en la distribución muestral, los modelos teóricos de las distribuciones, el estimador de un parámetro, el intervalo de confianza, el coeficiente de confianza, la relación entre la amplitud de intervalo, la precisión y tamaño muestral y probabilidad condicional.

Los procedimientos para el campo de problemas del tipo B1 son la tipificación, la lectura de tablas de Distribución Normal, el cálculo de probabilidades en intervalos, el cálculo de percentiles en la Distribución Normal y el cálculo de extremos en intervalo de confianza.

En el gráfico “Distribución de Campos de problemas en el tiempo” (2003 a 2014), presentado en la investigación “Estudio Empírico de los problemas sobre intervalos de confianza en las Pruebas de Acceso a la Universidad” (López Martín *et al.*, 2016) se observa que en todos los años analizados, excepto 2014, el número de problemas propuestos del tipo B2 es siempre superior a los de tipo B1, dado que los ítems del tipo B2 permiten conseguir mayor precisión en la estimación realizada, donde el estudiante debe relacionar el nivel de confianza, error de estimación y tamaño muestral.

Se le solicita al estudiante mayor precisión en las resoluciones de los problemas, pretendiendo que encuentre la relación entre los objetos matemáticos intervinientes, siendo que muchas veces desconoce o no logra distinguirlos desde los conceptos y contenidos. De esta

manera no es posible observar el comportamiento de algunas variables ante la modificación de otras, desembarcando en errores y planteos dificultosos en el proceso de solución.

Este trabajo empírico sobre el análisis de los problemas de inferencia estadística en las Pruebas de Acceso a la Universidad aporta información original obtenida del campo de investigación, clasificando los problemas de inferencia y tratando de identificar las variables que los conforman.

4. Análisis sobre los problemas propuestos de intervalos de confianza y los contenidos evaluados en cada uno de ellos

Según el trabajo de investigación que estamos analizando, los alumnos tienen serias dificultades para resolver los problemas propuestos de inferencia estadística. Una de las causas, según los investigadores se atribuye a la cantidad de objetos matemáticos que intervienen en su resolución.

En España, los alumnos que cursan Bachillerato de Humanidades y Ciencias Sociales reciben menor formación en contenidos matemáticos que aquellos que cursan el Bachillerato de Ciencias. Desde ya es previsible que los estudiantes de Ciencias Sociales y Humanidades tengan mayores dificultades para comprender el tema de inferencia.

En algunos países, como en Argentina, son pocos los docentes que llegan a enseñar a los estudiantes de nivel medio conceptos básicos de Estadística y Probabilidad, ya sea por la falta de tiempo en sus clases, por los contenidos previos que los alumnos poseen, o bien por la existencia de un currículum prescripto vacuo.

Muchas veces, por lo mencionado anteriormente, el análisis estadístico de datos es enseñado y aplicado como un proceso mecánico, sin poder el alumno visualizar y descubrir situaciones de la realidad en las que pueda aplicar los conceptos aprendidos.

Entre los distintos tipos de campos de problemas que son tratados en las P.A.U., se seleccionaron únicamente los relacionados con el contenido de intervalo de confianza. Estos a su vez se clasificaron en dos categorías principales: tipo B1-cálculo o interpretación de un intervalo de confianza; y tipo B2-relación entre confianza, error de estimación y tamaño muestral en un intervalo de confianza.

Respecto a la categoría tipo B1 se analizó un problema incluido en la prueba de acceso en el curso académico 2005/2006 (López Martín *et al.*, 2016) cuyo enunciado decía: “Ítem 1A (2006). De 500 encuestados en una población, 350 se mostraron favorables a la retransmisión

de debates televisivos en tiempos de elecciones. Calcule un intervalo de confianza, al 99,5 % para la proporción de personas favorables a estas retransmisiones”.

El alumno luego de leer y comprender el texto del problema deberá identificar el estimador puntual, la distribución asociada a este estimador que en este caso es la distribución binomial. Luego debe distinguir que como la muestra tiene tamaño mayor a 30, se distribuye aproximadamente según una distribución normal y podemos tipificar.

También debe distinguir el coeficiente de confianza, para luego poder hallar los límites del intervalo. Sumamos a todo lo enunciado el debido uso que debe poseer de la Tabla de distribución normal para encontrar los valores de Z .

Es importante además que pueda representar la situación en la –Campana de Gauss– con diversas variantes (unilateral o bilateral), ubicar la desviación del parámetro poblacional y utilizar este gráfico para dar sentido a los resultados obtenidos.

Otro objetivo a lograr por parte del estudiante es poder expresar en lenguaje natural lo que ha obtenido.

En el caso de este problema y luego de realizar los cálculos, diremos que: el intervalo de confianza, a un nivel de confianza del 99,5 %, para la proporción de personas favorables a las retransmisiones de debates televisivos en tiempos de elecciones es (0,6421; 0,7576).

Respecto a la categoría tipo B2: Se analizó un problema incluido en la prueba de acceso en el curso académico 2007 (López Martín *et al.*, 2016) cuyo enunciado decía: “Ítem 3B (2007). Se sabe que (45,13; 51,03) es un intervalo de confianza, al 95 %, para la media de una variable aleatoria que sigue una distribución Normal con desviación típica 15. a) ¿Cuál es el error cometido? b) Calcule, con el mismo nivel de confianza, el tamaño muestral mínimo necesario para que el error no sea superior a 1,8”.

El estudiante después de leer y comprender la situación problemática deberá diferenciar entre lo que es un estadístico (media de la muestra) y un parámetro (media de la población); comprender el significado del intervalo de confianza, ya que en este caso no sabemos si incluye o no el parámetro.

El enunciado no da información acerca de la distribución de la población, pero si afirma que la media de una variable aleatoria, sigue una distribución Normal.

También para el segundo apartado debe recordar, que el Error de estimación (E) es la mitad de la amplitud del intervalo de confianza. En este mismo apartado se pide disminuir el error, manteniendo el intervalo de confianza, con lo cual debe plantear la fórmula y darse cuenta de cuál es la única variable que puede calcular que es el tamaño de la muestra.

Es importante que los estudiantes puedan comprender la relación que existe entre el error de estimación y el tamaño muestral. A mayor tamaño de la muestra, mayor precisión en la estimación realizada porque el intervalo será más estrecho.

Los investigadores consideran que existen conceptos y propiedades comunes evaluados en las categorías de problemas analizadas, los de población, muestra, distribución muestral y la relación entre los estadísticos de la muestra, la población y la distribución muestral.

En el estudio empírico se elabora una tabla con contenidos matemáticos evaluados en los diferentes campos de problemas (tipo B1 y tipo B2) dividida en dos partes, la primera parte muestra los conceptos y propiedades de los objetos matemáticos que aparecen en los problemas de las pruebas de selectividad. Por ejemplo Población y muestra, modelos teóricos de distribución, estimador de un parámetro, intervalos de confianza, etc.

En la tabla se expresa que la relación entre la amplitud del intervalo, la precisión y el tamaño muestral es evaluada como concepto y propiedades, solamente en los problemas de la categoría tipo B2, como se observa en el problema antes presentado.

Los procedimientos evaluados forman la segunda parte de la tabla, la tipificación, la lectura de tablas y el cálculo de probabilidades y extremos en los intervalos de confianza. Todos ellos son utilizados y evaluados en la categoría de problemas B1, siendo la lectura de tablas de la distribución Normal el único procedimiento compartido entre los dos campos de problemas tipo B1 y tipo B2 ya que se aplica en ambos.

El estudiante tiene que diferenciar la distribución de datos en la muestra, la distribución de la variable estudiada en la población y la distribución muestral del estadístico en todas las

posibles muestras seleccionadas de la población. Se observa que los alumnos cometen errores confundiendo el estimador con la media poblacional.

“Los dos tipos de problemas que se analizan se apoyan en la Probabilidad condicional, interviniente en la definición de las distribuciones muestrales (condicionada al valor del parámetro), intervalos de confianza (condicionados con el valor del estimador) como en el contraste de hipótesis (regiones de aceptación y rechazo se calculan bajo la condición de ser cierta la hipótesis nula)” (López Martín *et al.*, 2016).

Se entiende que la probabilidad es condicionada por diferentes objetos matemáticos y eso hace más difícil comprender los objetos matemáticos que el estudiante tiene que comprender para resolver los problemas en las P.A.U..

“Dado que el concepto de probabilidad condicional es difícil y se han descripto números sesgos sobre el mismo, los profesores deberán prestar especial atención a su uso en el tema de inferencia, para asegurar un aprendizaje significativo en los estudiantes, que influirá en la correcta solución que realicen de los problemas propuestos sobre las pruebas de acceso” (Díaz *et al.*, 2012).

Las P.A.U. han demostrado que los estudiantes tienen grandes dificultades en el momento de resolver los problemas de inferencia estadística propuestos, no sólo en relacionar variables sino en razonar los efectos que producen las fluctuaciones de algunas de ellas permaneciendo las restantes sin cambios.

5. Acerca de la importancia de la Estadística y Probabilidad en las Pruebas de Acceso a la Universidad (P.A.U.)

En países como Estados Unidos, Inglaterra, Dinamarca, Suecia, Alemania y otros del primer mundo, gobernantes, instituciones educativas, bancos, negocios y las sociedades en su conjunto conciben a la Estadística como una recopilación y organización de datos numéricos, encuestas, tablas y diversos tipos de gráficos que ayudan a tomar decisiones ante las diferentes alternativas que se presentan a diario en los escenarios políticos, sociales, empresariales y educativos.

La base de las acciones administrativas y políticas en un Estado con cultura Estadística están centradas en el estudio de los datos estadísticos, ya que existe la posibilidad de cuantificar la realidad y disponer de los elementos para actuar de forma más coherente con eficacia y eficiencia.

La Estadística permite que situaciones impredecibles se puedan volver cada vez más explicativas. Ella se expresa a través de los números, muchas veces como un medio, como una herramienta que nos invita a tomar una decisión lo más cercana a la realidad y otras veces no tanto, dando respuesta a muchas de las necesidades que las sociedades actuales pueden experimentar.

El Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas de la Universidad Nacional Autónoma de México expresó que "día a día se acumula más información y por ello, es necesario formar gente capaz de poder analizarla e interpretarla, para que ésta "llegue a los tomadores de decisiones".

Para extraer información precisa y útil de fenómenos complejos y con grandes volúmenes de datos es necesario que los principios estadísticos sean acompañados de algoritmos computacionales de aprendizaje y obtención de conocimiento, es por ello el surgimiento de la minería de datos (data mining).

Ya hace aproximadamente treinta años que la Estadística viene ocupando un lugar muy

importante en el desarrollo social de las naciones, “específicamente por su contribución a la generación de distintos indicadores para medir el bienestar social de los pueblos” (Barreto-Villanueva, 2012).

Desde principio de este siglo se están realizando diversos estudios e investigaciones acerca de las nuevas formas de medición, aspectos teóricos, metodológicos y de aplicación, presentando oportunidades para nuevos lineamientos y procedimientos que forman la metodología estadística, como así también la formación de líneas independientes de la Estadística.

Para darle el verdadero valor a la Estadística y Probabilidad como objeto de estudio matemático es imprescindible que una sociedad tenga confianza en ella, tanto en las mediciones de las variables de las muestras elegidas, su análisis estadístico con sus test de bondades, sus métodos y distribuciones estadísticas, y sobre todo que se tomen decisiones que no se aparten del mundo real.

Si partimos que en los países desarrollados se buscan estadísticos para entender el mundo, ante la gran masa de datos que requiere una transformación rápida en información para la toma de decisiones, nos da lugar a los estudiantes de profesorado, profesores e investigadores en Matemática a contar con un currículo prescripto que incluya a la Estadística y la Probabilidad en los planes de estudio de edades escolares más tempranas, que formen a ciudadanos que comprendan, interpreten, entiendan los problemas y luego elijan las alternativas que más se ajustan a la realidad.

La preocupación por propiciar una cultura estadística que forme parte de la cultura general para las sociedades del conocimiento y de la información, prevé “una tarea más generalizada, que desde luego involucrará a los profesionales de la Estadística, pero sobre todo a los administradores, a los actores de sistemas educativos y a los medios de comunicación” (Barreto-Villanueva, 2012).

En ese contexto, los alumnos podrían desarrollar sus competencias y conocimientos guiados hacia estadística y probabilidad no sólo como una herramienta que utilizarán para resolver

situaciones problemáticas de la vida cotidiana, en post de tomar medidas o decisiones lo más eficientes posibles, sino como una ciencia en la que existe un método para el avance o mejoramiento de alguna situación en un área determinada.

La Estadística de este siglo permite medir otras variables, que ya no son materiales ni estrictamente económicas, aceptando la construcción de indicadores alternativos del progreso económico y la prosperidad de las sociedades actuales.

En Argentina los resultados de las “Evaluaciones Aprender 2016-2017” (Polack, 2019) indican que son muy pocos los estudiantes egresados del nivel medio que han aprendido Estadística y Probabilidad en la escuela media.

Es necesaria una revisión a fondo de todos los contenidos de Estadística y Probabilidad a enseñar, sobretodo en una sociedad que solicita empleos calificados, impulsados por el desarrollo tecnológico, por los volúmenes de la nueva información y por las vertiginosas comunicaciones.

Más allá de los resultados, se observa una amplia y clara diferencia entre las escuelas medias de Argentina sobre el estudio de la Estadística y Probabilidad, su interpretación y utilidad respecto a otros países, como España, que incorporan problemas sobre intervalos de confianza en las Pruebas de Accesos a la Universidad (P.A.U.).

Conclusión

Los autores concluyen en su estudio que es necesario que los alumnos del nivel medio que deben aprobar las Pruebas de Acceso a la Universidad (P.A.U.) comprendan los múltiples objetos matemáticos que intervienen en la construcción de los intervalos de confianza, y en la relación del nivel de confianza, error de estimación y tamaño muestral.

Los escasos contenidos de estadística y probabilidad enseñados por el poco tiempo que disponen los estudiantes para dominar la inferencia estadística, y la falta de aplicaciones prácticas en el bachillerato, muestra el fracaso en los resultados de las pruebas de acceso a la universidad española.

Los distintos significados sobre un objeto matemático, la correcta aplicación de los contenidos, propiedades y procedimientos llevaron a conformar dos campos de problemas para su estudio, relacionados con los intervalos de confianza en las pruebas de selectividad.

Desde una perspectiva Ontosemiótica del conocimiento matemático, dependiendo del significado institucional evaluado en las P.A.U. sobre el objeto matemático intervalos de confianza, se determinará la planificación del tema a enseñar y la formación que los estudiantes de nivel medio recibirán.

Las nociones teóricas y el conjunto de normas que condicionan la enseñanza y los aprendizajes deben permitir intervenciones de profesores y estudiantes, con el fin de mejorar el funcionamiento y el control de los sistemas didácticos, en busca de la evolución de los significados personales hacia los significados institucionales pretendidos.

Se necesita que los profesores y los formadores del bachillerato tengan presente la configuración epistémica del objeto “Intervalo de Confianza”, eso permitirá que durante el proceso de enseñanza-aprendizaje emerjan objetos matemáticos subyacentes, con sus propias definiciones, lenguajes, proposiciones y argumentos.

Para que el proceso de enseñanza-aprendizaje sea enriquecedor es necesario que el estudiante plantee un enunciado sobre una situación que considere, y luego la resuelva aplicando

conceptos estadísticos adquiridos previamente, dado que el razonamiento estadístico va más allá del conocimiento matemático y de la comprensión de los conceptos y procedimientos.

Los docentes y los futuros profesores de matemática tienen el desafío de formar sujetos que tienen excesiva cantidad de información con un fácil manejo de redes, pero que necesitan situaciones problemáticas llamativas, interesantes y con aplicación concreta en el mundo real. Se debe recurrir a la creatividad, a la innovación permanente, a la investigación, a la exploración de los objetos que forman parte de la realidad.

El uso de aplicaciones tecnológicas estadísticas, ya sea desde la computadora o desde el celular, les aportan al estudiante diferentes representaciones significativas sobre un objeto matemático, y de allí su interpretación y análisis.

El profesor tiene que promover el cambio de una representación a otra, desde su intervención activa en el aula, aplicando los criterios de idoneidad didáctica, logrando que los alumnos muestren creatividad en la resolución de problemas y precisión en la ejecución de las operaciones, no recayendo en resoluciones mecánicas.

Mientras que el análisis exploratorio de datos se apoya fuertemente en representaciones gráficas; en la estadística inferencial los gráficos guían al estudiante a construir otros objetos matemáticos en busca de una conclusión sobre la población.

Referencias

- Barreto-Villanueva, A. (2012). *El progreso de la estadística y su utilidad en la evaluación del desarrollo*. Papeles de Poblacion.
- Batanero, C. (2001). *Didáctica de la Estadística*. ISBN 84-699-4295-6. doi:10.13140/RG.2.1.3946.7044.
- Batanero, C. y Romero, L. (1995). *Aleatoriedad, sus significados e implicaciones educativas*. Uno.
- Díaz, C., Contreras, J. M., Batanero, C., y Roa, R. (2012). *Evaluación de Sesgos en el Razonamiento sobre Probabilidad Condicional en Futuros Profesores de Educación Secundaria*. Bolema - Mathematics Education Bulletin. doi:10.1590/S0103-636X2012000400006.
- Godino, J. y Batanero, C. (2007). *Un enfoque ontosemiótico del conocimiento y la instrucción matemática*.
- Hacking, I. (1995). *La tesis de Cassirer*. En *La domesticación del azar. La erosión del determinismo y el nacimiento de las ciencias del caos*.
- López Martín, M., Díaz-Levicoy, D., Gea, M., y Arteaga, P. (2016). *Estudio empírico de los problemas sobre intervalos de confianza en las pruebas de acceso a la universidad*. pages 81–92.
- Mallart Solaz, A. (2014). *La resolución de problemas en la prueba de Matemáticas de acceso a la universidad: procesos y errores*. Educatio Siglo XXI. doi:10.6018/j/194171.
- Mendez, H. (1991). *Understanding the Central Limit Theorem*. University of California, Santa Barbara.
- Polack, M. E. (2019). *Desigualdad en matemática entre alumnos de la ciudad y el conurbano*. Diario La Nación, Argentina.