

USO DEL RITUAL DE LA SIGNIFICANCIA ESTADÍSTICA Y SU IMPACTO SOBRE EL APRENDIZAJE DE LA MISMA

¹Javier Alonso Trujillo
alonsot1212@yahoo.com.mx

Recibido: 10/01/2017
Aprobado: 13/06/2017

²Leticia Cuevas Guajardo

³Abraham Alonso Ricardez

1. Doctor en Educación. Profesor titular "B" TC, Facultad de Estudios Superiores Iztacala, UNAM. México.
2. Doctora en Educación. Profesor titular "C" TC, Facultad de Estudios Superiores Iztacala, UNAM. México.
3. Profesor de asignatura "A". Facultad de Estudios Superiores Iztacala, UNAM. México.

Para citar este artículo:

Alonso J, Cuevas L, Alonso A. Uso del ritual de la significancia estadística y su impacto sobre el aprendizaje de la misma. Cuidarte. 2017; 6(12): 16-29.
DOI: <http://dx.doi.org/10.22201/fesi.23958979e.2017.6.12.69127>

ARTÍCULO ORIGINAL

.....
REVISTA ELECTRÓNICA DE INVESTIGACIÓN EN ENFERMERÍA FESI-UNAM.
Se autoriza la reproducción parcial o total de este contenido,
sólo citando la fuente y su dirección electrónica completas.

USO DEL RITUAL DE LA SIGNIFICANCIA ESTADÍSTICA Y SU IMPACTO SOBRE EL APRENDIZAJE DE LA MISMA

THE STATISTICAL SIGNIFICANCE PROCESS AND ITS IMPACT ON THE TEACHING OF STATISTICS

RESUMEN

El uso de pruebas de hipótesis en investigación de Enfermería y en general en el área de la salud, es una práctica que desde hace más de 100 años ha fortalecido la presentación de evidencias, así como el rigor metodológico y pensamiento crítico de los investigadores. Objetivo: Asociar el manejo del formato del ritual de la significancia estadística (FRSE) y aprendizaje de la Estadística. Metodología. Estudio cuasi-experimental, longitudinal, prospectivo y analítico. Intervención educativa: Se implementó el FRSE para la resolución de tres ejercicios de entrenamiento y dos cuestionarios. Tamaño muestral de 36 sujetos seleccionados según criterio del investigador. Resultados. 13.9% tuvo problemas para plantear hipótesis. 80.5% comprendió lo que es el nivel de significancia estadística. 8.3% de los sujetos tuvo problema para elegir la prueba de hipótesis. 77.7% dijo leer correctamente el "p" valor y el 8.3% dijo haber tenido problemas para tomar decisiones estadísticas. Las calificaciones medias de los ejercicios fueron 7.2 ± 2.2 , 7.9 ± 2.0 y 7.3 ± 3.0 . Las calificaciones medias de los cuestionarios fueron 5.7 ± 2.0 y 7.5 ± 2.5 ($p = 0.001$). No se demostró que exista asociación entre el manejo del formato y el rendimiento académico ($p = 0.087$). Conclusión. El manejo del formato es independiente del aprendizaje de la estadística. El aprendizaje de la estadística se observa cuando los estudiantes trabajan en equipo no es así cuando el trabajo es individual y los ejercicios son complejos.

Palabras clave. Aprendizaje, Estadística, Pruebas de hipótesis.

ABSTRACT

The use of statistical hypothesis testing in nursing research—and, more broadly, in the field of health—is a practice that has, for over one hundred years, strengthened the reliability of evidence as well as the methodological rigor and critical thinking of researchers. Goal: To associate mastery of the statistical significance process with the teaching of statistics. Methodology: Quasi-experimental, prospective, and analytical cross-study. The statistical significance process was applied, involving three training exercises and two questionnaires. The sample consisted of 36 subjects selected according to the researcher's criteria. Results: 13.9% of subjects faced problems in expressing a hypothesis. 80.5% understood the notion of statistical significance. 8.3% of subjects faced difficulties in selecting the statistical hypothesis test. 77.7% claimed to read the p value correctly, and 8.3% claimed to have had trouble making statistical decisions. Mean scores on the exercises were 7.2 ± 2.2 ; 7.9 ± 2.0 ; and 7.3 ± 3.0 . Mean scores on the questionnaires were 5.7 ± 2.0 and 7.5 ± 2.5 ($p = 0.001$). It was not proven that there exists an association between mastery of the statistical significance process and academic success ($p = 0.087$). Conclusion: Mastering the process and learning statistics occur independently. Statistical learning is observed when students work in teams—rather than individually—on complex exercises.

Keywords: learning, statistics, statistical hypothesis testing.

DOI: <http://dx.doi.org/10.22201/fesi.23958979e.2017.6.12.69127>

INTRODUCCIÓN

Los elementos lógicos que dieron origen a las pruebas de hipótesis fueron presentados en artículos científicos a principios del siglo XVIII. Sin embargo, formalmente las pruebas de hipótesis surgen en las décadas de 1920 y 1930 como resultado del trabajo de dos grupos o escuelas de pensamiento: por un lado, Ronald Fisher (1890-1962), y por el otro, Jerzy Neyman (1894-1981) y Egon Pearson (1895-1980), es decir, son ya casi 100 años que estas pruebas son utilizadas para fortalecer la presentación de evidencias, así como el rigor metodológico y pensamiento crítico de los investigadores¹.

En la carrera de Enfermería, es frecuente observar una problemática común a otras disciplinas. Los estudiantes han señalado que para ellos, el aprendizaje de la Estadística es irrelevante, poco atractivo y difícil. Según algunos estudios, su actitud hacia la Estadística en general es negativa^{2,3}. Muchos se preguntan: "si voy a ser Enfermero (a), ¿para qué quiero Estadística? La respuesta frecuente es "la tenemos que estudiar solo para recibirnos" ⁴. La experiencia docente, ha mostrado durante varios años, que por ejemplo, la comprensión del "p" valor y la interpretación contextual de las salidas del software SPSS, son un nodo problemático en la formación científica de los estudiantes de Enfermería. Ante esta situación, el docente debe reflexionar acerca de su práctica e implementar estrategias que faciliten el aprendizaje de aquellos temas y asignaturas que presentan rechazo, desinterés y una actitud negativa⁵. De hecho, se ha planteado que una actitud favorable hacia la Estadística se relaciona directa y positivamente con el rendimiento académico, aunque existen evidencias en las que dicha relación no ha podido ser demostrada^{3,6}. En este estudio se implementó como estrategia didáctica, el formato del ritual de la significancia estadística (FRSE). Esta estrategia pretendió cubrir dos necesidades; mejorar la comprensión del significado del "p" valor (concepto

de por sí polémico como lo han señalado algunos estadísticos) y mejorar la formación científica de los estudiantes de la carrera de Enfermería⁷. Según Silva, la única función que cumplen las pruebas de hipótesis, es la de valorar si existe o no suficiente evidencia muestral como para rechazar la validez de cierta conjetura; la llamada hipótesis nula, frecuentemente denotada como H_0 ⁸. Si se logra una adecuada comprensión del significado del "p" valor, podrá hacerse una adecuada lectura y aplicación del "p" valor durante la resolución de problemas de investigación⁹.

Según Rivera, la solución de problemas, toma de decisiones, interpretación de resultados y dominio de herramientas computacionales (como los paquetes estadísticos tipo SPSS), son sólo algunas de las principales competencias en investigación que los estudiantes de cualquier licenciatura deben de poseer^{10,11}. Orellana y Sanhueza señalan que un estudiante de pregrado de Enfermería, debe tener dominio técnico especializado en aspectos metodológicos de investigación, en este sentido se señala, como ya se mencionó, el manejo de paquetes estadísticos (software) que le permitan obtener información acerca del comportamiento de los datos obtenidos durante el trabajo de campo en una investigación científica¹².

Pues bien, la información de salida que arroja el software estadístico, por ejemplo SPSS, Excel y Stats entre otros, requiere de conocimientos teóricos acerca de conceptos fundamentales de Estadística. Uno de ellos es el "p" valor, concepto que no pocas veces causa malestar y confusión en los estudiantes en virtud de que tiene como base teórica a la probabilidad, rama de las matemáticas que permite predecir la ocurrencia de un evento⁷. En este sentido, un evento, en caso de ser posible su ocurrencia, puede estudiarse en función de la probabilidad que tiene para que éste ocurra. Es aquí donde el FRSE, desempeña un papel fundamental en el aprendizaje de pruebas de hipótesis y de la aplicación del "p" valor.

DOI: <http://dx.doi.org/10.22201/fesi.23958979e.2017.6.12.69127>

Con fines didácticos, el Ritual de la significancia estadística ha sido "encapsulado" dentro de un formato que permite desarrollar, paso a paso, el análisis de la información de salida que proporciona el software estadístico. Pero cuidado, cuando un software estadístico concluye su tarea, muchos estudiantes e incluso investigadores consolidados, creen que en ese momento ha terminado también la suya. Se trata de un grave error, es precisamente en ese momento donde comienza la función más importante del investigador, es decir, el análisis e interpretación contextual de la información de salida arrojada por el software¹.

El FRSE es una hoja dividida en 5 secciones que permiten llevar de la mano al estudiante que pretende tomar la mejor decisión acerca de aceptar o rechazar su hipótesis de investigación, aunque desde el punto de vista teórico, es la hipótesis nula la que se pone a prueba¹³. Tómese en cuenta que los elementos lógicos del FRSE resultan de la fusión de dos posiciones filosóficas; por un lado, Ronald Fisher y por el otro Jerzy Neyman y Egon Pearson, situación que desde el origen de la aplicación de las pruebas de hipótesis, ha conducido a diversas dificultades para su aplicación e interpretación¹.

La primera sección se denomina "Planteamiento de hipótesis", en ella que se escriben la hipótesis nula (H_0) y la alterna (H_1) o del investigador.

La segunda sección es el "Nivel de significancia". Aquí se escribe el porcentaje de error que el investigador está dispuesto a "tolerar" para rechazar H_0 y aceptar H_1 . La tercera sección la constituye un menú de pruebas de hipótesis entre las cuales el alumno debe distinguir cuál de ellas es la que está aplicando a su análisis. Se espera, por supuesto, que el profesor haya hecho una descripción lo suficientemente detallada del uso, aplicación y requisitos para cada una de ellas. La cuarta sección, quizá la de mayor dificultad, es la redacción del enunciado que permite describir la salida que el software estadístico arroja. La lectura

del "p" valor es con frecuencia errática para los estudiantes cuando se están familiarizando con el FRSE. Se han propuesto varias redacciones para esta sección.

Por último, la quinta sección ha sido nombrada "Toma de decisiones". La correcta toma de decisiones consiste en la comparación entre el nivel de significancia establecido en la segunda sección del FRSE y la lectura del "p" valor arrojado por el software. La simple comparación entre un valor y otro permite rechazar o no rechazar H_0 .

No obstante lo anterior, la interpretación de la toma de decisiones no es sencilla, pues el investigador debe contextualizar sus resultados estadísticos con la información descrita en la investigación, en el planteamiento del problema y en su hipótesis, pues en muchas ocasiones, como lo señala Silva, la aplicación de rituales "va generando así una suerte de costumbre a lo largo de cuya configuración mecánica se va supliendo al pensamiento crítico que supuestamente lo originó"¹⁴⁻¹⁹.

Por lo anterior, la problemática que da origen a este trabajo es la necesidad de mejorar la formación científica de los estudiantes de Enfermería, incorporando una estrategia didáctica "conductista" hasta cierto grado, ligada al desarrollo de un pensamiento constructivista. Se espera mejorar el aprendizaje de la Estadística, lo cual se reflejará en la comprensión del concepto "p" valor, la resolución de problemas y la interpretación de los resultados arrojados por el software estadístico. El mejor indicador del impacto de esta estrategia es el aprendizaje de la Estadística en estudiantes que cursan el módulo de Metodología de la Investigación.

El objetivo de este trabajo fue asociar el manejo del FRSE y aprendizaje de la Estadística de estudiantes de la carrera de Enfermería.

DOI: <http://dx.doi.org/10.22201/fesi.23958979e.2017.6.12.69127>

METODOLOGÍA

Estudio cuasi-experimental, longitudinal, prospectivo y analítico que se ubica en el nivel relacional según taxonomía de Supo¹³.

Diagrama del diseño de investigación:

G IE O1 O2 O3 O4 O5 O6 O7 >> "X" → "Y"

Diseño según la nomenclatura propuesta por Campelly Stanley²⁰. El grupo experimental (G) estuvo integrado por 36 sujetos sin asignación aleatoria. Se aplicó la intervención educativa (IE) que consistió en impartir clase sobre fundamentos estadísticos y manejo del FRSE. O1 fue la entrevista sobre el manejo del FRSE. O2, O3 y O4 son las evaluaciones de los ejercicios realizados en equipo. O5 y O6, representan la aplicación de los cuestionarios 1 y 2 resueltos individualmente. O7 es el rendimiento académico (concepto que incluye el aprendizaje de la Estadística aplicada a problemas concretos de forma integral). "X" representa los resultados de la entrevista acerca del manejo que dijeron tener los estudiantes sobre el FRSE. "Y" representa al conjunto de Evaluaciones de Ejercicios, cuestionarios y rendimiento académico.

La unidad de estudio estuvo representada por cada estudiante de la carrera de Enfermería matriculado en el módulo de Metodología de la Investigación en la Licenciatura en Enfermería que imparte la Facultad de Estudios Superiores Iztacala, durante el semestre 2016-1. Se requirió que las unidades de estudio hubieran recibido la IE. Para lograr lo anterior, se impartió teoría en estadística, se realizaron tres ejercicios de entrenamiento estadística inferencial (Pruebas de asociación, de comparación de medias y de independencia condicional), y se practicó el uso del software estadístico SPSS versión 20. La evaluación del impacto que tuvo la estrategia didáctica se fundamentó en la demostración de

competencias cognitivas y procedimentales, a través de la resolución de dos cuestionarios durante el semestre escolar 2016-1. El tamaño de la muestra fue de 36 sujetos seleccionados según criterio del investigador. Este estudio se ubica dentro de los lineamientos que rigen la investigación-acción que según Elliot, consiste en la "reflexión relacionada con el diagnóstico" ⁵. El aspecto conductista de la intervención educativa radica en el hecho de que se aplica un estímulo y se espera una respuesta, en este caso, una toma de decisión basada en la comparación simple del "p" valor obtenido contra el nivel de significancia. El aspecto constructivista consistió en que, a partir de un contexto que situaba la problematización de ejercicios y cuestionarios, el alumno interpretaba y contextualizaba la salida arrojada por el FRSE y la del software estadístico, lo que favorece la construcción del conocimiento partiendo de situaciones significativas para el estudiante, quien debe asumir un papel activo en su propio aprendizaje pues interactúa con el contexto de la problematización, la cual no le es ajena dado que se relaciona con la práctica de Enfermería. Como lo señalan algunos autores, la cognición situada transmite la idea de que el conocimiento está anclado y conectado con el contexto en el que el conocimiento se construyó^{16,17,21,22}.

Diseño del instrumento. Se procedió al diseño y validación de un instrumento de medición documental de la variable "Manejo del FRSE". La técnica de recolección de datos fue a través de una encuesta en la que los participantes opinaron acerca del grado de manejo que tenían del FRSE. La validación de contenido se basó en la propuesta de Fisher, Neyman y Pearson y el análisis hecho por Prieto y Herranz, así como también se contó con la opinión de profesores expertos en Estadística¹⁷.

El instrumento presentó varianza diferente de cero en cada uno de sus ítems. La correlación ítem-total arrojó correlaciones positivas y significativas en todos los casos (Tabla 1). La confiabilidad fue

DOI: <http://dx.doi.org/10.22201/fesi.23958979e.2017.6.12.69127>

determinada por Alfa de Cronbach obteniendo en la prueba piloto ejecutada con diez sujetos un valor de 0.879

Procedimientos. Se impartieron los contenidos relacionados con Estadística inferencial procurando que los ejercicios de entrenamiento de la aplicación del "p" valor, tuvieran una estrecha relación con la práctica de Enfermería, lo que algunos autores denominan "Aprendizaje experiencial"^{16,17,19}.

Tabla 1. Correlaciones con coeficiente de Spearman entre cada ítem y el total del instrumento. Todos los ítems correlacionaron positiva y significativamente.

Instrumento	r_s	p
Ítem 1	.423	.010
Ítem 2	.614	.000
Ítem 3	.545	.001
Ítem 4	.573	.000
Ítem 5	.440	.007
Ítem 6	.561	.000
Ítem 7	.739	.000
Ítem 8	.611	.000
Ítem 9	.410	.013
Ítem 10	.507	.002

Fuente: Trabajo de campo. FES Iztacala, UNAM, 2016.

Se aplicaron y evaluaron dos cuestionarios teórico-prácticos durante el semestre 2016-1. La variable rendimiento académico fue definida como "Nivel académico alcanzado por un estudiante y que es cuantificado por dos evaluaciones sumativas (cuestionarios 1 y 2) realizadas durante el semestre lectivo 2016-1. Cada una de las evaluaciones incluyó aspectos teórico-metodológicos de tipo estadístico y manejo del Software Microsoft Excel 2010 y SPSS versión 20. Todas las evaluaciones estuvieron sustentadas por los contenidos temáticos del

Programa de Metodología de la Investigación de la licenciatura en Enfermería de acuerdo al Plan de estudios vigente en 2016²³.

Aspectos éticos. Los sujetos participantes fueron invitados a participar y aceptaron a responder a la encuesta en la que opinaron acerca del manejo que tenían del FRSE.

Plan de análisis estadístico. Con los datos obtenidos se procedió a obtener las frecuencias importantes para cada sección del FRSE. Con el puntaje total del instrumento de opinión se clasificó a los sujetos de acuerdo al manejo que dijeron tener del FRSE, obteniéndose las categorías ordinales "Bueno", "Muy bueno" y "Excelente". Se realizaron correlaciones bivariadas no paramétricas (correlación de Spearman) entre el grado de manejo del FRSE, las calificaciones de los tres ejercicios de entrenamiento y los dos cuestionarios de integración y evaluación sumativa. Se comprobó el supuesto de normalidad de los datos con la prueba de Shapiro-Wilk. Se compararon las calificaciones medias del cuestionario 1 y 2 de integración utilizando la prueba de Mann Whitney y de los ejercicios de entrenamiento con Kruskal-Wallis. El rendimiento académico dicotomizado se asoció con el manejo de FRSE dicotomizado utilizando Chi cuadrada de independencia. El nivel de significancia fue establecido ≤ 0.05

Los procedimientos estadísticos descriptivos e inferenciales se ejecutaron con el software Excel y SPSS versión 22.

RESULTADOS

Manejo del FRSE (O1). El 100 % de los sujetos dijo haber utilizado el FRSE para resolver sus ejercicios y cuestionarios.

A continuación se describen los resultados obtenidos después de la aplicación del instrumento,

DOI: <http://dx.doi.org/10.22201/fesi.23958979e.2017.6.12.69127>

con el cual evaluó la opinión de los estudiantes respecto a la lo difícil que les resultó desarrollar las secciones que integran el FRSE. Véase la figura 1.

Sección 1. Plantear hipótesis: El 13.9% de los estudiantes les resultó frecuentemente difícil plantear la hipótesis nula y la hipótesis alterna. Este es el punto de partida para el correcto desarrollo de una prueba de hipótesis.

Sección 2. Entender el concepto "Nivel de significancia estadística": El 80.5% dijo haber entendido frecuentemente el concepto "significancia estadística", paso fundamental para la resolución de los ejercicios y cuestionarios. En esta sección del FRSE, sólo el 2.8% dijo nunca haber entendido el significado concepto, por lo que no tenían la certeza de la razón por la cual en FRSE se establece un nivel de significancia ≤ 0.05 .

Sección 3. Selección de la prueba de hipótesis: Al 8.3% de la muestra. Le resultó frecuentemente difícil elegir correctamente la prueba de hipótesis que debería aplicar en cada ejercicio de entrenamiento y

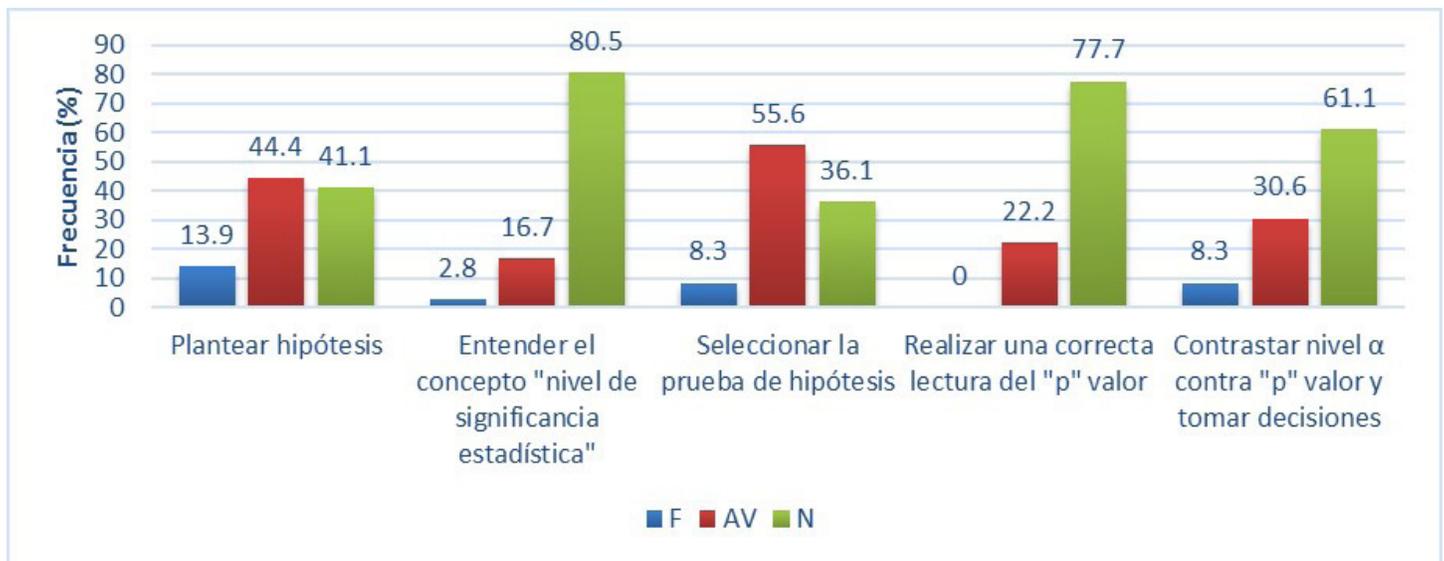
en los cuestionarios de integración. Al respecto, cerca de una tercera parte de la muestra (36.1 %) nunca tuvo problemas para elegir la prueba estadística que correspondía a cada ejercicio y cuestionarios.

Sección 4. Realizar una correcta lectura del "p" valor: El 77.7% dijeron que nunca les fue difícil haber hecho un lectura correcta del "p" valor obtenido con el software. De hecho, a nadie le resultó frecuentemente difícil realizar esta lectura.

Sección 5. Contrastar nivel α contra "p" valor y tomar decisiones: Solo el 8.3% dijo haber tenido problemas frecuentemente para tomar decisiones, último paso del FRSE.

Sin embargo, hay que considerar que el 30.6% de los estudiantes señalaron que algunas veces se les dificultó tomar la decisión correcta para aceptar o rechazar la hipótesis nula. Esto deja ver que en realidad a un tercio aproximadamente de la muestra, les resulta difícil comparar el "p" valor contra el nivel de significancia y luego tomar la decisión correcta acerca de cuál de las hipótesis rechazar o

Figura 1. Encuesta acerca de la opinión de los estudiantes respecto a la frecuencia con la les resultó difícil desarrollar las secciones del FRSE.



F= Frecuentemente AV= Algunas veces N=Nunca
Fuente: Trabajo de campo. FES Iztacala, UNAM, 2016-1.

DOI: <http://dx.doi.org/10.22201/fesi.23958979e.2017.6.12.69127>

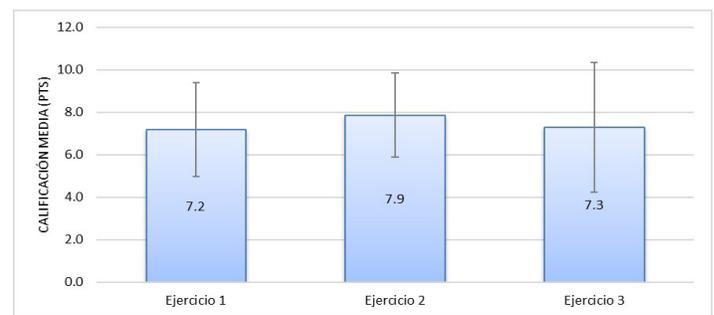
no rechazar.

El resultado de la encuesta permitió sumar los puntos obtenidos en los ítems que integraron al instrumento de medición por cada sujeto, lo cual representa, la intensidad de la perspectiva que tienen los estudiantes acerca de la dificultad o facilidad para resolver el FRSE. Los puntajes obtenidos permitieron a su vez, clasificar en cinco categorías a los participantes, a saber, "Manejo pésimo" con ningún caso, "Manejo suficiente" con ningún caso, "Manejo bueno" con el 27.8% de los sujetos, 52.8% un "Manejo muy bueno" y el 19.4% un "Manejo excelente". Véase figura 2.

Ejercicios de entrenamiento (02-04). Los ejercicios de entrenamiento (que fueron resueltos de manera colaborativa entre los miembros de los equipos de trabajo formados al inicio del semestre) fueron evaluados y comparadas sus calificaciones medias. Se observó que no se presentaron diferencias significativas entre las tres evaluaciones realizadas ($p = 0.329$). En el primer ejercicio la calificación media fue 7.2 ± 2.2 , el ejercicio 2 obtuvo 7.9 ± 2.0 y el ejercicio 3 obtuvo 7.3 ± 3.0 , por lo que puede señalarse que el grado de dificultad para resolver los ejercicios de entrenamiento fue homogéneo y las calificaciones revelan un rendimiento apenas regular. Los ejercicios tuvieron un grado de dificultad

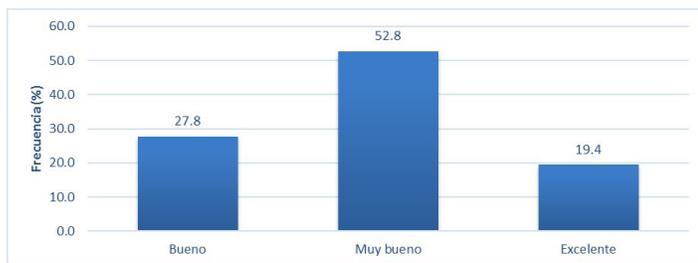
ascendente, ya que se consideró que el ejercicio 1 de baja complejidad, el ejercicio 2 de mediana complejidad y el ejercicio 3 de elevada complejidad, no obstante el trabajo colaborativo para haber influido en el rendimiento de los estudiantes. Véase figura 3.

Figura 3. Calificación promedio para cada uno de los ejercicios de entrenamiento (Pruebas de asociación, comparación de medias e independencia condicional). La prueba de Kruskal-Wallis demostró que no existe diferencia significativa entre las tres evaluaciones ($p = 0.329$). Media \pm Desviación Estándar.



Fuente: Trabajo de campo. FES Iztacala, UNAM, 2016-1.

Figura 2. Manejo del FRSE según la opinión de los estudiantes observados. Nótese que no hubo casos de "Manejo pésimo" ni "manejo suficiente".



Fuente: Trabajo de campo. FES Iztacala, UNAM, 2016-1.

Recuérdese que se exploraron tres tipos de pruebas de hipótesis; Pruebas de asociación, pruebas de comparación de medias y pruebas de independencia condicional. Al correlacionar las calificaciones de los estudiantes con el grado de manejo del FRSE, se observó que en el caso de las pruebas de asociación y la comparación de medias, la correlación fue significativa ($p = 0.006$ y $p = 0.030$ respectivamente) como se puede apreciar en la tabla 2. El ejercicio de independencia condicional ya no tuvo correlación con el grado de manejo del FRSE.

Cuestionarios de integración (05-06). En la tabla 2, se observa que existe correlación significativa ($p = 0.003$) entre el resolver un cuestionario alta complejidad que incluya pruebas de independencia condicional y el grado de manejo del FRSE, sin embargo tal correlación desaparece

DOI: <http://dx.doi.org/10.22201/fesi.23958979e.2017.6.12.69127>

Tabla 2. Correlación entre el grado de manejo del FRSE (Bueno, Muy bueno y Excelente) con respecto a la calificación obtenida en los ejercicios de entrenamiento y cuestionarios de integración.

Pruebas	Ejercicio 1 Asociación	Ejercicio 2 Comparación de medias	Ejercicio 3 Independencia condicional	Cuestionario 1 Independencia condicional	Cuestionario 2 Asociación y comparación de medias	Rendimiento académico (calificación final)
<i>r_s</i>	.452	.362	.186	.483	.263	.289
<i>p</i>	.006	.030	.277	.003	.121	.087

Fuente: Trabajo de campo. FES Iztacala, UNAM, 2016-1.

($p = 0.121$) cuando se resuelven cuestionarios de complejidad moderada en los que se incluyen pruebas de asociación y comparación de medias.

Estos cuestionarios (1 y 2) junto con otros aspectos, fueron considerados para la evaluación sumativa del semestre, por lo que los estudiantes ya no podían trabajar en equipo al intentar resolverlos, como lo hicieron al realizar los ejercicios de entrenamiento. Para la resolución de cuestionarios, se vigiló que los estudiantes no se comunicaran entre ellos y demostrarán así sus competencias cognitivas y procedimentales de forma individual, pues cada alumno contó con una computadora y el mismo software (SPSS y Excel) para resolverlos.

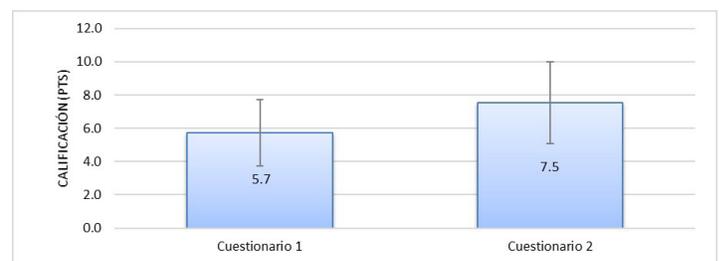
La calificación media en el cuestionario que exploró sus competencias para resolver problemas en los que se incluyeron pruebas de independencia condicional (con alto grado de complejidad) fue de 5.7 ± 2.0 , mientras que cuando se evaluó asociación y comparación de medias (complejidad moderada) fue de 7.5 ± 2.5 (Figura 4).

El análisis estadístico mostró que la calificación media del cuestionario 1 (que evaluó pruebas de independencia condicional), fue menor estadísticamente que la calificación media del cuestionario 2 ($p = 0.001$), lo que indica que, efectivamente, un cuestionario de alta complejidad

resulta más difícil de resolver, sin embargo, la calificación del cuestionario 1 si se correlacionó con el grado de manejo que el estudiante tuvo del FRSE.

Figura 4. Calificaciones medias obtenidas en los cuestionarios de integración. La comparación de medias mostró que existe diferencia significativas entre los dos cuestionarios ($p = 0.001$), resultando una mayor calificación media al resolver el cuestionario que incluyó pruebas de asociación y comparación de medias. Para el cuestionario que incluyó la prueba de independencia condicional, la calificación media fue de 5.7 ± 2.0 .

Media \pm Desviación Estándar.



Fuente: Trabajo de campo. FES Iztacala, UNAM, 2016-1.

Rendimiento académico (Aprendizaje de la Estadística) (O7).

El rendimiento académico fue el resultado de la

DOI: <http://dx.doi.org/10.22201/fesi.23958979e.2017.6.12.69127>

sumatoria de las evaluaciones de realizadas durante el semestre escolar, a saber, realización y exposición de protocolo de investigación (20%), realización y exposición de informe final de investigación (20%) y las calificaciones obtenidas en los cuestionarios 1 y 2 (60%) que se han señalado anteriormente. Aplicando el reglamento general de exámenes de la UNAM, la calificación mínima para un rendimiento académico aprobatorio es 60%, menos de esta calificación el rendimiento académico es no aprobatorio²⁴.

Asociación entre Manejo del FRSE y Rendimiento académico.

Como se mencionó en la metodología, el rendimiento académico representa tanto el aprendizaje de la Estadística aplicado a la resolución integral de problemas de investigación como a la calificación media de los cuestionarios 1 y 2 más otros factores, tales como la realización de un protocolo de investigación, su informe final y

exposición frente a grupo. El rendimiento académico es la calificación que representa el "Nivel académico alcanzado por un estudiante" durante el semestre lectivo y en especial, su aprendizaje de la Estadística como disciplina.

Dado que esta investigación buscaba demostrar que el manejo del FRSE mejoraba el rendimiento académico, se procedió a correlacionar y a buscar si había asociación entre ambas variables. No se pudo demostrar que a manejo excelente del FRSE correspondiera un rendimiento académico excelente ($p = 0.087$), tampoco se pudo demostrar que los alumnos que dijeron tener un manejo excelente del FRSE aprobaran el módulo de Metodología de la Investigación al finalizar el semestre ($p = 0.797$), considerando una muestra de 36 sujetos. Véanse tablas 2 y 3 respectivamente.

No obstante, ejercicios de complejidad baja y regular trabajados en forma colaborativa si se correlacionan con el manejo del FRSE, así como también la resolución del cuestionario 1 de alta complejidad.

Tabla 3. Grado de asociación entre el manejo del FRSE y el rendimiento académico. No se pudo demostrar con una muestra de 36 estudiantes, que exista asociación entre el decir tener un excelente manejo del FRSE y un mejor aprendizaje de la Estadística aplicada a la resolución de problemas de forma integral al final del semestre ($p = 0.797$)

Categorías		Rendimiento académico		Total
		Aprobado	No Aprobado	
Manejo del FRSE	Excelente	3	4	
	Bueno a Muy Bueno	14	15	
Total		17	19	

Fuente: Trabajo de campo. FES Iztacala, UNAM, 2016-1.

DISCUSIÓN

El 100% de los sujetos utilizaron el FRSE, en virtud de que precisamente esta intervención educativa es la diferencia, respecto a los cursos de metodología de la investigación, que clásicamente se imparten en la institución en la cual se realizó el estudio. Lo anterior es una condición necesaria para el análisis de este trabajo.

El FRSE tiene cinco secciones, que en teoría y desde principios del siglo XX, tiene la función de apoyar la comprensión de las pruebas de significancia estadística tal y como alguna vez lo planteó Fisher, Neyman y Pearson, pero que ha sido severamente criticado por Silva y ampliamente analizado por Prieto y Herranz^{1,7}. Es decir, si bien en un principio el FRSE tiene tintes conductistas, la transición hacia el aspecto constructivista y sobre todo al aprendizaje situado, depende en alto grado, no de la opinión que el estudiante tenga de sí mismo (postura subjetiva), sino de las competencias que para investigar haya desarrollado durante su carrera o en el peor de los casos, durante el semestre escolar en el que aprende específicamente a investigar. Algunas de esas competencias sin duda son el trabajo en equipo, pensamiento crítico, manejo del software estadístico, comprensión de teorías y de manera muy especial, interpretación contextualizada de los resultados^{10,11}.

La figura 1 debe leerse con mucho cuidado, pues es pieza clave en este análisis. En ella se destaca que la sección con mayor dificultad fue el planteamiento de las hipótesis nula y alternativa. Esto puede tener su explicación en el hecho de que el alumno debe de transformar la hipótesis de investigación (contextualizada desde el planteamiento del problema) en hipótesis estadísticas, situación que probablemente resulta complejo por el cambio conceptual que debe de realizarse.

En contraste, la sección del FRSE que resultó "más fácil" de resolver fue la lectura del "p" valor,

pues ningún estudiante dijo haber tenido dificultad para resolver este paso, no obstante que es un momento crucial y altamente complejo, pues la lectura del "p" valor representa la probabilidad de rechazar o no rechazar la hipótesis nula y pensar en cuál es la hipótesis que entonces resulta más probable de ser cierta. Una posible explicación puede ser que el trabajo en equipo ayuda a resolver este paso del FRSE y además se cuenta con la asesoría del docente durante el desarrollo de los ejercicios de entrenamiento, pero obsérvese que cuando los alumnos trabajan solos y bajo la presión de la evaluación sumativa, solo los alumnos que realmente tuvieron un "Manejo excelente" del FRSE obtuvieron las calificaciones más altas en el cuestionario que presentaba alta complejidad (Pruebas de independencia condicional).

Es así como, probablemente, el profesor pensaría que el FRSE es un elemento conductista y hasta cierto punto memorístico del aprendizaje, cuando en realidad, tan sólo para iniciarlo y luego desarrollarlo, se requieren de competencias cognitivas especializadas. También es necesario señalar, que la toma de decisiones resulta frecuentemente y algunas veces difícil en términos generales para el 38.9% de los participantes. Con este nivel de desarrollo del FRSE, es claro que existe una sobrestimación de los alumnos, quienes por sus respuestas, pueden ser clasificados como sujetos que manejan bien, muy bien y excelentemente bien el FRSE. Los resultados de los ejercicios de entrenamiento podrían ser compatibles con esta auto-apreciación, pero considérese que para resolverlos, los estudiantes trabajaron en equipo y sin la presión que significa saber que sólo son ejercicios de entrenamiento sin valor sumativo en su rendimiento académico.

Debe señalarse también, que los cuestionarios 1 y 2 presentaron distinto grado de dificultad. Por un lado, se resolvían individualmente sin opción a consultar apuntes ni a sus compañeros de clase ni

DOI: <http://dx.doi.org/10.22201/fesi.23958979e.2017.6.12.69127>

al profesor. Por otra parte, tenían carácter sumativo para su evaluación del rendimiento académico y además diverso grado de dificultad, siendo relativamente más difícil el cuestionario 1 respecto al 2, situación que se observó claramente en la calificación media, donde se pudo comprobar que la calificación del cuestionario 1 tuvo una calificación media de 5.7 ± 2.0 y la del cuestionario 2 fue de 7.5 ± 2.5 ($p = 0.001$) Véase figura 4.

Finalmente, podemos preguntarnos ¿Por qué razón no existe asociación entre el excelente manejo del FRSE y el aprendizaje de la Estadística o rendimiento académico? ¿Es acaso que el FRSE no contribuye a mejorar el aprendizaje estadístico?

No se puede afirmar que un excelente manejo del FRSE resulte en un mejoramiento del aprendizaje estadístico, es decir, dominar de memoria el FRSE no garantiza que el estudiante pueda resolver problemas de investigación utilizando herramientas estadísticas, tanto descriptivas como inferenciales. La explicación puede residir en que el FRSE únicamente apoya las competencias procedimentales, pero no las cognitivas, es más, supone que el estudiante posee competencias teóricas sobre el planteamiento de sus hipótesis, conocimientos sobre pruebas estadísticas, nivel de significancia y comprensión clara del "p" valor y su lectura, pues estos elementos permiten tomar decisiones adecuadas. En consecuencia, el excelente manejo del FRSE (según la encuesta aplicada a los estudiantes) no garantiza un rendimiento académico aceptable, dado que todo apunta a que los estudiantes fueron sobrevalorados con el instrumento que daba cuenta de su opinión acerca del manejo del FRSE y el profesorado supuso que el nivel de competencias cognitivas era suficiente para la aplicación del FRSE. A la explicación anterior, habrá que sumarle la presión psicológica que implica la evaluación sumativa sin el apoyo del trabajo colaborativo ni la consulta de apuntes u otro recurso que si podían utilizar en los ejercicios de entrenamiento²⁵⁻²⁷. No puede descartarse que para la comprensión

de las pruebas de hipótesis, se requiere integrar conceptos como población, muestra, estadístico de prueba, distribución muestral del estadístico de prueba, nivel de significancia, hipótesis nula, hipótesis alternativa, "p" valor, entre otros. Inzunza y Jiménez demostraron que las pruebas de hipótesis son un concepto complejo para los estudiantes universitarios, incluso para los de la carrera de Matemáticas¹.

CONCLUSIONES

La sociedad actual necesita ciudadanos estadísticamente cultos. Los estudiantes universitarios y sus profesores deben incorporar competencias investigativas relacionadas con el manejo de la Estadística.

Nuestros resultados permiten concluir que la valoración de los estudiantes al autocalificarse como sujetos con manejo excelente del FRSE, sólo se correlacionó significativamente con una buena calificación en el cuestionario 1 (alta complejidad), pero no con la calificación del cuestionario 2 (moderada complejidad). No existe asociación entre excelente manejo de FRSE y un mejor aprendizaje de la Estadística aplicado a la resolución de problemas de manera integral, dado que ambas variables dependen de competencias no sólo procedimentales sino también cognitivas, las cuales se dan por hecho cuando el estudiante usa el FRSE.

Las buenas notas se obtienen cuando los estudiantes trabajan en equipo y tienen la oportunidad de consultar diversos apoyos didácticos, no obstante, cuando resuelven cuestionarios de manera individual y con la presión psicológica que implica la evaluación sumativa, el rendimiento académico puede descender hasta niveles no aprobatorios.

El FRSE resultó útil en la resolución de los cuestionarios de integración de alta complejidad

DOI: <http://dx.doi.org/10.22201/fesi.23958979e.2017.6.12.69127>

(independencia condicional), a pesar de que en esta categoría, hubo un promedio no aprobatorio (5.7 ± 2.0), es decir, los estudiantes que aprobaron el cuestionario de alta complejidad con 6 o poco más de 7 de calificación, simultáneamente dijeron haber tenido un excelente manejo del formato.

El FRSE está lejos de ser una aplicación didáctica conductista en sí misma, pues ha quedado claro, que requiere necesariamente la incorporación de competencias cognitivas especializadas que permitan al alumno comprender las cinco secciones que lo integran, como por ejemplo, se requiere haber comprendido y saber integrar conceptos como población, muestra, hipótesis nula, hipótesis alternativa, nivel de significancia, "p" valor, criterios de rechazo y no rechazo de hipótesis y tener la capacidad de realizar una interpretación contextual de resultados.

Cuando se inició la aplicación del FRSE en nuestra práctica docente, se pensó que podría ser utilizado como simple receta para resolver problemas de investigación, sin embargo, esta investigación dejó claro que su uso requiere del desarrollo de un razonamiento estadístico especializado, que además traerá como consecuencia una cultura estadística, útil para entender informaciones técnicas en la vida cotidiana.

La recomendación que los autores de este trabajo hacen a los profesores de Metodología de la Investigación, Seminario de Tesis, Métodos cuantitativos o aquella que tenga incorporados contenidos de Estadística inferencial, es que se estudie perfectamente la teoría que subyace a los 5 pasos del FRSE, antes de su implementación y se enseñe a los estudiantes a razonar y comprender más que a memorizar.

Por último, los autores deseamos que esta línea de investigación sea abordada con la lógica implícita en los niveles de investigación, es decir, el aprendizaje de la estadística, la formación científica de los

estudiantes de Enfermería, debe abordarse según se avance en cada uno de los niveles investigativos, desde el exploratorio hasta el aplicativo. El diseño de investigación experimental deberá incorporar a un grupo de control y la asignación aleatoria, pues de éste modo, se podrá evidenciar, demostrar y probar que el FRSE es un factor que incide positivamente en el mejoramiento de la enseñanza y del aprendizaje de la Estadística, no sólo en Enfermería, sino en general en las ciencias de la salud y la educación

REFERENCIAS BIBLIOGRÁFICAS

1. Alonso TJ, Alonso RA, Valadez DD. Actitud hacia la investigación científica de estudiantes de Enfermería. *Rev Cuid El Arte Cuid.* 2015;4(7):24-35.
2. Alonso TJ. Actitud hacia la investigación y logro escolar durante el aprendizaje de la Estadística de estudiantes de Enfermería. *Rev Cuid El Arte Cuid.* 2015;4(8):8-16.
3. Ferreri N, Panzeri A, Santone B, Storza M, De Giorgio N, Elchaimi MM, et al. Enseñanza de la Estadística: Desde el ámbito laboral al aula. En Universidad Nacional de Rosario, Argentina; 1999. p. 117-21.
4. Elliott J. La investigación-acción en educación. Ediciones Morata; 1990. 340 p.
5. Pulido, JE. Enseñanza de la Estadística a partir de la actitud del alumno. *Laurus.* 2009;15(30):42-70.
6. Prieto-Valiente, L, Herranz-Tejedor, I. ¿Qué significa estadísticamente significativo? España: Ediciones Díaz de Santos.; 2010.
7. Silva Aycaguer, LC. Las pruebas de significación estadística: seis décadas de fuegos artificiales. *Rev Fac Nac Salud Pública.* 2016;34(3):372-9.
8. Fisher, RA. *Statistical methods for research workers.* London: Hafner Press; 1925. 166 p.
9. Rivera HM, Arango PL, Torres VK, Salgado BR, García GMF, Caña DL. Competencias para la investigación. Desarrollo de habilidades y conceptos. Trillas / Universidad Simón Bolívar; 2009. 159 p.
10. Quezada UR, Medina AG. Competencias para la investigación y Conectivismo. Una relación que ya no puede esperar. *CuidArte.* 29 de abril de 2014;3(5):60-8.

DOI: <http://dx.doi.org/10.22201/fesi.23958979e.2017.6.12.69127>

11. Orellana YA, Sanhueza AO. Competencias en Investigación en Enfermería. *Cienc Enferm.* 2011;17(2):9-17.
12. Inzunza-Cazarez, S, Jiménez-Ramírez, JV. Caracterización del razonamiento estadístico de estudiantes universitarios acerca de las pruebas de hipótesis. *Rev Latinoam Investig En Matemática Educ.* 2013;16(2):179-211.
13. Supo CJ. Seminario de investigación científica. [Internet]. Arequipa, Perú: Bioestadístico.com; 2014. Disponible en: Bioestadístico.com; 2014. Recuperado a partir de: <http://seminariosdeinvestigacion.com/sinopsis>
14. Silva Aycaguer, LC. Una Ceremonia Estadística para Identificar Factores de Riesgo. *Salud Colect.* diciembre de 2005;1(3):309-22.
15. Dorado Pereda, C. Aprender a aprender. estrategias y técnicas [Internet]. 1997 [citado 1 de mayo de 2014]. Disponible en: <http://www.xtec.cat/~cdorado/cdora1/esp/habcog.htm>
16. Alonso TJ, Valadez DD, Carrasco YM, Guzmán GAL. Aprendizaje experiencial aplicado al proceso enseñanza-aprendizaje de Estadística Inferencial. Experiencia docente. En Puebla, México: Facultad de Ciencias Físico Matemáticas de la Benemérita Universidad Autónoma de Puebla; 2016.
17. Díaz Barriga Arceo F. Enseñanza situada: Vínculo entre la escuela y la vida. México: McGraw-Hill; 2006. 191 p.
18. Amaya Guerra, J, Prado Maillard, E. Estrategias de aprendizaje para universitarios. Un enfoque constructivista. México: Trillas; 2002. 198 p.
19. Díaz Barriga AF, Hernández RG. Estrategias docentes para un aprendizaje significativo. Una interpretación constructivista. México: McGraw-Hill; 1998. 14 p.
20. Campbell, DT, Stanley, JC. Diseños experimentales y cuasiexperimentales en la investigación social. Buenos Aires: Amorrortu Editores; 1993.
21. Rodríguez EF. Aprendizaje experiencial, investigación-acción y creación organizacional de saber: la formación concebida como una zona de innovación profesional. *Rev Electrónica Interuniv Form Profr.* 2009;12(3):3.
22. Sánchez PR. Enseñar a investigar. Una didáctica nueva de la investigación en ciencias sociales y humanas. 2a ed. México: Plaza y Valdés Editores; 2010. 190 p. (Educación).
23. Lara BA, Alonso TJ, Pineda OJ, Ibarra ChL, López SG, Paredes BL. Programa del módulo de Metodología de la investigación en Enfermería. FES Iztacala, UNAM; 2006.
24. Universidad Nacional Autónoma de México. Reglamento General de Exámenes. [Internet]. UNAM; 2015. Disponible en: <http://www.ddu.unam.mx/index.php/reglamento-general-de-examenes>
25. Silva-Aycaguer LC. Límites y mitos de la estadística en la valoración de los resultados de los ensayos clínicos y otros estudios. *Aten Primaria.* 2007;39(2):12.
26. Silva-Aycaguer LC. Las pruebas de significación estadística: seis décadas de fuegos artificiales. *Rev Fac Nac Salud Pública.* 2016;34(3):372-9.
27. Heredia J, Carolina V. Los conflictos escolares en el aula durante el período de exámenes y el estrés en adolescentes de décimo de básica del Colegio Experimental «Lev Vygotsky» durante el año 2013 [Internet] [Thesis]. PUCE; 2013 [citado 18 de febrero de 2017]. Disponible en: <http://repositorio.puce.edu.ec/handle/22000/12361>