

DEPARTAMENTO DE DIDACTICA DE LAS MATEMÁTICAS



**LA ESTADÍSTICA COMO HERRAMIENTA EN LA
INVESTIGACIÓN PSICOLÓGICA: UN ESTUDIO
EXPLORATORIO**

Trabajo final de Máster

Cuauhtémoc Gerardo Pérez López

Directora: Dra. Carmen Batanero

Granada, Septiembre de 2010

PROYECTO SEJ2007-60110/EDUC, MEC, FEDER

PROYECTO EDU2010-14947 Ministerio de Ciencia e Innovación

Área Académica. Aprendizaje y Enseñanza en Ciencias, Humanidades y Artes,

Universidad Pedagógica Nacional, Unidad Ajusco, México.

Yani, Yutzi, Salvi

Gracias por dejarme usar parte del tiempo de ustedes en esta aventura, que ahora también es suya.

Introducción	1
1. El problema de investigación	3
1.1. Introducción	3
1.2. La estadística en la investigación psicológica y educativa	3
1.3. Formación estadística en los estudios de Educación y Psicología	5
1.3.1. Características generales de plan de estudios	9
1.3.2. Las tesis de licenciatura	10
1.4. Razonamiento estadístico y sus componentes	12
1.4.1. El ciclo de investigación	12
1.4.2. Modos fundamentales de razonamiento estadístico	15
1.4.3. Ciclo interrogativo	16
1.4.4. Disposiciones	17
1.5. Objetivos del trabajo	17
2. Antecedentes	21
2.1. Introducción	21
2.2. La controversia sobre la metodología estadística en Psicología	22
2.2.1. Introducción	22
2.2.2. Críticas al contraste de hipótesis	23
2.2.3. Recomendaciones para mejorar el uso de la estadística en Psicología	25
2.3. Errores en inferencia estadística	27
2.3.1. Introducción	27
2.3.2. Significancia estadística	27
2.3.3. Hipótesis y lógica del contraste	28
2.3.4. Intervalos de confianza	29
2.4. Análisis empíricos del uso de la estadística en la investigación	30
2.4.1. Introducción	30
2.4.2. Estudios en Didáctica de la Matemática	31
2.4.3. Estudios en otras disciplinas	39
3.El estudio empírico	46
3.1. Introducción	46
3.2. Metodología del estudio	46
3.3. Muestra de trabajos analizados	49
3.4. Análisis de las tesis. Resultados	54
3.4.1. Planteamiento del problema	54
3.4.2. Planificación	60
3.4.3. Datos	62
3.4.4. Análisis	68
3.4.5. Interpretación y conclusiones	71
3.4.6. Modos fundamentales de razonamiento estadístico	73
4. Conclusiones	77
Referencias	81
Anexo 1. Cursos metodológicos en Psicología Educativa	i
Anexo 2. Las asignaturas de estadística en la licenciatura de Psicología Educativa	Ix

INTRODUCCIÓN

En este trabajo se analizan algunas tesis de licenciatura presentados por los estudiantes de Psicología Educativa en la Unidad Ajusco de la Universidad Pedagógica Nacional de México. La finalidad es estudiar el uso de la estadística, dentro del proceso de investigación llevado a cabo por estos estudiantes en sus trabajos.

El interés por esta temática surge de nuestra concepción de la íntima relación entre estadística y metodología de investigación y de la creencia que la verdadera comprensión de la estadística se alcanza al aplicarla dentro de un problema en el campo de investigación del estudiante. Como indican Batanero y Díaz (2004), hay que diferenciar entre conocer y ser capaz de aplicar un conocimiento. La habilidad para aplicar los conocimientos matemáticos es frecuentemente mucho más difícil de lo que se supone, porque requiere no sólo conocimientos técnicos (cómo preparar un gráfico o calcular un promedio), sino también conocimientos estratégicos (saber cuándo hay que usar un concepto o gráfico dado). La evaluación a través de tareas estándares cubre tan sólo los conocimientos técnicos de la estadística, mientras que la evaluación a partir de proyectos y trabajos de investigación cubre también los conocimientos estratégicos.

Entendemos en este trabajo el razonamiento estadístico en un sentido amplio, siguiendo las directrices de Wild y Pfannkuch (1999), quienes en su modelo describen varios ciclos que juntamente componen el razonamiento estadístico. Uno de estos componentes es el ciclo de investigación que abarca desde la definición del problema a la obtención de conclusiones. Por este motivo, se analiza en los trabajos cada una de las fases descritas por los autores en su investigación y no sólo el análisis de datos. Nos interesamos también en los modos fundamentales de razonamiento estadístico que componen el tercer componente del modelo de Wild y Pfannkuch y analizamos algunos de los indicadores de dichos razonamientos puestos de manifiesto en las tesis de licenciatura.

La investigación realizada pretende contribuir a una mejora de la enseñanza de las disciplinas metodológicas en la Licenciatura de Psicología Educativa, incluyendo las asignaturas de Estadística. El interés por el tema surge de mi experiencia como docente en los cursos metodológicos y por las numerosas dificultades en la aplicación de la estadística y la metodología de investigación descritas en el campo de la Psicología (véase por ejemplo Wilkinson, 1999 o Fernández-Cano y Fernández, 2009). Estos

últimos fueron objeto de estudio en una Roundtable conference de IASE (Batanero, 2001).

Este interés nos llevó a contactar con el Grupo de Educación Estadística de la Universidad de Granada en el que se han realizado trabajos sobre enseñanza de la estadística a nivel universitario (Vallecillos, 1994; Alvarado, 2007; Díaz, 2007; Olivo, 2008). El trabajo tiene carácter exploratorio, ya que el tiempo disponible para la realización de la Memoria de Máster sólo nos ha permitido analizar una muestra reducida de trabajos y, asimismo, una muestra reducida de variables entre las que potencialmente se podrían tener en cuenta en el análisis. No obstante los resultados obtenidos hasta la fecha indican el interés del tema y posiblemente nos llevará a extenderlo con vistas a una futura tesis doctoral. El trabajo se organiza en la forma siguiente:

- En el primer capítulo se introduce el problema de investigación y se describe el contexto institucional en que se llevan a cabo los proyectos fin de carrera. Se analiza en forma somera la formación metodológica de los licenciados en Psicología Educativa en la Universidad Pedagógica Nacional y se describen los condicionantes de las tesis de Licenciatura. Se hace una breve síntesis del modelo de razonamiento estadístico de Wild y Pfannkuch (1999) que sirve de apoyo teórico a nuestro trabajo y se finaliza presentando los objetivos y algunas expectativas iniciales.
- El segundo capítulo incluye un breve estado de la cuestión en el que tratamos específicamente tres puntos: a) las controversias sobre el uso de la estadística en psicología y recomendaciones de mejora de la práctica metodológica; b) los errores y dificultades descritos en el uso de la estadística en la investigación; c) los estudios empíricos que analizan el uso de la estadística en la investigación y que son antecedentes del nuestro.
- En el tercer capítulo incluimos un breve estudio empírico en que se analizan algunas características metodológicas y el uso de la estadística en 65 tesis de licenciatura. Se describe la muestra y metodología de análisis para pasar a analizar el planteamiento del problema, muestras, instrumentos, métodos estadísticos empleados y razonamiento estadístico de los autores.

Presentamos finalmente algunas conclusiones e ideas para extender este trabajo, junto dos anexos que describen con más detalle las asignaturas de metodología y estadística.

I. EL PROBLEMA DE INVESTIGACIÓN

- 1.1. Introducción
- 1.2. La estadística en la investigación psicológica
- 1.3. Formación estadística en los estudios de Psicología
 - 1.3.1. Características generales del plan de estudios
 - 1.3.2. Las tesis de licenciatura
- 1.4. Razonamiento estadístico y sus componentes
 - 1.4.1. El ciclo de investigación
 - 1.4.2. Modos fundamentales de razonamiento estadístico
 - 1.4.3. Ciclo interrogativo
 - 1.4.4. Disposiciones
- 1.5. Objetivos del trabajo

1.1. INTRODUCCIÓN

Consecuencia de su diverso uso para explicar hechos de la vida cotidiana y el mundo científico, el conocimiento estadístico juega un papel cada vez más relevante en la sociedad actual. Un análisis del entorno es suficiente para percatarse que en casi todos los medios impresos y electrónicos existe información con datos estadísticos, gráficas y tablas, que hacen referencia a tendencias en el consumo de diferentes productos, a la distribución de la población y el uso de recursos naturales, entre otros.

La cultura estadística de un país tiene relación con su desarrollo y el grado en que su sistema estadístico produce estadísticas completas y fiables. Según Batanero (2004) esta información es necesaria para tomar decisiones acertadas de tipo económico, social y político. La formación adecuada, no sólo de los técnicos que producen estas estadísticas, sino de los profesionales y ciudadanos que deben interpretarlas y tomar a su vez decisiones basadas en esta información, así como de los que deben colaborar en la obtención de los datos requeridos, es por tanto, un objetivo educativo relevante.

1.2. LA ESTADÍSTICA EN LA INVESTIGACIÓN PSICOLÓGICA

Aunque en la actualidad continúa el debate sobre los contenidos pilares en la formación del psicólogo, un elemento común en los diversos currícula de psicología, en México, España y diversos países, es la base científica (De Vega, 2005). Por ello, estos currículos incluyen cursos en los que se revisa los métodos de la psicología científica y las técnicas derivadas de ella. Se trata de preparar al profesional para acceder y generar conocimiento científico, discernir entre las falsas teorías y las mejor fundamentadas,

utilizar y seleccionar instrumentos y aplicarlos significativamente.

En psicología se acepta la diversidad de procedimientos en la investigación (Seoane, 2005). Por ello debe fomentarse el aprendizaje significativo del componente metodológico, la importancia de plantear preguntas y convertirlas en problemas que, a través de la investigación permitan encontrar relaciones y aplicaciones del campo de investigación. La actividad profesional del psicólogo consiste en el estudio empírico de casos, individuales o de grupo, lo cual conlleva la recopilación de datos, análisis de los mismos, diseño de una intervención y evaluación de la misma (Carpintero, 2005). El psicólogo debe también saber “investigar” en sentido amplio. Por tal razón;

... debe saber evaluar resultados de su investigación o de su intervención de forma cuantitativa. Requiere aprender a emplear técnicas y procedimientos experimentales psicométricos u observacionales; dominar el uso de estadísticos; hacer búsquedas bibliográficas, leer y asimilar información publicada, redactar informes sobre los resultados de su investigación o intervención. Sin embargo, aunque se enfatiza en los currícula la formación científica, en el contexto laboral, este tipo de conocimiento (formación), subyace a la práctica profesional misma (De Vega, 2005, pp. 159).

En consecuencia, para una buena parte de los psicólogos, el análisis estadístico se precisa, tanto en el trabajo profesional como en la investigación, bien a nivel de interpretación de los resultados estadísticos en revistas o informes de colegas, bien en la producción de sus estudios estadísticos (Batanero, Godino y Vallecillos, 1992). Incluso el caso de colaboración con estadísticos profesionales no elimina la necesidad de conocimiento estadístico, pues el psicólogo debe plantear su problema, describiendo sus variables, hipótesis y datos al experto en estadística e interpretar los resultados que éste le proporciona (Barnett, 1988).

El conocimiento de la estadística interviene en todas las fases del trabajo de investigación, desde la decisión sobre las variables que se investigan y la planificación de la forma en que se han de recoger los datos, hasta la interpretación de los resultados obtenidos en el análisis de los mismos. Aunque en estas fases hay aspectos exclusivamente estadísticos, como serían los relativos al diseño del muestreo, o la elección del método concreto de análisis y su realización, otros dependen del problema planteado: elección de las unidades de análisis, variables, escalas de medida; proceso de

categorización, etc. Estos puntos no pueden dejarse a la responsabilidad del estadístico, puesto que dependen de las preguntas de investigación, que sólo tienen sentido profundo para la persona que las ha planteado (Batanero, Godino y Vallecillos, 1992).

1.3. FORMACIÓN ESTADÍSTICA EN LOS ESTUDIOS DE PSICOLOGÍA

Las anteriores razones, según Sánchez y Estepa (2000) y Estrada (2000) son suficientemente válidas para incorporar los contenidos estadísticos en los programas de estudio en la formación de profesionales. De hecho la mayor parte de las licenciaturas, estudios de máster y doctorado incluyen alguna asignatura de estadística, siendo esta materia la más enseñada a nivel universitario. Según Harris y Kanji (1988), un conocimiento básico estadístico es necesario a los profesionales para realizar las tareas estadísticas rutinarias, para reconocer los errores estadísticos en los trabajos de su especialidad y las situaciones en que precisan ayuda de un estadístico profesional. Schuyten (1990) señala las siguientes competencias estadísticas necesarias en el campo de la psicología:

1. *Competencias requeridas para interpretar la literatura de investigación:* Se necesita conocer y comprender los métodos usuales en psicología, su dominio de aplicación y las hipótesis en que se basan. Se requiere capacidad para juzgar las decisiones metodológicas de otros investigadores y las posibilidades de generalizar un estudio. Precisamente el campo de la psicología es uno de los que ha dado mayor importancia al empleo correcto de métodos estadísticos (Wilkinson, 1999). Es importante, entonces que el psicólogo pueda diferenciar aquellas investigaciones con una buena validez y fiabilidad de las que son irrelevantes.
2. *Competencias requeridas para el desarrollo de la investigación.* Hoy no se requiere destreza de cálculo, gracias a los paquetes estadísticos, pero se requiere elegir la técnica apropiada e interpretar sus resultados. El investigador tendrá que seleccionar sus muestras, codificar sus datos e introducirlos en el ordenador, realizar los cálculos sencillos (como medias, gráficos o correlaciones). También tendrá que escribir las conclusiones de su estudio y presentar los datos estadísticos de una forma apropiada.

Santisteban (1990) analizó diferentes programas de las facultades de Psicología de las universidades españolas, mostrando las similitudes de los cursos, y sus lagunas.

La mayoría de las universidades españolas ofrecen al menos dos cursos de Estadística, ubicados en los primeros dos periodos lectivos de la carrera. Esos contenidos se amplían y se profundizan en algunas asignaturas de semestres posteriores. Aunque al estudiar estadística al principio de la carrera todavía los estudiantes no aprecian su utilidad para temas de psicología, es necesario que se imparta al inicio, pues asignaturas como la Psicometría requerirán de estos conocimientos.

Los temas del primero de los cursos de estadística en la Licenciatura hacen referencia a la Estadística Descriptiva y el Cálculo de Probabilidades. Los contenidos que se revisan suelen ser: Tablas y gráficos, Estadísticos de posición y dispersión; Distribuciones de frecuencias; Correlación lineal simple; Regresión lineal simple y Probabilidad. Se pretende repasar los conocimientos que los estudiantes han adquirido en el Bachillerato, así como aplicarlos en el área de psicología.

En el segundo curso, se enseña la Estadística inferencial. Los programas más completos suelen incluir: Introducción al muestreo; Estimación de parámetros: puntual y mediante intervalos de confianza; Introducción a los contrastes paramétricos de hipótesis; Contrastes de medias, varianzas proporciones y coeficientes de correlación y regresión; Pruebas de independencia y de bondad de ajuste; Introducción al Análisis de la Varianza y de la Covarianza y Técnicas no paramétricas. Aunque como vemos, el temario es amplio, los cursos hacen una revisión muy general de la inferencia centrados en la propuesta por Fisher y Neyman y Pearson. En ninguno de los programas se revisa la Inferencia Bayesiana.

Esta aproximación ha sido discutida posteriormente por Díaz y Batanero (2006), quienes indican que la visión subjetiva amplía las aplicaciones de la probabilidad y la inferencia, no siendo ya un requisito la repetición de una experiencia en las mismas condiciones. Dichas autoras sugieren la necesidad de que en los cursos de pregrado de psicología se introduzcan algunas ideas de inferencia bayesiana, junto con los de inferencia clásica, sugiriendo el siguiente contenido en un curso de inferencia:

- Conceptos básicos fundamentales: Población, parámetro, muestra estadística, función de verosimilitud, distribución inicial y final.
- Estimación puntual: Métodos clásicos y bayesianos.
- Estimación por intervalos: Intervalos de confianza y de credibilidad.
- Contrastes de hipótesis: Pruebas clásicas y bayesianas, problemas de decisión múltiple.

En algunas universidades en los cursos de Psicología experimental, cuyos docentes son psicólogos se hace la revisión de temas como ANOVA y el diseño de experimentos. Finalmente, las Técnicas de Análisis Multivariado, Teoría de la Decisión, Análisis de Datos, Teoría de la Información, entre otros temas de uso en la psicología, son contenidos que solamente se imparten en algunas universidades.

Un elemento común en las universidades es "...Introducir a los psicólogos en el manejo y aplicación de paquetes de programas estandarizados con objeto de que analicen sus datos y sepan interpretarlos..." (Santisteban, 1990, p. 469). La autora indica que debe tenerse cuidado de no caer en el supuesto que el ordenador, resuelve la enseñanza de la Estadística, que no debe limitarse a la presentación de la técnica estadística como específica de un caso. De ese modo, el estudiante podrá comprender su aplicación en otras situaciones durante su desempeño profesional.

Algunos psicólogos profesionales o profesores de las asignaturas hacen comentarios al documento de Santisteban. Ubicado en una posición pragmática, Carpintero (1990) se pregunta cuánto debe profundizarse en la enseñanza de la estadística si se busca formar un psicólogo usuario de esa disciplina y no un especialista en ella. El autor menciona que, para utilizar la estadística de manera adecuada, el psicólogo necesita conocer y diferenciar las características numéricas y métricas de los instrumentos que maneja, así como sus restricciones. Requiere, además de conocer la lógica general de la investigación, distinguir el tipo de escala de medida y relaciones entre las variables; si son muestras relacionadas o independientes; si se supone una distribución paramétrica de la variable que se mide. Sobre todo precisa comprender qué es y qué no es estadísticamente significativo y la manera en que esto puede variar de acuerdo con el tamaño de las muestras o con el uso de una prueba con mayor potencia. Para aprender a utilizar técnicas como el análisis de varianza, la correlación o la regresión, el autor sugiere contar con un modelo que funcione como paradigma, que los estudiantes usen con sus datos. La tarea es elaborar programas con situaciones educativas que, de manera gradual, conduzcan al estudiante, mediante el contacto con procedimientos cada vez más complejos, a consolidar el pensamiento estadístico.

Para decidir la amplitud y profundidad de los programas de Estadística en la formación de psicólogos es indispensable considerar dos formas de aplicarla (Ponsoda, 1990; Yela, 1990). La primera manera, propia de los trabajos que se publican, puede ser

inferencia estadística a partir del contraste de hipótesis o estudios de tipo exploratorio o descriptivo en los que no se requiere el contraste de hipótesis. En ambos casos se utiliza la estadística como herramienta para el análisis de los resultados. Otra vertiente es el uso de la estadística en la formulación de teorías o modelos que requiere un conocimiento de la estadística mucho más amplio.

Ponsoda y Yela plantean que la formación estadística del psicólogo debe orientarse a que éste pueda analizar los datos que él genere en su práctica profesional y, del mismo modo, comprender los resultados en los documentos publicados en las distintas áreas de la psicología. En conclusión, el criterio debe ser que en los cursos se enseñe las técnicas de mayor uso en la Psicología. Para identificarlas es pertinente revisar los procedimientos estadísticos más utilizados en los informes psicológicos. Pero debe atenderse en no convertir la formación en un círculo reiterado. Es decir, enseñar a los psicólogos las técnicas estadísticas que aparecen en las revistas; ellos las usarán cuando publiquen un informe, por lo cual sólo se publicará lo que se sabe, esto es, lo que se enseña. Es tarea del profesor estar al tanto de los procedimientos estadísticos novedosos y útiles para la psicología, de modo que sea factible la actualización de los contenidos.

Yela (1990) sugiere que la estadística debe ser una herramienta para su desempeño profesional. En ese sentido, la enseñanza no puede recaer en personas que sólo sean matemáticos o expertos en estadística, ni tampoco en psicólogos que no tengan suficiente conocimiento de esa disciplina. Al respecto el autor observa una participación, cada vez más frecuente, de profesores especialistas en la asignatura, provenientes de otros campos disciplinares en particular de las Ciencias Exactas. Asimismo, destaca el papel que juegan los psicólogos con formación matemática quienes promueven y apoyan la enseñanza de esos contenidos. Con el transcurrir del tiempo, esta relación de cercanía entre profesionales de dos o más campos científicos ha generado las condiciones para realizar trabajo conjunto y en colaboración indispensable para impulsar no sólo la formación de los estudiantes, sino para el desarrollo de la investigación y la docencia.

En lo que sigue analizamos brevemente formación y las tesis de Licenciatura en la Licenciatura en Psicología Educativa en la Universidad Pedagógica Nacional, Unidad Ajusco, de México. En los anexos se hace una descripción detallada de las asignaturas metodológicas y estadísticas en dicho plan de estudio.

1.3.1. CARACTERÍSTICAS GENERALES DEL PLAN DE ESTUDIOS

En el Decreto de Creación de la Universidad Pedagógica Nacional (López-Portillo, 1978) se plantea que su objetivo es formar profesionales de la educación cuyo desempeño profesional ofrezca propuestas de atención a los distintos problemas de la educación nacional. Los planes de estudio 1979 para formar profesionales de la educación, se crearon con un diseño interdisciplinario. Se incorporaron cursos generales de metodología y matemáticas, y se dio fuerte énfasis al contexto mexicano. Los contenidos conformaron la estructura curricular común para las cinco licenciaturas de la universidad: Administración Educativa, Educación Básica, Pedagogía, Sociología de la Educación y Psicología Educativa. Los estudiantes de esta institución eran profesores de educación primaria en servicio, quienes pretendían adquirir el grado de licenciatura. Esto llevó a incorporar materias de iniciación en los estudios universitarios durante los primeros semestres y, posteriormente, únicamente en los dos últimos años, para ofrecer una formación profesional. El plan de estudios de la Licenciatura en Psicología Educativa (Plan 1979), constaba de 32 cursos y sólo 12 se dedicaban a la formación profesional en la disciplina. El objetivo de formar con un conocimiento básico de la disciplina y la urgencia de profesionalizar con ciertas herramientas de intervención fue apenas logrado.

El Plan de estudios 1990 fue resultado de la evaluación de la licenciatura y de la preocupación por integrar en la formación, distintos conocimientos, habilidades y herramientas para intervenir profesionalmente en problemas educativos. Los cursos con contenidos psicológicos y educativos aumentaron en forma considerable; se integraron en áreas curriculares y se planteó una visión mutidisciplinar. Las herramientas metodológicas y de intervención profesional se integraron con el estudio de las prácticas y los contextos educativos (Hernández, Sánchez, Ortega y Millán, 1990). El diseño de este plan de estudios adoptó un enfoque constructivista del aprendizaje y la enseñanza. Con este marco se hizo una selección de contenidos de la Psicología Evolutiva y Cognitiva, y se vincularon con el análisis de problemas del diseño y evaluación de la enseñanza y de la organización escolar para potenciar el aprendizaje de los estudiantes.

La organización de estos contenidos en los distintos cursos promovió la integración de contenidos derivados de la investigación y diseño de aplicaciones

enfocadas en los procesos de formación de los sujetos en situaciones educativas; así como de valores profesionales de honestidad y respeto al valor de las personas, sistematicidad y rigor metodológico en una intervención profesional y una actitud de servicio para atender los problemas de la Educación Básica. Las áreas curriculares que definen la agrupación de las materias, de acuerdo con las temáticas y los problemas afines que conducen a una profundización en los contenidos, son las siguientes:

- a. *Psicología Evolutiva*, agrupa materias relacionadas con los procesos de cambio del educando y la forma como se relaciona con distintos contextos.
- b. *Psicología Educativa*, agrupa materias vinculadas con los procesos de aprendizaje y de diseño de la enseñanza.
- c. *Metodológica*, agrupa materias relacionadas con la formación de habilidades básicas para estudiar e intervenir en procesos de enseñanza aprendizaje en contextos educativos. Las asignaturas relacionadas se describen con detalle en los anexos 1 y 2.
- d. *Psicosocial*, agrupa materias acerca de los procesos de aprendizaje e interacción social en el grupo escolar y la escuela como organización.
- e. *Currículo*, agrupa materias relacionadas con el diseño de planes y programas, el desarrollo de proyectos y la evaluación curricular.
- f. *Integración educativa*, agrupa cursos vinculados a la atención a la diversidad y el diseño de intervenciones psicopedagógicas.
- g. *Socio histórica*, agrupa cursos acerca de las funciones sociales de la educación y los proyectos educativos de la sociedad mexicana.

1.3.2. LAS TESIS DE LICENCIATURA

En el reglamento de titulación de Licenciaturas de la Universidad Pedagógica Nacional se establece que la titulación es la fase final de la formación profesional y que la realización de una Tesis de Licenciatura es la forma de evaluar la capacidad crítica, reflexiva y transformadora que los egresados poseen. Asimismo se indica que dicho trabajo es una actividad formativa que permite a los estudiantes poner en práctica los conocimientos y habilidades desarrolladas durante sus estudios (UPN, 2002).

Posteriormente establece que para la obtención del título de licenciatura se requiere tener tres requisitos: (a) cubrir el 100% de los créditos del plan de estudios de la carrera; (b) haber realizado el servicio social y (c) realizar el proyecto, obtener su

aprobación, elaborar el trabajo recepcional, obtener los votos aprobatorios del jurado y acreditar el examen profesional. Para la Licenciatura en Psicología Educativa existen tres modalidades (Bollás, Ruiz, Castillo y Pedraza, 2005). Una de ellas es la tesis, la cual se caracteriza de la siguiente manera:

“El objetivo de la tesis es la construcción de un problema de investigación dentro del campo temático o problemática psicoeducativa. El trabajo consiste en desarrollar un argumento lógico a través de la estructuración, ordenamiento y definición de los conceptos centrales, estableciendo las relaciones entre ellos con la finalidad de fundamentar teóricamente la investigación y derivar una pregunta a la que debe darse respuesta.

Para dar respuesta a la pregunta de investigación se deberá proporcionar evidencia obtenida de manera directa y sistemática en el contexto educativo. Dicha evidencia puede ser contrastada con datos obtenidos por otras investigaciones. La presentación de tesis puede ser individual o en grupo: de ser en grupo, el número de integrantes no excederá de tres.” (p. 7).

Es tarea de la Comisión de titulación de la Licenciatura dictaminar los proyectos de trabajo fin de carrera. Por tal razón en las instrucciones se propone un formato para elaborar el proyecto de trabajo recepcional. Este proyecto, así como la tesis se desarrollan con la asesoría de un profesor de la licenciatura o de la universidad, o en pocos casos, por un profesional de la educación externo a UPN. Tanto en el reglamento de titulación, como en las instrucciones de titulación, la descripción de tesis es muy general, el estudiante tiene amplia libertad para elegir el tema, enfoque y los aspectos metodológicos. El estudiante elige libremente el tema y busca un asesor; la mayor parte de las veces no suele ser el profesor de estadística o metodología, pues en general los profesores de los primeros cursos son matemáticos o estadísticos, no psicólogos. La profundidad de la orientación y apoyo para el estudiante es variable, pues el número de estudiantes asesorados por un profesor puede fluctuar entre 1 y 10.

Una oportunidad para intentar definir qué es una tesis y cómo se organiza, se proporciona a los estudiantes dentro de los cursos Seminario de tesis I y Seminario de tesis II (ver anexos). Estos cursos se diseñaron para apoyar a los estudiantes en el desarrollo y elaboración de su trabajo recepcional (tesis o tesina). En la organización curricular, existen cursos específicos de cada especialidad, denominados Seminarios

taller de concentración I y II, que 'aportan' el contenido temático: orientación educativa, problemas del escolar, aprendizajes escolares, currículo, grupos y organización escolar. Los objetivos de formación que se persigue son diferentes en cada área de concentración profesional. En el capítulo III se hace una breve descripción.

1.4. RAZONAMIENTO ESTADÍSTICO Y SUS COMPONENTES

En el desarrollo de la Tesis y en las asignaturas de metodología y estadística descritas en los Anexos, son objetivos importantes no sólo la adquisición de conocimientos conceptuales o procedimentales, es decir, conocimientos técnicos, sino también conocimientos estratégicos, que se ponen de manifiesto cuando el estudiante elige una opción entre muchas posibles (por ejemplo, elige sus variables, diseño, instrumento o métodos estadísticos). Se trata de desarrollar su razonamiento estadístico.

Conviene, por tanto, en esta sección analizar en qué consiste este tipo de razonamiento. Diversos autores (Moore y Cobb, 2000; Rossman, Chance y Medina, 2006; Scheaffer, 2006) han analizado las diferencias y complementariedades entre el razonamiento matemático y estadístico. Otros han descrito modelos más o menos sofisticados sobre el razonamiento estadístico, siendo el más conocido y citado el de Wild y Pfannkuch (Wild y Pfannkuch, 1999; Pfannkuch y Wild, 2004). En todo trabajo estadístico se lleva a cabo un sistema de procesos interconectados. Los autores describen varios componentes, que usa un estadístico simultáneamente durante su trabajo y que conjuntamente constituyen su modelo de razonamiento estadístico: (a) el ciclo de investigación; (b) los modos fundamentales de pensamiento estadístico; (c) el ciclo de interrogación y (d) disposiciones del estadístico.

1.4.1. EL CICLO DE INVESTIGACIÓN

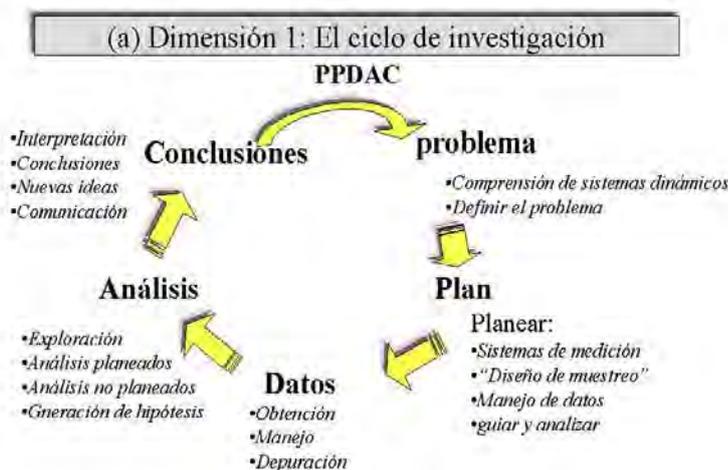
La primera dimensión del modelo es el ciclo de investigación (Figura 1.4.1); ésta abarca cinco fases: Planteamiento del problema, Planificación, Recogida de Datos, Análisis y Obtención de Conclusiones (PPDAC), que tienen correspondencia con las fases del "método de investigación científica". Nosotros tomaremos este ciclo y sus componentes como primer elemento de análisis de las Tesis de Licenciatura en el estudio empírico que se reporta en el Capítulo 3. Como vemos, el ciclo tiene también similitudes con las etapas recomendadas por Polya (1944) en la resolución de problemas: Entender el enunciado; Diseñar un plan de actuación; Ejecutar el plan;

Comprobar y validar. Sin embargo, se ha separado la recogida y análisis de datos, dentro de la fase de ejecución y se da gran importancia a la parte de extracción de conclusiones.

El problema

Un ciclo PPDAC tiene que ver con resolución de un problema inherente a una problemática real, en nuestro caso del campo de la Psicología Educativa. La fase de planteamiento del problema es una de las más difíciles, pues según Wild y Pfannkuch los problemas estadísticos raramente están completamente definidos o claramente formulados. Al iniciar su tesis de licenciatura, los estudiantes podrían comenzar sin preguntas claramente definidas. Para poder concretarlas, tendrá que comprender el sistema que quieren investigar, sus condicionantes y variables y elegir o formular las que tomarán como dependientes o independientes en su trabajo. Comenzarán por la elección de la temática, la decisión sobre las variables que se estudiarán y el papel de cada una en la investigación, la definición de objetivos o preguntas de investigación, así como la elaboración de conjeturas o hipótesis sobre lo que esperan encontrar.

Figura 1.4.1. El ciclo de investigación (Wild & Pfannkuch, 1999, p. 7, trad., E. Sánchez)



El plan

Definido el problema, hay que trazar un plan de resolución. La planificación abarca diferentes puntos. Se trata de anticipar los pasos en la resolución del proyecto, en sus diferentes dimensiones: poblaciones y muestras que se elegirán, técnicas de análisis,

gestión de los datos, etc. Un punto importante será el tipo de diseño elegido, puesto que la misma temática de investigación puede enfocarse de forma diferente.

Los datos

Batanero y Díaz (2009) indican que a veces los datos de un proyecto estadístico son proporcionados por el profesor. Otros son recogidos por los estudiantes mediante la realización de una encuesta o a través de un experimento. La encuesta requerirá la elaboración de un cuestionario, fijando los objetivos del mismo, eligiendo las variables explicativas y redactando las preguntas que permitan obtener la información deseada de una forma clara y concisa. Si se pretende extender los resultados más allá de la muestra, la selección de una muestra representativa plantea problemas de tipo teórico y práctico, relacionados con la población objetivo y alcanzada, el marco de muestro, los métodos de selección, la administración del cuestionario y los problemas de no respuesta.

La información recogida puede corresponder a diversos niveles que se corresponden con diferentes técnicas de obtención de datos: información consciente y conocida (encuesta), información desconocida, pero que puede deducirse de la observación e información no consciente ni observable (medida). Finalmente es importante considerar la naturaleza de las escalas de medida y tipo de variable, de ellas depende el método de análisis de datos que se puede aplicar. La elección del conjunto de datos es crítica, pues no todas las técnicas son aplicables a cualquier tipo de dato.

El análisis

La naturaleza de los datos, variables y tipos de hipótesis influye en los tipos de análisis estadísticos. Si no se tiene ninguna hipótesis previa, tal vez se aborde simplemente un estudio exploratorio cuya finalidad sea precisamente generar las hipótesis. En otros casos será necesario métodos inferenciales, que, dependiendo del nivel de medición de la variable serán paramétricos o no paramétricos. En cualquier caso, el proceso no es lineal y pueden generarse nuevas hipótesis del mismo análisis estadístico con lo cual, además de los análisis planeados será necesario efectuar análisis no planeados. Será importante la correcta elección del método, en función de las variables, el tipo de pregunta de investigación y los supuestos del método.

Las conclusiones

Esta es una fase crucial, pues hay que interpretar los resultados del análisis estadístico, en función de la pregunta planteada originalmente y dotarlos de sentido en el contexto de los datos. Es la última fase en el ciclo de investigación y también en el ciclo de modelización que, según, Dantal (1997), consta de los siguientes pasos: (a) Observación de la realidad; (b) Descripción simplificada de la realidad; (c) Construcción de un modelo; (d) Trabajo matemático con el modelo, (e) Interpretación de resultados en la realidad. El autor indica que en la enseñanza sólo se consideran los pasos 3 y 4 que son los más sencillos de enseñar. Sin embargo todas las etapas son igualmente importantes en el aprendizaje si queremos que realmente los alumnos lleguen a comprender la utilidad y la razón de ser de las matemáticas.

1.4.2. MODOS FUNDAMENTALES DE RAZONAMIENTO ESTADÍSTICO

Este componente del modelo de razonamiento estadístico de Wild y Pfannkuch (1999) es el más importante de todos, pues incluye las estrategias que se pueden seguir para resolver el problema. La mayoría de estas estrategias son las usuales en resolución de problema: por ejemplo, las recomendadas por Polya (1944); análisis y síntesis; descomponer un problema en partes, buscar un problema similar; fijar variables; generalizar. No analizaremos estas estrategias en este trabajo, aunque podrían considerarse para un análisis futuro. En este segundo componente, los autores incluyen también, las siguientes formas fundamentales de razonamiento estadístico:

- *Reconocer la necesidad de los datos:* La base de la investigación estadística es la hipótesis de que muchas situaciones de la vida real sólo pueden ser comprendidas a partir del análisis de datos que han sido recogidos en forma adecuada. La experiencia personal o la evidencia de tipo anecdótico no es fiable y puede llevar a confusión en los juicios o toma de decisiones.
- *Transnumeración:* Los autores usan esta palabra para indicar la comprensión que puede surgir al cambiar la representación de los datos. Al considerar un sistema real desde la perspectiva de modelización, puede haber tres tipos de transnumeración: (1) a partir de la medida que “captura” las cualidades o características del mundo real, (2) al pasar de un listado de datos desordenados a una representación tabular o gráfica que permita extraer sentido de los mismos; (3) al comunicar este significado que surge de los datos, en forma que sea comprensible a otros.

- *Percepción de la variación:* La variación resulta cuando dos o más cosas son diferentes y es característica de la estadística que es concebida por Moore y Cobb (2000) como la ciencia que estudia la variabilidad. La variación aparece en los datos, las muestras, las variables estadísticas y aleatorias, los instrumentos usados, las respuestas obtenidas, etc. La recogida adecuada de datos y los juicios a partir de los mismos requieren la comprensión de la variación que hay en los datos, y de la incertidumbre originada por la variación no explicada. La estadística permite hacer predicciones, buscar explicaciones y causas de la variación y aprender del contexto.
- *Razonamiento con modelos estadísticos.* Cualquier útil estadístico, incluso un gráfico simple puede considerarse como modelo, pues es una forma de representar la realidad. Lo importante es diferenciar el modelo de los datos y, al mismo tiempo, relacionar el modelo con los datos. Aunque en matemáticas también se usan modelos, la semejanza entre modelos y datos es mucho mayor que en el caso de la estadística. Por ello la modelización es más difícil en esta segunda materia.
- *Integración de la estadística y el contexto:* Es también un componente esencial del razonamiento estadístico. Mientras que en matemáticas lo esencial es los conceptos y las relaciones entre ellos, en estadística juega un papel mucho más importante el contexto, de modo que el mismo concepto puede tener diferente significado, dependiendo del mismo.

1.4.3. CICLO INTERROGATIVO

La tercera dimensión en el trabajo de Wild y Pfannkuch (1999) es el ciclo interrogativo (Figura 1.4.3); ésta es una serie encadenada de razonamientos que se usan constantemente en la resolución de problemas estadísticos. Este ciclo se aplica en un nivel micro, dentro del ciclo 1 de investigación, pero también en todo aquél ciclo, pues es recursivo. En cada paso, el estadístico debe comenzar generando posibilidades de: (a) estrategias para resolver el problema, (b) posibles explicaciones o modelos que respondan a los datos recogidos y (c) de nueva información que se necesita para continuar.

El investigador tendrá que buscar con frecuencia información e ideas, tanto sobre los datos, sobre el problema o sobre los modelos (por ejemplo, cómo usar la paquetería, qué diferentes tipos de contraste podría aplicar). Esta información necesita

ser interpretada, leída (a veces traducida, tanto de un idioma a otro, como al propio conocimiento del investigador). La información discordante tendrá que ser conectada, comparada y asimilada respecto a la previa. Es importante la crítica del propio trabajo, por medio de referentes externos e internos, comprobando las conjeturas iniciales y las que se van generando en el proceso. Finalmente llegará el proceso de decisión sobre qué aspectos tomar cuáles descartar y cómo seguir al paso siguiente.

Figura 1.4.3. El ciclo interrogativo (Wild & Pfannkuch, 1999, p. 18, trad., E. Sánchez)



1.4.4. DISPOSICIONES

La cuarta dimensión del modelo de Wild y Pfannkuch son las disposiciones. Éstas se refieren a las actitudes de los estadísticos. Incluye una serie de características, como curiosidad y compromiso para mantener constancia en el trabajo. Se requiere imaginación y creatividad, para generar las ideas y para conseguir un trabajo original. Pero también se necesita una dosis de escepticismo, pues siempre hay que revisar los resultados con espíritu crítico, hasta que la evidencia nos dé un apoyo a nuestras conclusiones. Finalmente añade la necesidad de ser lógico, una propensión para buscar un significado más profundo de lo que los datos parecen arrojar a primera vista, apertura a nuevos puntos de vista y perseverancia en el trabajo. Todo lo anterior lleva al *compromiso*. Cuando el investigador llega a estar intensamente interesado en un problema, se desarrolla una elevada sensibilidad y conciencia hacia la información relacionada con el problema. El compromiso intensifica cada uno de los elementos de la “disposición”: curiosidad, conciencia, imaginación y perseverancia.

1.5. OBJETIVOS DEL TRABAJO

En términos generales, nos planteamos iniciar una investigación con una componente metodológica, a la vez que con una componente didáctica. Desde el punto de vista metodológico, el trabajo se orienta a iniciar una evaluación del uso de las herramientas que los estudiantes han adquirido en su formación del área metodológica y estadística en los trabajos de Tesis de licenciatura de la titulación en Psicología Educativa Universidad Pedagógica Nacional en México, D.F. La finalidad, desde este punto de vista metodológico sería evaluar el uso correcto de las técnicas de investigación en los egresados, asegurando, de este modo que han adquirido las competencias necesarias para abordar otras investigaciones en su trabajo futuro. En caso de encontrar un uso deficiente, el trabajo serviría para evaluar los puntos más débiles, y pensar en posibles acciones de mejora. Contribuiríamos de este modo a mejorar la práctica estadística y metodológica en Psicología, cuya necesidad es resaltada desde la American Psychological Association (Wilkinson, 1999).

Desde un punto de vista didáctico, y, como profesor de estas asignaturas, la finalidad última sería evaluar la efectividad de la enseñanza recibida por los egresados, del Área Metodológica, que incluyen los cursos de estadística y metodología, así como en los seminarios de investigación (ver anexos). Por un lado nos interesa saber si los métodos y las técnicas utilizadas coinciden con las enseñadas en dichas asignaturas, así como si se siguen las directrices metodológicas generales que tratamos de impartir a los titulados y que son también parte del razonamiento estadístico en el modelo de Wild y Pfannkuch (1999). Las posibles dificultades o errores que encontremos permitirán evaluar en qué contenido o grado sería necesario cambiar o complementar los actuales programas de las asignaturas. Un uso inadecuado de las técnicas nos alertaría de posibles errores conceptuales o de conceptos no suficientemente adquiridos y nos proporcionaría también una base para mejorar nuestra acción didáctica.

Puesto que este objetivo general es demasiado ambicioso, en el trabajo de Máster sólo lo abordaremos parcialmente, aunque lo podremos retomar en una futura tesis doctoral. Para este trabajo señalamos dos objetivos principales:

- El primer objetivo es *iniciar un estado de la cuestión* sobre la problemática del uso adecuado de métodos estadísticos en la investigación (con particular énfasis en la investigación psicológica). Consideramos este objetivo necesario para comenzar cualquier investigación más profunda sobre el tema, en cuanto nos proporcionará

una base para abordar el problema, elegir las variables del presente estudio e interpretar sus resultados. Este objetivo se aborda en el Capítulo 2, donde se elabora una síntesis sobre esta problemática que será ampliada en el futuro.

- *El segundo objetivo es iniciar un estudio limitado* (en cuanto al número de documentos analizados y las variables consideradas) de evaluación del uso de las herramientas metodológicas y estadísticas en las Tesis de Licenciatura. Tendrá también un carácter exploratorio, nos servirá para iniciarnos en la metodología de un estudio futuro y para comenzar a conocer los materiales que posteriormente podríamos analizar. Se aborda en el capítulo 3 y servirá para adquirir una experiencia investigadora en el periodo de Máster.

Los objetivos planteados se ubican en el campo de investigación denominado etnoestadística (Gephart 1988; 2006; Mills, Weatherbee y Colwell, 2006). En la etnoestadística: *“La preocupación fundamental de la etnoestadística es describir, analizar, explicar y comprender cómo se usa realmente la estadística en el proceso de investigación. La etnoestadística examina los aspectos cualitativos de la realización de análisis estadísticos, y utiliza este examen como fuente de comprensión de los procesos sociales que subyacen en el conocimiento científico”* (Gephart, 1988, p. 11). El autor menciona que existen tres tipos de estudios etnoestadísticos:

1. *Análisis de la producción de datos estadísticos.* En esta clase de estudios se investigan las actividades, significados y contextos implicados en la producción de la información estadística que ha de servir de base para el análisis. Se observan las normas sociales, convenios y prácticas de los productores de datos estadísticos o de los grupos y ambientes de investigación naturales en los que se producen y usan los datos estadísticos. Nosotros no tocamos este punto, pues no tenemos acceso al momento en que los datos de los estudiantes fueron recogidos.
2. *Adecuación del uso de los métodos e hipótesis.* Se evalúa cómo el investigador, los datos, el instrumento de evaluación y el análisis realizado afectan a los resultados de las investigaciones. Se realizan análisis críticos de las propiedades técnicas del uso de la estadística en todo el proceso de investigación o en algunas fases. Esta es una preocupación característica de las investigaciones metodológicas, y en este trabajo, nosotros iniciamos un acercamiento a este problema.
3. *Retórica de la estadística.* Se analiza el uso de la estadística como argumentación,

para convencer o persuadir al lector del informe. Aunque en esta memoria no abordamos este punto, no lo descartamos para el futuro.

Se relaciona también con los estudios de cienciometría, en los que se analizan cuantitativamente documentos y artículos de investigación, con finalidad de evaluación de la misma. Consideraremos en este trabajo algunos estudios realizados en Didáctica de la Matemática por Torralbo y colaboradores (Torralbo, 2001; Torralbo, Vallejo y Fernández, 2001; Torralbo, Fernández, Rico, Maz y Gutiérrez, 2003; Torralbo, Vallejo, Fernández y Rico, 2004; Vallejo, 2005; Vallejo, Fernández, Torralbo y Maz, 2007), que, aun cuando estos trabajos tienen un contenido más amplio que el nuestro, también analizan las características metodológicas de trabajos de investigación y consideran algunas de nuestras variables.

Nuestra finalidad es, sin embargo, algo diferente de los estudios de etnoestadística o cienciometría, pues es principalmente didáctica. Se trata, por un lado, de valorar de algún modo el aprendizaje de los egresados en las áreas de metodología y estadística para en el futuro utilizar esta información con objeto de mejorar la docencia en estas disciplinas. Por otro, también se trataría de analizar si es necesario mejorar la cultura estadística de los profesores no especialistas en metodología, con objeto de ofrecerles cursos de apoyo o consultoría estadística en la labor de asesoramiento de las tesis de licenciatura.

CAPÍTULO 2. ANTECEDENTES

- 2.1. Introducción
- 2.2. La controversia sobre la metodología estadística en Psicología
 - 2.2.1. Introducción
 - 2.2.2. Críticas al contraste de hipótesis
 - 2.2.3. Recomendaciones para mejorar el uso de la estadística en psicología
- 2.3. Errores en inferencia estadística
 - 2.3.1. Introducción
 - 2.3.2. Significación estadística
 - 2.3.3. Hipótesis estadísticas y lógica de contraste
 - 2.3.4. Intervalos de confianza
- 2.4. Análisis empíricos del uso de la estadística en la investigación
 - 2.4.1. Estudios en Didáctica de la Matemática
 - 2.4.2. Estudios en otras disciplinas

2.1. INTRODUCCIÓN

La investigación en distintas ciencias experimentales y sociales y, en particular, en Psicología Educativa necesita recoger y analizar datos sobre los fenómenos que estudia. En consecuencia, los conceptos y procedimientos estadísticos y las técnicas de análisis de datos, tanto cualitativos como cuantitativos, deben constituir un objeto de estudio por parte de la comunidad de investigadores. También deben ser tema de análisis por los profesores encargados de la formación en esta materia.

Estas consideraciones nos han llevado a iniciar un estado de la cuestión, para recoger y resumir las investigaciones que traten este tema y nos permitan elaborar y fundamentar nuestro trabajo. Contamos con algunos artículos de síntesis que resumen la investigación relacionada, entre ellos los de Batanero (2000), Díaz y de la Fuente (2004), Batanero y Díaz (2006b), Castro-Soto, Vanhoof, Van den Nororgate y Onghena, (2007) y Díaz, Batanero y Wilhelmi (2008), que hemos sintetizado y complementado con algunos trabajos no considerados por estos autores. Un análisis inicial de la literatura relacionada nos ha llevado a identificar los siguientes temas relacionados con el que nos ocupa y que se analizan en este capítulo:

1. La literatura relacionada con la crítica y sugerencias para mejorar el uso de la estadística en Psicología. Estas publicaciones se inician con la controversia surgida en relación con el abuso que se hace del contraste estadístico de hipótesis y la confusión entre diversas metodologías de inferencia
2. Los estudios de evaluación sobre aprendizaje y errores en inferencia estadística, que son relativamente abundantes. Nos centraremos preferentemente en los que tienen

como sujetos de estudio investigadores.

3. Los estudios empíricos documentales que analizan el uso de la estadística en la investigación, tanto en Didáctica de la Matemática, como en otras ciencias.

2.2. LA CONTROVERSIAS SOBRE LA METODOLOGÍA ESTADÍSTICA EN PSICOLOGÍA

2.2.1. INTRODUCCIÓN

En primer lugar nos preocupamos por describir muy resumidamente el debate en torno al contraste de hipótesis que lleva ya más de 30 años de historia. La estadística ha jugado siempre un papel destacado en psicología, aunque su uso no siempre ha sido correcto o bien comprendido por los investigadores (Morrison y Henkel, 1970; Batanero, 2000; Batanero y Díaz, 2006a; Fernández-Cano y Fernández, 2009). Ello ha ocasionado una situación peculiar, como analizan Díaz y de la Fuente (2004), donde, por un lado, la estadística se usa mal, pero por otro, la obtención de resultados estadísticamente significativos se ha convertido en un requisito frecuente para que los trabajos sean aceptados en las revistas científicas o congresos del área.

Las autoras sugieren que esta costumbre ha sido muy criticada, pues algunos investigadores toman la significación estadística como criterio para valorar si una investigación (significativa/no significativa) tiene calidad, sin preocuparse de la significación práctica de sus resultados. Por otro lado, los investigadores se limitan a mencionar el nivel de significación, sin incluir intervalos de confianza para los valores medios de las variables u otra información sobre los efectos producidos (Morrison y Henkel, 1970; Vallecillos, 1994; Borges, San Luis, Sánchez y Cañadas, 2001; Díaz, 2007).

Los investigadores no siempre son conscientes de la existencia de estas diferentes escuelas de inferencia estadística y a menudo se usa la inferencia estadística mezclando conceptos y supuestos de las mismas. De acuerdo con Gingerenzer (1993), Neyman y Pearson idearon el contraste de hipótesis como un proceso de decisión entre dos hipótesis. Sin embargo algunos investigadores emplean el contraste de hipótesis para proporcionar evidencia en contra de una única hipótesis sin especificar cuál es la hipótesis alternativa, lo que es propio de la filosofía de Fisher. Otros investigadores dan una interpretación Bayesiana a los resultados de los contrastes de hipótesis clásicos.

2.2.2. CRÍTICAS AL CONTRASTE DE HIPÓTESIS

Todos estos errores han ocasionado fuertes críticas al uso actual del contraste de hipótesis (Morrison y Henkel, 1979; Borges, San Luis, Sánchez y Cañadas, 2001) que son revisados por Batanero (2000), Díaz y de la Fuente (2004) y Díaz (2007) y resumimos a continuación.

- *Se argumenta que la lógica del contraste de hipótesis es inválida, pues la hipótesis nula siempre es falsa.* Las hipótesis experimentales en la investigación psicológica siempre proponen un efecto de una determinada variable (por ejemplo, mejora del rendimiento con las horas de estudio), pero la hipótesis nula se plantea como si no hubiera efecto. Es decir, el investigador asume que la hipótesis nula que plantea a priori es falsa. Todo el razonamiento matemático del contraste de hipótesis parte del supuesto de que la hipótesis nula es cierta. Sería por tanto inválido usar el contraste de hipótesis, puesto que se construye sobre un supuesto falso.

De la Fuente y Díaz indican que esta crítica no tiene en cuenta la lógica del contraste de hipótesis, pues éste no depende de que la hipótesis nula sea falsa o verdadera. El hecho de que un resultado significativo sea improbable si la hipótesis nula fuese cierta se deduce matemáticamente de la distribución muestral del estadístico, que es válida incluso cuando la hipótesis nula es falsa. Es decir, nos basamos en una probabilidad condicional del tipo $P(A/B)$ y aunque B sea falsa en un caso particular, la probabilidad $P(A/B)$ no cambia.

- *Se argumenta que el contraste de hipótesis es inválido porque las dos hipótesis no juegan el mismo papel.* Algunos investigadores piensan que no se debiera usar el contraste de hipótesis, porque no se trata igualmente a las hipótesis nula y alternativa. No se entiende por qué cuando obtenemos un resultado significativo rechazamos la hipótesis nula, pero cuando obtenemos un resultado no significativo sólo decimos que no hemos podido rechazarla, sin llegar a aceptarla.

Aunque es cierto que el papel de las dos hipótesis en el contraste no es simétrico, ello no invalida la utilidad del contraste de hipótesis. Los resultados significativos y no significativos no tienen el mismo valor de prueba, debido al tipo de razonamiento que se hace, similar (aunque no equivalente) a la prueba por contradicción (Gigerenzer, 1993). En una prueba por contradicción razonamos de la siguiente forma: Si A

es cierta, B es falsa, entonces si B ocurre, podemos concluir que A es falsa. Pero cuando no ocurre B no podemos deducir que A sea necesariamente cierta; hay una asimetría entre las consecuencias de B y no B . Sin embargo, este hecho no invalida la prueba por contradicción y lo mismo sucede en el contraste de hipótesis (Batanero, 2000).

- *No debiera aceptarse un procedimiento matemático que fija un valor arbitrario para realizarse: El nivel de significación.* Es cierto que no hay una regla matemática para establecer el valor del nivel de significación. Por tanto los mismos datos podrían ser estadísticamente significativos o no, dependiendo del nivel fijado. No parece propio de un procedimiento matemático, pues la solución es algo subjetiva.

Esta crítica es cierta, según Díaz (2007), pero el componente subjetivo de la estadística no implica que el contraste de hipótesis sea inútil. Además, aunque en la práctica actual se sugiere elegir el nivel de significación antes de recoger los datos para ser más objetivos, si se sigue el enfoque de Fisher, sería posible también usar el p-valor exacto para rechazar las hipótesis nulas a diferentes niveles de significación.

- *Se presta demasiada atención al error tipo I y poca al error tipo II.* Muchos investigadores prestan excesiva atención al error tipo I e ignoran el error tipo II. Sin embargo, en algunas investigaciones, el error tipo II podría ser más importante que el tipo I.

Esta crítica está justificada, pero, aunque las probabilidades de los dos tipos de error están inversamente relacionadas, hay una diferencia fundamental entre ellas. Mientras que la probabilidad de error Tipo I α es una constante, que puede elegirse antes de realizar el experimento, la probabilidad de error Tipo II es función del verdadero valor del efecto, que el investigador no conoce. Para resolver este problema, podría usarse el análisis de la potencia, donde, se hacen supuestos sobre los valores posibles del parámetro y se calcula la probabilidad de error Tipo II para estos diferentes valores. Sin embargo, no es común el análisis de la potencia (Borges, San Luis, Sánchez y Cañadas, 2001).

- *No podemos calcular la probabilidad de que la hipótesis sea cierta.* Esta crítica también es fundamentada, pues la significación estadística no informa de la probabilidad de que la hipótesis sea cierta (Lecoutre, 2006).

Sólo en la inferencia bayesiana podemos calcular probabilidades a posteriori de

la hipótesis y probabilidades de que el efecto tenga un cierto tamaño. Sin embargo, también el método Bayesiano tiene sus problemas, pues no hay acuerdo en cómo asignar las probabilidades iniciales, lo que da a los resultados un carácter subjetivo. Díaz y Batanero (2006) sugieren que una posible solución es comenzar la aplicación de los métodos bayesianos con distribuciones a priori no informativas (uniformes), lo que formalmente es equivalente al caso de equiprobabilidad para la probabilidad clásica. Sin embargo, también señalan las autoras, en muchas ocasiones, no tenemos información a priori, pero la hipótesis de equiprobabilidad no puede asumirse.

2.2.3. RECOMENDACIONES PARA MEJORAR EL USO DE LA ESTADÍSTICA EN PSICOLOGÍA

Las anteriores críticas han llevado a muchos psicólogos a replantear sus prácticas en el uso de la estadística en la investigación. Uno de los primeros fue Cohen (1994), quien concluyó que la interpretación errónea del nivel de significación está muy extendida y sugirió hacer un uso mayor del análisis exploratorio de datos, así como proporcionar intervalos de confianza, junto con los contrastes de hipótesis. También recomienda replicar las investigaciones, pues afirma que con los resultados de sólo una investigación sobre un tema, no se puede ser concluyente.

Recomendaciones de la American Psychological Association

La controversia sobre el uso de los métodos estadísticos se recoge también en el volumen 8, número 1 de la revista *Psychological Science*, en un número monográfico publicado en 1997, en el que varios investigadores dan su opinión sobre las pruebas de hipótesis. Una síntesis de las conclusiones en este monográfico son las siguientes:

- Hunter (1997) indica que, mientras la mayoría de investigadores suponen que el 5% de los contrastes serán erróneos, los estudios empíricos de la investigación publicada en psicología muestran que la tasa de contrastes erróneamente aceptados en las revistas psicológicas es aproximadamente del 60%.
- Abelson (1997) piensa que, mientras el uso de los contrastes como regla de decisión dicotómica lleva a muchos resultados erróneos, hay otros casos en que puede ser muy valioso, por ejemplo, para comprobar la bondad de ajuste de los datos a un modelo estadístico.
- Harris (1997) cree que las pruebas de significación todavía pueden ser útiles en

psicología, si en lugar de considerarse solamente dos alternativas (aceptar o rechazar la hipótesis nula) se realiza una prueba de tres alternativas: hipótesis nula de no diferencia, hipótesis de efecto positivo e hipótesis de efecto negativo.

- Estes (1997) sugiere que no debe imponerse ningún método estadístico, sino que se debe favorecer la adaptación de la metodología al avance de la psicología. También indica que los esfuerzos para mejorar la práctica de la estadística en psicología podrían empezar con la mejora de su enseñanza.

Además de estas sugerencias, la American Psychological Association creó una comisión específica: The Task Force on Statistical Inference. Después de un estudio, publican sus conclusiones (Wilkinson, 1999), donde sugieren describir adecuadamente los datos, incluyendo las omisiones o no respuesta y asegurarse que los resultados estadísticamente significativos no se producen debido a anomalías o valores atípicos o a problemas en la selección de los datos (Wilkinson, 1999). Asimismo hacen las siguientes recomendaciones sobre el uso de la estadística en psicología:

- *Uso del contraste de hipótesis:* Se recomienda las estimaciones de los efectos y los intervalos de confianza y no usar nunca la expresión “aceptar la hipótesis nula”.
- *Comparaciones múltiples:* Se deben usar los procedimientos específicos (por ejemplo el test de Tukey en Análisis de varianza) para tratar las situaciones en que tenemos que hacer varias comparaciones en la misma muestra.
- *Potencia y tamaño de muestra:* Además de proporcionar información sobre el tamaño de muestra, se ha de realizar, cuando sea necesario, el análisis de la potencia.
- *Población y muestra:* Definirlas con claridad, así como el procedimiento de muestreo o la forma de asignar sujetos a los tratamientos. También recomienda controlar las variables relevantes cuando no ha habido diseño aleatorio.

Otros autores sugieren complementar o sustituir la inferencia clásica con inferencia bayesiana, que proporciona una respuesta más ajustada que la clásica a las necesidades del investigador (Lindley, 1993; Lecoutre, 2006). Por un lado, la interpretación intuitiva, aunque incorrecta, que muchos científicos dan a las probabilidades frecuenciales asociadas con el contraste de hipótesis, sería correcta en la óptica bayesiana (Gingerenzer, 1993; Lecoutre, 2006; Díaz y Batanero, 2006). Por ello,

es de esperar menos errores de interpretación de resultados estadísticos bajo este paradigma.

Por otro lado, el estudio del efecto, que se recomienda como veremos con insistencia desde la Psicología, se puede obtener en forma sencilla a partir de la distribución a posteriori en los métodos bayesianos, que consideran el parámetro como variable aleatoria. El intervalo de credibilidad proporciona también los límites en los que el parámetro está incluido con una cierta probabilidad (Lecoutre, 2006). En inferencia bayesiana es posible calcular probabilidades finales de la hipótesis y probabilidades de que el efecto tenga un cierto tamaño (Lindley, 1993). Es posible también comparar la probabilidad del suceso observado bajo la hipótesis nula y bajo diferentes hipótesis alternativas. También podemos tener en cuenta la información previa, cosa que no es posible en inferencia clásica (Díaz, 2007).

2.3. ERRORES EN INFERENCIA ESTADÍSTICA

2.3.1. INTRODUCCIÓN

Nuestro trabajo también se relaciona con los errores de aprendizaje de la estadística a nivel universitario, especialmente los relacionados con inferencia, pues es sobre esta parte de la estadística donde se producen más errores en este nivel educativo. Sin pretender ser exhaustivos describimos a continuación aquellos que han recibido más atención en la investigación en didáctica de la estadística.

2.3.2. SIGNIFICACIÓN ESTADÍSTICA

El error más común en inferencia afecta a la interpretación del nivel de significación y el valor p , así como a la interpretación de la significación estadística. Una interpretación incorrecta del nivel de significación α (que se define como la probabilidad de rechazar la hipótesis nula siendo cierta), consiste en cambiar los términos de la probabilidad condicional que interviene en esta definición.

Así, Birnbaum (1982) ya indicó la interpretación incorrecta, pues se asume la siguiente definición: “*Un nivel de significación del 5% indica que, en promedio, 5 de cada 100 veces que rechazamos la hipótesis nula estaremos equivocados*”. De hecho, no podemos calcular la probabilidad de estar equivocados si rechazamos la hipótesis nula, a menos que trabajemos en inferencia Bayesiana (Lecoutre, 2006). El trabajo de

Birnbaum fue replicado posteriormente por muchos investigadores. Por ejemplo, Vallecillos (1994) encontró este error en su estudio con 436 estudiantes universitarios de diferentes especialidades (estadística, medicina, psicología, ingeniería y empresariales). Lecoutre (2006) encontró los mismos tipos de errores en profesores de metodología e investigadores en Psicología.

Respecto al p -valor (probabilidad de obtener el valor dado del estadístico, en caso de ser cierta la hipótesis nula) se piensa que este valor indica la probabilidad de que el valor obtenido del estadístico se deba al azar. No obstante, cuando rechazamos la hipótesis nula, no podemos inferir la existencia de una causa particular que llevó al resultado (Batanero, 2000). Por ejemplo, una diferencia significativa de las medias de un grupo experimental y otro de control puede ser debida a un tratamiento particular, pero también puede ser que la razón de las diferencias sea intrínseca a los grupos, de ahí la importancia del control de las variables extrañas.

La interpretación incorrecta del nivel de significación se une normalmente a la confusión entre *significación estadística* y *significación práctica*. Un resultado significativo sólo indica el porcentaje de veces que rechazaríamos la hipótesis nula cierta a la larga, cuando fuese cierta (*error tipo I*). La significación práctica implica un efecto experimental suficientemente elevado, es decir una diferencia de medias o correlación suficiente para fines prácticos (Díaz, Batanero y Wilhelmi, 2008). La significación estadística no es equivalente a la significación práctica, pues podemos encontrar datos estadísticamente significativos con un pequeño efecto experimental, siempre que tomemos una muestra grande (Frías, Pascual y García, 2000). Lecoutre (2006) indica que los psicólogos no comprenden esta diferencia y tienen una confianza excesiva en los contrastes, por lo cual, olvidan mencionar y considerar el tamaño de la muestra en la interpretación de los resultados significativos.

2.3.3. HIPÓTESIS Y LÓGICA DEL CONTRASTE

Es también usual confundir los diversos tipos de hipótesis que aparecen en la investigación y en el contraste estadístico (Vallecillos, 1994; Chow, 1996). Hay varias hipótesis implicadas en los distintos niveles de abstracción de la investigación experimental orientada a la confirmación de teorías:

- *Hipótesis substantivas o teóricas*: sobre constructos teóricos inobservables (e.g., El conocimiento que un estudiante tiene del lenguaje escrito favorece la comprensión

de textos);

- *Hipótesis de investigación*: Se deducen de las hipótesis substantivas, concretándolas en una variable relacionada más concreta, teóricamente observable, pero difícil de observar en la práctica (e.g., el conocimiento y el uso de estrategias para enfrentar un texto mejorará la comprensión de textos expositivos);
- *Hipótesis experimental*: Una hipótesis todavía más concreta sobre una variable relacionada con la hipótesis de investigación (e.g., si dos grupos de estudiantes de bachillerato reciben una enseñanza idéntica para la comprensión de textos narrativos, excepto en la enseñanza de la elaboración de resumen, la calificación que obtendrá en una prueba de comprensión lectora, el grupo que recibió entrenamiento para elaborar resumen será mayor);
- *Hipótesis estadística alternativa*: Una hipótesis referida a un resumen estadístico relacionado con la variable que interviene en la hipótesis experimental (e.g., la calificación media será mayor en los estudiantes que recibieron entrenamiento para elaborar resúmenes);
- *Hipótesis estadística nula*: Puesto que el contraste estadístico se basa en hipótesis nulas (no diferencia), esta es la negación de la hipótesis alternativa y es la hipótesis que usamos para el análisis estadístico (e.g., no existen diferencias en la calificación media de la comprensión lectora cuestionario en alumnos que fueron entrenados para elaborar resumen y quienes no fueron entrenados).

La teoría estadística se ocupa del último nivel (hipótesis estadísticas nula y alternativa), pero algunos estudiantes e investigadores suelen confundir todas estas hipótesis y cuando encuentran un resultado significativo, piensan que se refiere a la hipótesis de investigación o incluso a la substantiva (Chow, 1996). Es decir, si se rechaza que la media de los dos grupos de estudiantes en la prueba de comprensión lectora es igual, se deduce (injustificadamente) que el conocimiento del lenguaje (en general) o bien la cultura lingüística de uno de los grupos es mayor que la del otro.

2.3.4. INTERVALOS DE CONFIANZA

Los intervalos de confianza también tienen interpretaciones erróneas entre los investigadores. Cumming, Williams y Fidler (2004) estudian los errores de interpretación de intervalos de confianza, para el caso particular de la media. La

mayoría de investigadores en su estudio esperaban (erróneamente) una alta probabilidad de replicación, esperando que en una nueva muestra, la media cayera de nuevo en el intervalo de confianza original. Otra creencia errónea muy común en los investigadores fue que los intervalos de confianza de dos medias de muestras independientes son sólo significativamente diferentes cuando se tocan justo extremo con extremo. También confunden el cálculo de intervalos de confianza para medias independientes y relacionadas (Olivo, 2008).

2.4. ANÁLISIS EMPÍRICOS DEL USO DE LA ESTADÍSTICA EN LA INVESTIGACIÓN

2.4.1 Introducción

Nuestro trabajo puede también encuadrarse en la investigación de carácter metodológico, que examina el análisis del papel de la estadística y otros métodos en la investigación experimental. El estudio del uso de métodos estadísticos, su adecuación al problema, el empleo correcto y otros aspectos tiene algunos antecedentes, tanto en Didáctica de la Matemática, como en otras áreas, por ejemplo, Medicina o Psicología. A continuación se hace un resumen de algunas de estas investigaciones, que, están relacionadas con nuestro trabajo de investigación.

Batanero, Godino y Vallecillos (1992) analizan el papel de la estadística en la investigación, que, según ellos, está condicionado por las actitudes de los investigadores. Algunos piensan que cualquier investigación será buena con unas técnicas apropiadas de análisis, olvidando la importancia de la teoría como base del conocimiento. Lo peligroso de esta postura es que los paquetes estadísticos, como SPSS, permiten la aplicación de la estadística a personas sin conocimiento profundo de la misma. La Psicología y otras ciencias de la conducta tienen mucha dificultad en obtener nuevo conocimiento, debido a que son ciencias empíricas y deben apoyarse en identificar las relaciones entre muchas variables complejas, incluso cuando los métodos estadísticos se usan en forma adecuada. Si el empleo de las técnicas de investigación se hace inadecuadamente, la probabilidad de una conclusión errónea es bastante elevada.

Batanero, Godino y Vallecillos (1992) sugieren que otros autores rechazan el paradigma cuantitativo en casos donde sería conveniente aplicarlo, centrándose sólo en estudios de caso, en ocasiones en que sería necesario usar una muestra más amplia y el empleo de métodos estadísticos. Esta postura supone una ilusión de determinismo,

fallando en percibir la variabilidad de los fenómenos psicológicos o de otras ciencias sociales. Citan a Goetz y Lecompte (1988), quienes recuerdan que investigación cualitativa no es sinónimo de investigación anumérica, pues todo proceso de toma de datos supone un muestreo. Fijado el sujeto, se muestrean los tiempos, las circunstancias, las preguntas planteadas y el control de esta variabilidad o al menos su medición pasa por el uso de métodos estadísticos. Por otro lado, entre las técnicas aplicables al análisis cualitativo de datos, se van incorporando poco a poco algunas que clásicamente se han considerado como cuantitativas. Como ejemplo, Miles y Huberman (1984) en una exposición sistemática de tales técnicas mencionan algunas de este tipo, como la aplicación de técnicas estadísticas a las frecuencias obtenidas en las diversas categorías, el análisis cúmulo o análisis de correspondencias.

2.4.2. ESTUDIOS EN DIDÁCTICA DE LA MATEMÁTICA

En Didáctica de la matemática son pocos los estudios que se centran en el tema. En particular hemos encontrado el de Torralbo, Fernández, Rico, Maz y Gutiérrez sobre análisis metodológico de las tesis doctorales en Didáctica de la Matemática, el de Hill y Shih sobre métodos estadísticos en la revista *Journal for Research in Mathematics Education*; y los de Arteaga, quien analiza los proyectos estadísticos realizados por futuros profesores.

Estudio de Torralbo, Fernández, Maz, Rico y Vallejo

Estos autores (Torralbo, 2001; Torralbo, Vallejo y Fernández, 2001; Torralbo, Fernández, Rico, Maz y Gutiérrez, 2003; Torralbo, Vallejo, Fernández y Rico, 2004; Vallejo, Fernández, Torralbo y Maz, 2007) han llevado a cabo estudios sistemáticos de la producción sobre Didáctica de la Matemática, desde diversos puntos de vista. Por ejemplo, Torralbo y cols. (2003) estudian aspectos conceptuales que se investigan en las tesis doctorales a lo largo de veintidós años y algunas características cuantitativas de las mismas, con el fin de construir una visión general y sistemática de cómo se ha venido abordando la investigación en el área de didáctica de la matemática.

Se analizó el objeto de estudio de tales tesis en relación con la enseñanza y el aprendizaje de las matemáticas en cualquiera de los niveles educativos, con la innovación, el estudio y el desarrollo curricular, la formación del profesorado de matemáticas o la fundamentación teórica del área de conocimiento de didáctica de la matemática, para establecer si correspondían o no a la población elegida. Se localizaron

las tesis a través de la base de datos TESEO, centro de documentación Thales y consultas a algunas revistas españolas, así como por las citas cruzadas desde las tesis localizadas en primer lugar.

Entre otros aspectos, se analizó la productividad diacrónica, respecto a instituciones, departamentos y directores de tesis; los patrones de citas (revistas y autores citados), temática abordada (tipo de estudio y campo de aplicación). Este estudio, realizado a través del análisis conceptual de categorías temáticas, mostró que, pese a que en España la didáctica de la matemática como área de conocimiento es reciente, se ha investigado en sus diversos campos temáticos en la mayoría de las universidades españolas. Sin embargo, el autor cita 39 categorías temáticas que, en el momento del estudio no habían sido objeto de investigación, sugiriendo la posibilidad de crecimiento del área para llegar al desarrollo propio de cualquier disciplina científica. El estudio muestra el liderazgo de los departamentos de didáctica de la matemática y de didácticas específicas en educación matemática, desplazando a los departamentos de pedagogía, como consecuencia de la implantación de los estudios de doctorado en didáctica de la matemática a partir de 1988.

Torrallbo, Vallejo, Fernández y Rico (2004) analizan las características metodológicas de 135 tesis doctorales de Didáctica de la Matemática producidas en el periodo 1976-1998. Éstas, se localizan a través de la base de datos TESEO, mediante consulta a varias revistas de educación matemática que suelen publicar resúmenes de tesis doctorales y mediante análisis de citas cruzadas de las mismas tesis y literatura relacionada. Llevan a cabo un estudio mucho más amplio y detallado que en nuestro caso, analizando una serie de indicadores de las tesis, que organizan de acuerdo con las fases de la investigación, del modo siguiente:

- *Ubicación de la Investigación:* Paradigmas o enfoques metodológicos; Teoría desde o para la que se trabaja; Marco metodológico; Revisión de la literatura y referencias bibliográficas. Nosotros sólo consideramos el tema general del trabajo (que correspondería a la ubicación de la investigación).
- *Definición del problema:* Cuestión de la investigación o problema. Enunciado de los objetivos generales; Exposición de los objetivos específicos; Hipótesis; Definición de términos. En nuestro caso sólo se considera el establecimiento de hipótesis y los objetivos, así como la definición de las variables.

- *Diseño y trabajo de campo:* Uso de instrumentos. Instrumento específico. Instrumento no estandarizado vs. estandarizado. Validez del instrumento. Tipo de validez. Fiabilidad del instrumento. Procedimiento de fiabilidad. Unidad básica de análisis. Nivel académico. Identificación de la población. Tipo de selección muestral. Técnicas específicas de muestreo. Aleatorización de la muestra. Tipo de diseño general. Tipo de diseño específico. Temporalización. Amenazas al diseño. Control de las amenazas. En nuestro caso, en relación con la muestra, sólo se considera si se determina de manera específica; se revisa también, el tipo de instrumento utilizado, si es construido por el autor de la tesis, o se usa uno existente y, finalmente, los procedimientos para determinar la validez y fiabilidad de los instrumentos.
- *Tratamiento y análisis de los datos:* Estadísticos descriptivos. Tipos de estadísticos descriptivos. Uso de valores p . Inferencia estadística. Técnicas de inferencia estadística. Técnicas correlacionales y multivariadas. Tipos de técnicas correlacionales y multivariadas. Análisis cualitativo. Triangulación. En nuestro caso se revisa la cantidad diferente de procedimientos estadísticos utilizados, el tipo de procedimiento estadístico y las categorías de los métodos estadísticos. Finalmente se determina la correcta o no elección del procedimiento estadístico, considerando el tipo de estudio, nivel de medición y diseño.
- *Discusión de los resultados.* Hallazgos. Cuestiones abiertas. Implicación a la teoría. Implicación a la práctica. En nuestro caso se hace el análisis acerca de si se realiza o no una interpretación adecuada de los resultados obtenidos después de aplicar el procedimiento estadístico.

Vallejo (2005) realizó el análisis longitudinal de las tesis de doctorado en Educación Matemática defendidas en España de 1975 a 2002. Su objetivo fue describir la evolución del campo de estudio desde el nivel metodológico, cuantitativo y conceptual. La autora revisó 241 documentos, de cuyo análisis concluye, en el aspecto diacrónico-metodológico, que enunciar el problema es un elemento común y frecuente, por lo cual se puede suponer que en el futuro, aparecerá en las tesis de doctorado. Del mismo modo, supone que en los trabajos, se priorizará el planteamiento de objetivos generales sobre los específicos. Se observó un lento descenso en el planteamiento de

hipótesis, pero se puede predecir que en las tesis posteriores se continuará planteándolas.

En cuanto a la metodología utilizada, Vallejo menciona que es posible predecir el predominio de los estudios de corte cualitativo, sobre los cuantitativos. En relación con el instrumento, se observa que en la mayor parte de las tesis doctorales se utiliza el cuestionario, y continuará usándose en trabajos posteriores. Con los hallazgos encontrados no es posible suponer que habrá un mayor interés por atender las cuestiones de validez y fiabilidad de los datos. Finalmente, el alumno ha sido y será la principal unidad de análisis en las tesis de doctorado. Otras conclusiones de esta serie de trabajos, aplicables a nuestro estudio, son las siguientes:

- En las tesis analizadas se usan, en gran medida, instrumentos contruidos por los autores, pero no estandarizados.
- En una gran parte de las tesis no se estudia la validez y no se controlan las amenazas de validez o bien no se informa sobre este punto. Otro porcentaje alto tampoco estudia la fiabilidad del instrumento utilizado.
- Hay un claro predominio del uso de métodos estadísticos descriptivos frente a los correlacionales o inferenciales.
- Hay una contradicción de tesis que dicen trabajar en un estudio descriptivo o exploratorio, pero luego informan de los valores p (lo que indica una intención de generalizar los resultados).

El trabajo de Hill y Shih

Hill y Shih (2009) hacen un análisis de la calidad técnica de la investigación estadística publicada en *Journal for Research in Mathematics Education*, en una década (1997-2006). El análisis se hace considerando los criterios propuestos por American Educational Research Association, American Psychological Association, National Research Council on Education y American Statistical Association, indicando que la mayoría de los trabajos revisados carecen de uno o más de esos criterios. Se menciona que la investigación publicada que usa estadística comprende entre el 25-50% de la producción anual de JRME, aunque se observa un crecimiento del porcentaje, si se toma en cuenta las publicaciones recientes. Aclaran que los estudios cualitativos aportan valiosa información que permite enriquecer el conocimiento, pero el objetivo de su estudio es revisar los artículos que han realizado análisis estadístico.

Para el análisis se consideran cinco variables, cada una de las cuáles se relaciona con una recomendación para los investigadores: (a) uso del mismo nivel de análisis en todo el trabajo, (b) vinculación de los efectos del tratamiento respecto a los efectos del instructor, (c) inclusión de controles adecuados en los estudios cuasi-experimentales, (d) información acerca de la fiabilidad y validez de los instrumentos y (e) utilización adecuada de las escalas de medición.

Respecto al nivel de análisis, la ASA propone conservar la misma unidad de análisis en todo el estudio. Esto hace referencia a que en algunos estudios se mide individuos y posteriormente, en las conclusiones las medidas se interpretan en referencia a un grupo, que es una unidad estadística de nivel superior. Otras veces se mide un grupo y el contraste se hace entre individuos. También sucede que en el caso de medidas repetidas se utiliza el promedio de éstas para realizar el análisis. Se encontró que en los artículos revisados, sólo 5 de 15 conservaron el mismo nivel de análisis en la investigación.

En los trabajos en los que se lleve a cabo una intervención en el salón de clase o escuela, la ASA plantea que debe desligarse el efecto del tratamiento de la variación, que puede aparecer de manera natural. No se debe obviar que hay profesores más hábiles para trabajar con alguna innovación; asimismo, cuando la muestra es pequeña la variación juega un papel preponderante al “ocultar los efectos del tratamiento”. En los artículos revisados (47), no se hace mención al tamaño de las muestras ni a cómo se controló el efecto; se puede observar que en varios casos el efecto es producto de las características del instructor. Por ejemplo, se obtienen resultados diferentes en intervenciones con un número reducido de profesores, intervenciones en las que se compara el grupo del investigador (en su papel de profesor), con otro instructor, estudios que comparan pequeños grupos de voluntarios y estudios que comparan escuelas. En ninguno de los casos hubo evaluación previa que permitiera medir la eficiencia del profesor o la escuela.

En caso de utilizar diseños cuasiexperimentales es importante recordar que se requiere el control máximo de las variables que puedan incidir en los resultados. En escala de 0 a 7, en los estudios revisados, en promedio sólo en el 1.6 de ellos, se realizó un análisis de covarianza, que es el método estadístico recomendado para mantener este control.

Una de las principales recomendaciones de la AERA/APA/NCME (1999) es informar acerca de la validez y fiabilidad de los instrumentos de medición empleados en los estudios. En esto se incluye, no sólo proporcionar información acerca de los estudiantes evaluados, sino también sobre las variables independientes clave (afectos, conocimiento del profesor y alumnos, condiciones de la escuela, actitudes de los padres). En los trabajos revisados, sólo en ocho de 47 apareció información al respecto. Cuando aparece se hace exclusivamente de la *validez psicométrica* (análisis factorial y explicación de cómo se agrupan los reactivos). No se hace referencia a otros aspectos del proceso de aplicación, ni de variables.

Se observó también, que algunos de los autores de los artículos revisados realizan el análisis estadístico con los datos “en bruto”. Es decir, cuando se utiliza una escala o de un instrumento que requiere de un procedimiento de calificación, éste se obvia o se omite. En consecuencia, los resultados de los análisis son espurios. Asimismo, se encontró análisis estadísticos realizados con los valores porcentuales o con datos agrupados (organizados en tablas), lo cual genera conclusiones erróneas. Un último aspecto es que en algunos artículos, se realizó análisis estadísticos asumiendo que la escala de medición era de intervalo o de razón, cuando sus características métricas no permitían hacer ese supuesto.

Trabajo de Arteaga

Otras investigaciones que se centran en el uso de la estadística son las realizadas por Arteaga (2008) y Batanero, Arteaga y Ruiz (2010). El objetivo de estos trabajos es analizar los gráficos estadísticos elaborados por futuros profesores de educación primaria en el proceso de resolver un Proyecto de Análisis de Datos. Este trabajo se considera aquí, porque podemos suponer que un proyecto, al contener todos sus componentes -planteamiento de un problema, recogida de datos, análisis de éstos y obtención de conclusiones, es por sus características, en sentido limitado, una investigación. El interés no son los investigadores, sino los futuros profesores de Educación Primaria. El proyecto no es elegido por los estudiantes, como en nuestro trabajo, sino dado por el profesor, es común para todos los futuros profesores. No obstante, es un antecedente que podremos tener en cuenta en el futuro, en particular, si examinamos los gráficos producidos en las tesis de licenciatura.

Con la realización del proyecto, los autores pretenden indagar las concepciones

sobre la aleatoriedad de los futuros profesores y su comprensión de los gráficos estadísticos, tanto en el nivel de construcción, como en el nivel de lectura. Debe mencionarse que los estudiantes tienen libertad total para resolver el problema. En la solución del proyecto, una de las actividades es coleccionar datos durante la clase; con base en esto, los estudiantes comparan tres pares de variables estadísticas.

En el informe de Batanero, Arteaga y Ruiz el análisis se hace sólo sobre los gráficos estadísticos producidos. Para realizar lo anterior, se define un nivel de complejidad semiótica con lo cual es posible elaborar una clasificación de los gráficos y relacionarla con el nivel de lectura y las conclusiones elaboradas por los estudiantes. En la definición del nivel de complejidad se consideró el conocimiento de objetos matemáticos que se debe tener para construir el gráfico, así como el nivel de lectura que, de acuerdo con la siguiente clasificación, es posible comprender un gráfico (Bertin, 1967).

1. Extracción de datos, que consiste en leer literalmente un dato de la gráfica, poniendo en relación un elemento de un eje con el de otro eje.
2. Extracción de tendencias, cuando se es capaz de percibir en el gráfico una relación entre dos subconjuntos de datos que pueden ser definidos a priori o visualmente.
3. Análisis de la estructura de los datos, comparando tendencias o agrupamientos y efectuando predicciones.

De manera análoga, Batanero, Arteaga y Ruiz definen los siguientes niveles para calificar la construcción de gráficos:

- *Nivel 1. Representa sólo sus resultados individuales.* Algunos estudiantes producen una gráfica donde únicamente representan los datos obtenidos en su experimento particular, sin considerar los de sus compañeros. La gráfica elaborada permite únicamente un nivel de lectura de *extracción de datos* (Bertín, 1967)
- *Nivel 2. Representa los valores individuales de la variable sin llegar a formar la distribución.* Algunos estudiantes representan todos los datos en el orden en que fueron compilados, sin llegar a agrupar los valores iguales de la variable. El nivel de lectura de los gráficos que se ha clasificado en el nivel 2 es superior al de los de nivel 1, puesto que permite visualizar todos los valores obtenidos de la variable, así como su variabilidad (se observa claramente el valor máximo y mínimo). Sin embargo aún no es posible percibir claramente la estructura o tendencia de los datos.

- *Nivel 3. Produce gráficos separados para cada distribución.* El alumno puede representar un gráfico para cada una de las dos variables. Los gráficos clasificados en el Nivel 3, permiten responder preguntas hasta el nivel de *extracción de tendencias* en la categorización de Bertin (1967), ya que es posible percibir directamente una relación entre dos o más subconjuntos de datos.
- *Nivel 4. Produce un gráfico conjunto de las dos distribuciones.* Se alcanza un nivel más alto de complejidad cuando el estudiante logra formar las distribuciones de las dos variables y las representa conjuntamente en el mismo gráfico, estrategia que facilitará la comparación. Los gráficos clasificados en el Nivel 4 permiten el nivel superior de lectura en la categorización de Bertin (1967), es decir se puede hacer el *análisis de la estructura* pues permite comparar tanto tendencias como variabilidad en las dos variables en una única imagen.

Batanero, Arteaga y Ruiz mencionan que los resultados encontrados muestran que sólo parte de los participantes son capaces de producir un gráfico con la complejidad necesaria para obtener una conclusión y que una parte quienes lo consiguen, no alcanzan un nivel suficiente de lectura, o bien no logran una conclusión completa sobre la pregunta planteada. Los hallazgos informados muestran que la construcción e interpretación de gráficos es una habilidad compleja para los futuros profesores, y aparecen muchos errores, tales como el uso de las escalas (escala no proporcional; escala demasiado amplia, no se representan correctamente los números en la recta real) o en la elección de un gráfico inapropiado al tipo de problema.

Finalmente, los autores sugieren que en el proyecto planteado los estudiantes recorren todos los pasos del método estadístico, desde el planteamiento del problema, la definición de las preguntas, recogida y análisis de datos y obtención de conclusiones. También se pone en práctica el proceso de modelización, donde la realidad se simplifica. Además de trabajar con las variables aleatorias y estadísticas correspondientes, los estudiantes han de interpretar los resultados del trabajo matemático realizado con el modelo (distribuciones de datos obtenidas) en el contexto del problema (traducir estos resultados a lo que indican respecto de las intuiciones de los estudiantes). Sin embargo, los estudiantes fallan precisamente en este último paso, limitándose a presentar sus gráficos sin comentarlos o no extrayendo una conclusión correcta sobre el problema planteado. Por ello no son capaces de completar el ciclo de

modelización en el proyecto.

2.4.3. ESTUDIOS EN OTRAS DISCIPLINAS

Respecto a otras disciplinas, los estudios son más numerosos. Entre ellos resumimos a continuación los trabajos de Valera, Sánchez Meca y Marín (2000) y Tomcho y Foels (2000) en el campo de la psicología, el de Pimenta (2005, 2006) en Fisioterapia y el de Harraway, Manly, Sutherland y Mcrae (2001) en Biología.

Trabajo de Valera, Sánchez-Meca y Marín

Valera, Sánchez-Meca y Marín (2000) analizan la aplicación del contraste de hipótesis para comprobar si su uso es correcto en la investigación psicológica; realizan un estudio de la potencia de las investigaciones publicadas en revistas españolas. La potencia de un contraste $1-\beta$, es la probabilidad de rechazar erróneamente una hipótesis nula falsa. Para la selección de los trabajos a analizar se eligieron las revistas españolas que aplican pruebas estadísticas inferenciales y que han sido publicadas en España en el año 1991. En total se revisaron 18 revistas que incluían 573 artículos.

Para calcular la potencia de cada estudio utilizan como unidad de análisis el experimento. Como tamaños del efecto emplean los convencionales -bajo, medio y alto- y el tamaño del efecto estimado. Fijan como nivel de significación el 5%, para contraste bilateral. Controlan la tasa de error Tipo I con el procedimiento Bonferroni no ordenado. Y además de todos los cálculos de potencia para cada tamaño del efecto, en cada experimento se obtuvo el tamaño muestral necesario para garantizar una potencia adecuada de .80. Destacan por sus mayores niveles de potencia logrados, las revistas Boletín de Psicología y Revista de Psicología del Trabajo y las Organizaciones.

El análisis de los 169 experimentos identificados, con un total de 5,480 pruebas de significación estadísticas, arrojó valores de potencia de 0.18, 0.58, 0.83 y 0.59, para tamaños del efecto bajo, medio, alto y estimado, respectivamente. Estos valores descendieron drásticamente en un 20% aproximadamente cuando se repitieron los cálculos controlando la tasa de error Tipo I mediante el ajuste de Bonferroni. Estos resultados, por lo demás similares a los encontrados en los estudios de potencia internacionales, deben hacer reflexionar a la comunidad científica sobre la necesidad de dedicar especial atención al control de la potencia cuando se diseña una investigación.

Agrupando la información según la disciplina el área de Metodología la que alcanza los mejores niveles de potencia, aunque sólo 4 estudios pudieron catalogarse como metodológicos y además son estudios correlacionales (psicométricos) que emplean un número de sujetos alto. En cuanto al tipo de diseño, los estudios correlacionales consiguen los niveles de potencia más adecuados para tamaños del efecto estimados. Las investigaciones por encuesta, aunque consiguen niveles de potencia muy altos para los tamaños del efecto convencionales, dado el alto número de sujetos que suelen emplear, no alcanzan una potencia media adecuada para tamaños del efecto estimados. Es necesario tener en cuenta que sólo dos investigaciones por encuesta se pudieron incluir en el estudio y que los tamaños del efecto estimados en estas investigaciones suelen ser muy bajos.

Otro de los objetivos de este estudio era comprobar si los tamaños muestrales resultan adecuados. Se observó que por término medio se emplean 102 unidades de observación por experimento, mientras que deberían utilizarse 227. Si se considera la mediana como estadístico más robusto, se emplean 53 sujetos por experimento, mientras que deberían emplearse 103. Es decir, deberían duplicarse los tamaños muestrales habitualmente empleados para conseguir una potencia adecuada.

Una buena parte de los problemas en la aplicación del contraste de hipótesis se debe, según Valera, Sánchez-Meca y Marín, a las dificultades de cálculo de la potencia. Indican que hay dos grupos de propuestas sobre esta problemática: (a) Las que pretenden mejorar el modelo del contraste de hipótesis y, (b) las que proponen sustituirlo por considerarlo carente de sentido científico. Entre los métodos que pueden añadirse al contraste de hipótesis para mejorarlo, los autores sugieren el uso de intervalos de confianza ya que éstos ofrecen más información. Del mismo modo, nos permiten la aceptación o rechazo de una hipótesis estadística, nos indican de qué magnitud es el efecto y acotan la precisión de la estimación. Otra propuesta como complemento al contraste es el tamaño del efecto. Nos indica el grado en que la hipótesis nula es falsa, el grado en que el fenómeno está presente en la población. Los tamaños del efecto resultan además fundamentales en el análisis de la potencia, porque es uno de los parámetros de los que la potencia es función.

Trabajo de Tomcho y Foels

Tomcho y Foels (2010) analizan la potencia estadística en los artículos publicados en la revista *Teaching of Psychology* en el periodo 1974 a 2006. En total analizan 197 artículos publicados en el periodo. Los autores mencionan que el uso de contrastes con insuficiente potencia arroja incertidumbre en la decisión de aceptar o rechazar la hipótesis nula y en la interpretación de los resultados y, en consecuencia, recomiendan utilizar una potencia al menos de 0.80. Como se ha indicado, la potencia estadística es básicamente función del tamaño del efecto, nivel de significación y tamaño muestral, siendo este último el único que puede controlar el investigador.

Sólo tres, entre todos los trabajos analizados, realizaron un cálculo de potencia o lo mencionaron. Los autores, calculan la potencia esperada del resto de los trabajos publicados en la revista, cuando dichos trabajos informaban del tamaño de muestra y el tipo de contraste empleado, de modo que era posible el cálculo. Usaron un valor $\alpha=0.05$ como nivel de significación y cuando en un trabajo se realizaba más de un contraste se ajustó para que el nivel de significación global fuese el 5%. Con base en este estudio, los investigadores serían capaces de detectar un efecto importante en el 81% de los casos, uno mediano en el 22% y uno pequeño sólo en el 5% de los casos.

Respecto al tamaño de muestra se comparó los tamaños usados en los estudios con los aconsejados para el tipo de contraste usado y para obtener un tamaño del efecto importante, medio o bajo. El tamaño de la muestra fue suficiente en dos tercios de los estudios, que reportaron un efecto grande. Sólo el 20% de los que reportaron un efecto mediano (con la excepción de los estudios que sólo usan el test *t* de diferencia de medias), tenía un tamaño suficiente de muestra. Para los efectos pequeños casi ningún estudio tenía tamaño suficiente de muestra. Los autores recomiendan la necesidad de planificar el análisis de potencia en la investigación, si no se quiere tener serias limitaciones en sus resultados. Advierten también la necesidad de usar muestras más grandes. Los investigadores realizan múltiples contrastes de hipótesis sin mencionar el ajuste de la tasa de error. Debería considerarse el cambiar el nivel de significación para tener en cuenta las comparaciones múltiples o bien realizar sólo los contrastes más importantes.

Trabajo de Pimenta

En 2006 Pimenta realizó un estudio exploratorio descriptivo con el objetivo de evaluar el uso que los estudiantes de segundo ciclo de la carrera Terapia Física hacen de los métodos estadísticos, cuando elaboran su proyecto Fin de carrera. En este sentido, su trabajo es el más similar al nuestro, aunque los estudiantes son diplomados y no graduados como en nuestro caso. Igual que en nuestro estudio, la realización del proyecto es obligatorio para los estudiantes que se diploman; asimismo, en su trabajo, Pimenta examina sólo una muestra de todos los producidos en el periodo de interés. Asimismo le interesaba saber si los estudiantes utilizaban los procedimientos que aprendieron durante su formación en la universidad en las asignaturas de Fundamentos de Diseño Experimental, donde el autor era profesor. Para lograr el objetivo revisó 35 proyectos finales, los cuales presentaban algún análisis estadístico de los resultados.

El autor analiza los métodos estadísticos utilizados en los proyectos para compararlos con los enseñados en la asignatura citada. Un primer elemento es que en 34 (97.1%) de los 35 proyectos se utiliza la estadística descriptiva. Se observa también que en poco más de la mitad (51.5%) se utiliza algún procedimiento de estadística no paramétrica (Wilcoxon, Mann-Whitney; Kruskal-Wallis o chi cuadrado), en relación con esto destaca que no se utilice McNemar o Fisher. Únicamente el 10% usa la prueba *t*, pero la mayoría de este porcentaje utiliza la de dos muestras apareadas, lo cual indica que los estudios son de intervención y por lo tanto, se evalúa a los mismos sujetos en dos momentos distintos. Es importante mencionar que un porcentaje reducido utilizó ANOVA (5.7%) o correlación (5.7%); además sólo en un caso se utilizó un procedimiento que no se revisó en la formación de los estudiantes. La conclusión respecto a este punto es que los métodos enseñados fueron suficientes en casi todos los casos para realizar el proyecto Fin de carrera.

Otro punto estudiado fue la adecuación de los métodos estadísticos, tanto al tipo de variable (cuantitativa, cualitativa o mixta), como al fin del estudio (predicción, estudio de relación, de diferencia, etc.). Se enfatiza que en todos los casos los estudiantes utilizaron la prueba adecuada para el nivel de medición de su(s) variable(s). En relación con lo adecuado o no del método utilizado en función del tipo de problema estudiado en el proyecto, al considerar si se trataba o no de una comparación, o de una correlación; o bien de una o dos o más muestras, independientes o relacionadas, Pimenta informa que en sólo el 20% de los casos se observa un uso no adecuado. Sin

embargo, al valorar el uso adecuado del estadístico considerando sus presupuestos, el 31.4% no lo hizo de manera correcta. Concluye que los estudiantes no llegaron a comprender bien el fin de los diferentes métodos estadísticos estudiados.

Otro rubro que Pimenta revisó fue el planteamiento de hipótesis, para ver si se planteaban o no y si el planteo fue correcto. En este caso se observa que sólo en 31 de los trabajos se plantea hipótesis; de éstos en 5 (16.1%) se encontraron hipótesis mal planteadas, bien por no ser exhaustivas, o por referirse a la muestra en vez de la población o por confundir hipótesis nula con alternativa. En el rubro de interpretación de resultados, se observa que en el 40% de los proyectos se realiza una manera incorrecta. Esto, se debe principalmente a cuatro errores: (a) se acepta incorrectamente H_0 , cuando lo único que puede decirse es que no hay evidencia para rechazarla; (b) confusión entre el valor de la prueba (por ejemplo, el valor de la distribución normal o de la distribución t) y el valor p asociado; (c) en el caso de pruebas no paramétricas se concluye considerando las medias de las muestras, olvidando que se utilizó un procedimiento para contrastar medianas y (d) las conclusiones se hacen para una muestra y no para la población, a pesar de tratarse de estudios de inferencia.

Otro elemento analizado es la interpretación del nivel de significación, donde, como hemos dicho, se producen muchos errores. El principal error mencionado por Pimenta, es considerar que el nivel de significancia representa la probabilidad de que una hipótesis nula sea cierta, es decir, invertir la condición y el condicionado en la definición del nivel de significación. Pimenta informa que en ninguno de los trabajos se interpretó de manera correcta; en tanto, en el 25.7% se interpreta de manera incorrecta y sorprende que en el 74.3% no hace interpretación de ese dato.

Como resumen, el autor informa que se observa serias dificultades en la aplicación de los procedimientos estadísticos, en la elección correcta de una prueba de hipótesis y en respetar los supuestos de los procedimientos estadísticos, así como en su interpretación. Todo ello apunta a la necesidad de mejorar la formación estadística de los diplomados en fisioterapia en la universidad donde se tomaron los datos.

Trabajo de Harraway, Manly, Sutherland y Mcrae

Harraway, Manly, Sutherland y Mcrae, (2001) se preguntan sobre la necesidad de la estadística en la formación de futuros profesionales de la biología y de las ciencias de la salud. Se plantean que una estrategia para conocer la amplitud y profundidad del

conocimiento estadístico de los estudiantes de dichas disciplinas, es analizar el tipo de procedimientos estadísticos que se utilizan en los artículos que publican los investigadores y expertos de esos campos del saber.

Los autores revisan 2997 artículos publicados en 1998 y los 1999, en las 16 revistas de investigación de mayor impacto en la biología y las ciencias de la salud. En promedio, el 82% del total de artículos publicados en estas revistas incluye análisis estadístico. En la mitad de las revistas, el 92% o más de los artículos tuvieron análisis estadísticos, en tres de ellas el 96% de los trabajos. En contra parte, sólo en tres revistas se tuvo una tasa menor del 80% de trabajos con análisis estadístico, una de ellas con menos de la mitad (72%, 51% y 35%). Esto es un indicador de la enorme importancia que los editores de las publicaciones periódicas conceden al uso de la estadística en las investigaciones de Biología.

Los procedimientos y pruebas utilizadas en cada artículo se clasificaron primero en 60 clases; sin embargo, posteriormente se organizaron en 14 categorías. Entre ellas se destaca el análisis de varianza (ANOVA) y sus contrastes; la correlación y regresión; los procedimientos de estadística descriptiva; métodos multivariados; tablas de contingencia y métodos estadísticos de medicina. Los autores estudiaron la frecuencia de uso de cada una de estas categorías por año y revista.

Se observa también la preponderancia de algún tipo de procedimiento estadístico en cada una de las revistas, esto refleja el interés de los editores por conducir a los autores en el uso de ese método en particular. Así, por ejemplo, de 88 artículos analizados en la revista *Journal of Vegetation Science*, el 38% tuvo algún procedimiento de ANOVA; el 24% realizó un análisis post hoc. Las publicaciones de la *Journal of Botany and Plant Physiology* utilizan predominantemente métodos específicos, estimaciones, análisis espacial y, en ocasiones, ANOVA. En *Journal of Dairy Science* y *Journal of Sensory Studies and Food Quality and Preference* los investigadores prefieren utilizar el análisis multivariado. En las revistas de nutrición se utiliza la estadística médica, la regresión logística, los modelos log-lineales y la estadística descriptiva.

Con el objetivo de conocer la opinión de los profesores acerca del conocimiento de los estudiantes de biología y de las ciencias de la salud, de la universidad de Otago, se entrevistó a 46 de ellos, integrantes de 5 de los departamentos. Todos ellos fueron investigadores que publicaron en alguna de las 16 revistas analizadas. En primera

instancia se pidió elaborar una lista de los métodos estadísticos que ellos usan y se comparó con otra lista que los autores pensaban que los futuros profesionales del campo deberían conocer, observando una correlación alta entre ambas listas. Se observa un uso frecuente de la prueba *t*, ANOVA, comparación múltiple, regresión y correlación. En oposición, aparecen poco análisis con los métodos más especializados como correlación canónica, series temporales, análisis de correspondencias o métodos matriciales.

Siete de cada diez entrevistados consideraron que el conocimiento estadístico de los posgraduados era insuficiente para poder realizar investigación. En consecuencia, nueve de cada diez profesores piensan que debería haber cursos posteriores para los graduados. Asimismo, el 30% menciona que es importante llevarlos a revisar investigación durante su estancia en la universidad. Por último, se destaca que el 65% de los entrevistados utiliza Excel para realizar sus análisis, el 24% SPSS, MINITAB (30%) y SAS (20%). Este dato se corresponde con el hecho de que el 44% de los estudiantes que asiste a talleres especializados en matemáticas o estadística, se interesa en aprender estos programas.

Por último, Harraway et al, mencionan la existencia de tres aproximaciones para acercar a los futuros profesionales al conocimiento y uso especializado de la estadística.

1. Aprender informalmente de la lectura de artículos, con la asistencia de investigadores consolidados.
2. Asistir a cursos o talleres breves impartidos por expertos en estadística o investigadores prestigiosos de sus disciplinas.
3. Asistir a cursos de lectura de literatura científica impartidos por expertos en estadística o investigadores consolidados.

Los autores comentan que es más común la primera aproximación, pero reconocen que una mezcla entre la segunda y la tercera es la opción que mayor éxito puede garantizar. En cuanto a la elaboración de programas de cursos, Harraway et al (2001) sugieren realizarlos de manera colaborativa con un equipo de expertos en estadística y un grupo de investigadores del campo de estudio. En los programas se debe recuperar los procedimientos que se usa comúnmente en la disciplina, pero también debe considerarse los métodos novedosos de modo que los investigadores puedan utilizarlos y con ello generar nuevas y mejores maneras de usar la estadística.

CAPITULO 3. ESTUDIO EMPÍRICO

3.1. Introducción
3.2. Metodología del estudio
3.3. Muestra de trabajos analizados
3.4. Análisis de las tesis. Resultados.
3.4.1. Planteamiento del problema
3.4.2. Planificación
3.4.3. Datos
3.4.4. Análisis
3.4.5. Interpretación y conclusiones
3.4.6. Razonamiento estadístico de los investigadores
Conclusiones

3.1. INTRODUCCIÓN

En este capítulo se presenta un estudio empírico exploratorio del uso de la estadística en una muestra de tesis de licenciatura de Psicología Educativa. Dado el tiempo limitado que se ha tenido para llevarlo a cabo, sólo incluye algunas de las posibles variables que se podrían considerar en un trabajo más amplio. La finalidad es adquirir experiencia en el estudio de los documentos, en la definición de las variables pertinentes para el análisis del uso de la estadística en la investigación en nuestra área de conocimiento y en la codificación, análisis e interpretación de los datos. A continuación se describen la metodología, muestra de trabajos y resultados obtenidos.

3.2. METODOLOGÍA DEL ESTUDIO

Diseño. Se trata de un diseño descriptivo observacional que comprende la descripción, el registro, el análisis y la interpretación de hechos o procesos (Batanero, 2002). “El método de observación considera los fenómenos tal como se presentan sin modificarlos, ni actuar sobre ellos” (Bisquerra, 1989, p. 7). Su objetivo es interpretar la realidad, describiendo algunas características significativas de los fenómenos, a través de criterios objetivos (Batanero, 2002). Una observación se transforma en un método científico si: (a) permite lograr un objetivo de una investigación; (b) se realiza de manera sistemática y (c) se controla sistemáticamente y se relaciona con proposiciones más generales u otras investigaciones, que es lo que trataremos de hacer en este trabajo.

Variables. Se consideran las siguientes variables relacionadas con cada fase del ciclo de investigación (Wild y Pfannkuch, 1999; Pfannkuch y Wild, 2004), que se definen y operacionalizan en la sección de resultados:

- *Planteamiento del problema:* variables, objetivos e hipótesis.
- *Planificación:* Tipo de diseño y descripción del diseño.
- *Datos:* Población y muestra, instrumentos de recogida de datos, construcción de instrumentos, fiabilidad y validez del instrumento.
- *Análisis:* Número de métodos, tipos de métodos y corrección en la elección.
- *Interpretación y conclusiones:* Corrección de la interpretación.
- Respecto a los *modos de razonamiento estadístico* del citado modelo se consideran las siguientes: Reconocer la necesidad de datos, transnumeración, percepción de la variación, razonamiento con modelos estadísticos, integración de la estadística y el contexto.

Expectativas iniciales

Puesto que el trabajo es exploratorio y descriptivo no formularemos hipótesis en el sentido formal del término. Además, no se dispone de antecedentes que examinen las tesis de licenciatura en Psicología Educativa o Psicología ni en México ni en otras universidades. Podemos tener algunas expectativas, en función de lo encontrado en los trabajos referenciados en el Capítulo II; sin embargo, la formación metodológica, estadística y profesional de los autores de los trabajos analizados en dichas investigaciones son diferentes a nuestro caso, por lo que los resultados que obtengamos podrían variar.

Respecto al análisis de datos, una primera expectativa es que los métodos estadísticos que se utilicen en las tesis serán básicamente aquellos que se enseñan, en coincidencia con lo mencionado por Santisteban (1990) y lo encontrado por Harraway y cols. (2001) y Pimenta (2005). Coincidiendo con Pimenta (2005) esperamos encontrar dificultades en la aplicación de los procedimientos estadísticos, en la elección correcta de una prueba de hipótesis y la interpretación de los resultados.

Otra expectativa se refiere a las variables que hemos descrito sobre el planteamiento del problema, la planificación y recogida de datos. Esperamos que nuestros resultados estén por debajo de los porcentajes reportados para variables similares, por Torralbo y sus colaboradores, pues en nuestro caso se trata de estudiantes de licenciatura (no de doctorado) y la asesoría que reciben es menor que en aquél caso.

Finalmente, esperamos mejor aplicación de los modos fundamentales de razonamiento estadístico que lo encontrado por Pimenta (2005) pues nuestros alumnos

reciben mayor número de cursos estadísticos y metodológicos.

Instrumento y análisis de datos

El método de recogida de datos fue la observación directa de los documentos. El instrumento utilizado fue un protocolo de observación, en que se sintetizó los resultados de observación de cada una de las variables, con el objeto de preparar un fichero para el tratamiento estadístico.

Se realizó un *análisis de contenido* y se obtuvieron tablas de frecuencias de resultados. El análisis de contenido pretende alcanzar la sistematización, la objetividad y la medición cuantitativa de los hechos comunicativos (Hernández, Fernández y Baptista, 2007). Se fundamenta en la idea de que las unidades del texto pueden clasificarse en categorías (Weber, 1985; Krippendorff, 1991). Identifica y enfatiza los elementos relevantes y significativos del discurso verbal o no verbal (Álvarez-Gayou, 2003), que se integran en categorías de análisis. Es un análisis *directo* (Visauta, 1989), puesto que describimos el contenido sin realizar inferencias sobre el mensaje subyacente. Se siguen las siguientes cuatro fases:

- *Determinar la población y muestra.* Se estudiarán 65 tesis entre las producidas en el periodo 2000-2007.
- *Definir el campo de observación.* En el caso del presente estudio, aquellas secciones que se relacionen con características metodológicas y estadísticas de la tesis.
- *Decidir los elementos del contenido que se estudia.* Siguiendo las fases del ciclo de investigación en el modelo de razonamiento estadístico de Wild y Pfannkuch (1999) se identificarán las variables descritas con anterioridad.
- *Proponer una forma de recabar la información.* Se realiza una observación directa a partir de la lectura de los documentos. Revisamos repetidamente todas las unidades de análisis, con el objetivo de conseguir una categorización adecuada mediante un proceso analítico e inductivo, revisando con ayuda del director del trabajo, hasta llegar a la categorización definitiva. Se diseñó un protocolo en el cual se sintetizó la información recogida, además de disponer un espacio para hacer transcripciones o anotar las observaciones relevantes.
- *Codificación de los dato.* Se elaboró un fichero de datos que posteriormente se analiza, para obtener tablas de frecuencias de las diferentes respuestas, tablas resumidas de frecuencias y extraer conclusiones.

El análisis estadístico se reduce a la elaboración y descripción de tablas de frecuencias de resultados.

Validez y fiabilidad

Goetz y Lecompte (1988) indican la necesidad de revisar la validez y fiabilidad en el diseño cualitativo. Para evaluar la fiabilidad será necesario que un segundo codificador analice parte de los documentos y aplicar a las dos codificaciones un coeficiente o medida de concordancia. En nuestro estudio se revisó por segunda vez la codificación de los datos, resolviendo los desajustes, con ayuda del director del trabajo, pero no se hizo el estudio completo de la fiabilidad. Respecto a la validez el autor diferencia entre interna y externa. La *validez interna* indica si el investigador observa o mide realmente lo que cree. La *validez externa* trata del grado en que los estudios son aplicables a otro grupo. Respecto a la validez interna, se controlaron posibles riesgos como selección de la muestra (tratando de que fuese representativa de la población), sesgo del observador (estableciendo un protocolo objetivo de observación) y conclusiones espurias (puesto que no hemos establecido relaciones causa-efecto no se aplica). La validez externa depende de la identificación y la descripción de las características más importantes de los fenómenos con vista a la comparación de éstos con otros de tipo similar. En nuestro caso se describen con detalle las variables y categorías y se comparan con las de otros estudios similares.

3.3. MUESTRA DE TRABAJOS ANALIZADOS

Durante el periodo 2000-2007 (ver tabla 3.3.1) se presentaron 432 tesis de Psicología Educativa en la Unidad Ajusco de UPN, realizadas por los alumnos que se formaron con el plan 90.

Tabla 3.3.1. Trabajos concluidos y egresados titulados por año

Año	Trabajos presentados	%	Egresados Titulados	%
2000	27	6.3	53	7.3
2001	32	7.4	57	7.8
2002	49	11.3	73	10.0
2003	61	14.1	98	13.4
2004	59	13.7	102	14.0
2005	55	12.7	92	12.6
2006	66	15.3	111	15.2
2007	83	19.2	145	19.8
Total	432	100.0	731	100.0

Las tesis se pueden elaborar individualmente o en grupo, lo habitual es que el

trabajo lo elaboren uno o dos autores. Hubo un incremento constante en el número de tesis y de egresados durante el periodo, de modo que del año 2000 al 2007 se triplicó en ambos casos. Aunque hay un trabajo de asesoría, lo firma el egresado, y como los autores del trabajo están de acuerdo con su contenido, el análisis de estas tesis refleja la cultura estadística adquirida por los egresados, tanto en el uso de la estadística para apoyar sus trabajos, como en la interpretación de resultados estadísticos.

En el periodo analizado, los autores de los trabajos se distribuyen, por género (ver tabla 3.3.2), de manera semejante a la población estudiantil de la Universidad Pedagógica Nacional (Lozano y Rodríguez, 2006), con predominio del mujeres, como en la mayoría de las escuelas de psicología en México (ANUIES, 2008).

Tabla 3.3.2. Distribución de titulados por género por año

Año	Femenino	Masculino	Total
2000	48	5	53
2001	51	6	57
2002	65	8	73
2003	92	6	98
2004	97	5	102
2005	84	10	94
2006	94	15	109
2007	135	10	145
Total	666	65	731

Temas de las tesis

En la tabla 3.3.3 presentamos los temas abordados en los trabajos, clasificados con base en los contenidos revisados en los Seminarios de Tesis I y II (ver Anexo 1). Cada área de formación profesional acuerda al interior de su colectivo la caracterización del trabajo. Se presenta a continuación una breve descripción de los aspectos particulares de cada área de formación profesional.

En el *área de orientación educativa* se pretende que el alumno desarrolle los conocimientos y habilidades profesionales relacionadas con el diagnóstico, diseño, intervención y evaluación en la Orientación Educativa. Se intenta formar psicólogos educativos, cuya misión es incidir de manera eficaz en los ambientes educativos que requieran de su intervención psicopedagógica. También, su intervención es apreciada en los contextos educativos para optimizar procesos de enseñanza y aprendizaje, desarrollar habilidades cognitivas, fortalecer autoestima y autoconcepto, entre otros aspectos.

En el área de *problemas del escolar*, se pretende que al concluir los cursos el alumno cuente con los conocimientos teóricos y el manejo de los procedimientos para atender alumnos con necesidades educativas especiales (NEE), en cuanto a la evaluación psicopedagógica y el diseño y aplicación de adecuaciones curriculares, considerando el contexto escolar y familiar de los alumnos, promoviendo su integración educativa y la igualdad y equidad de oportunidad educativas.

La problemática educativa del área de *Aprendizajes Escolares* es estudiar el proceso de enseñanza y aprendizaje y disponer de materiales educativos adecuados a las características de los educandos mexicanos congruentemente con el currículum. Se estudia la orientación y propuestas psicopedagógicas que rigen en los currícula, los materiales educativos, los procesos de enseñanza y aprendizaje. Se propone hacer intervención en la enseñanza de los contenidos en áreas como español, ciencias naturales, geografía, historia, matemáticas, ciencias sociales y civismo; al mismo tiempo desarrollar y evaluar materiales educativos. Se pretende que los egresados cuenten con los conocimientos sobre las teorías, modelos y estrategias de enseñanza, aprendizaje y evaluación de contenidos, medios y materiales didácticos en dominios específicos y asimismo con las habilidades de intervención psicopedagógica para atender problemas de la práctica educativa.

El área de *Currículo y Educación* está enfocada en proporcionar un conocimiento sobre los procesos de diseño, desarrollo y evaluación. La formación ofrecida en esta área considera las tareas que son competencia real del psicólogo y su trabajo coordinado con otros profesionales de la educación, especialmente con el maestro de aula, los cuerpos académicos y el director. Los propósitos formativos de esta área corresponden a un conocimiento general acerca las distintas teorías que conforman la reflexión sobre la educación y el campo curricular, así como su aplicación en los campos citados, asumiendo los valores de la igualdad y equidad de oportunidades educativas para los alumnos y la promoción de una calidad en los servicios educativos.

Finalmente, en el área *grupos y organización escolar* las actividades del psicólogo educativo se dirigen hacia la organización escolar, el aula y el contexto cultural de los procesos de socialización y comunicación en la escuela. Se trata de propiciar y promover reflexión e intervención psicoeducativa basada en el estudio de problemas educativos socialmente significativos para apoyar el quehacer educativo de los maestros, el aprendizaje de los alumnos y la participación de los padres y madres de familia. El alumno, una vez concluidos los cursos de esta área, dispondrá de los

conocimientos teóricos y habilidades para analizar e intervenir en los espacios institucionales del centro escolar y coadyuvar en el desarrollo de proyectos en colaboración con la comunidad escolar.

Tabla 3.3.3. Temas que se abordan en las tesis

Temas	Frecuencia	Porcentaje
Orientación educativa	135	31.3
Problemas del escolar	123	28.5
Aprendizajes escolares	106	24.5
Currículo y educación	32	7.4
Grupos y organización escolar	23	5.3
Otros temas	13	3.0
Total	432	100.0

Un dato relevante es que los egresados eligen el tema para realizar su tesis de acuerdo con los tópicos revisados a lo largo de su formación, pues sólo el 3% de las tesis se elaboró en un tema distinto. En la tabla 3.3.3 se observa que el 80% de los trabajos se realiza en tres de los temas. Uno de cada tres trabajos (29.6 %) trata la Orientación Educativa en temas como: el papel del psicólogo educativo en un departamento de Orientación Educativa, apoyo en la elección vocacional y desarrollo de programas de atención a alumnos, padres o profesores en aspectos como sexualidad, consumo de drogas, violencia o relaciones familiares.

El segundo tema más estudiado es Problemas del Escolar (27.6%), con indagaciones sobre: Estudio de actitudes de profesores, alumnos y padres hacia la integración escolar de alumnos con discapacidad; intervención psicopedagógica o estrategias de integración escolar para alumnos con necesidades educativas especiales. Uno de cada cuatro trabajos se realiza con el tema de Aprendizajes Escolares. Debido a que los temas de trabajo del área son temas que los estudiantes trabajan a lo largo de su formación, además de ser importante en las prácticas, son frecuentemente elegidos para elaborar la tesis. Los tópicos trabajados son los procesos de enseñanza-aprendizaje de contenidos curriculares: lengua, matemáticas, ciencias naturales, ciencias sociales, etc.

En el área de Currículo y Educación el porcentaje de trabajos refleja también la cantidad de alumnos que seleccionan el tema para su formación en los dos últimos semestres. El área de Grupos y Organización Escolar es un tema menos atendido en las tesis. Finalmente, los trabajos de tesis que se ubican en otros temas abordan tópicos relacionados con la psicología clínica: terapia familiar, ausencia de la figura paterna o en tópicos relacionados con la educación de adultos o ancianos.

Muestra seleccionada

Los egresados de la licenciatura entregaron su versión impresa de la tesis y también se dispone en la página electrónica de la universidad (www.upn.mx). El presente trabajo empírico comenzó diferenciando las tesis que tenían algún tratamiento estadístico (ver gráfico 3.3.1). A partir de la revisión se encontró que el 72.5% de las tesis presentan algún tipo de análisis estadístico, lo cual indica la necesidad que estos licenciados tienen de conocimientos estadísticos. De este conjunto se seleccionaron las 65 tesis que se analizaron en este trabajo. Como criterio para seleccionarlas, se consideró que, además de utilizar los procedimientos de estadística descriptiva, se utilizara otro procedimiento de análisis estadístico. La selección tuvo en cuenta la distribución de las tesis por temas y por año en el periodo 2000-2007 (ver tabla 3.3.4).

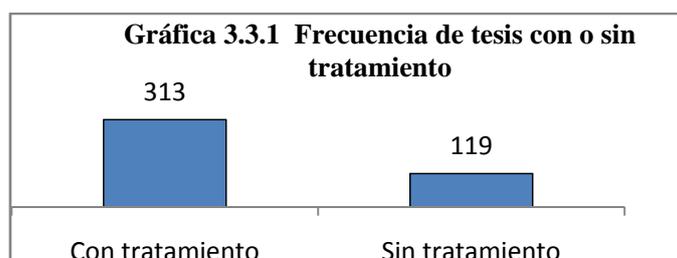


Tabla 3.3.4. Trabajos seleccionados y porcentaje por año

Año	Frecuencia	Porcentaje
2000	4	6.2
2001	5	7.7
2002	7	10.8
2003	10	15.4
2004	9	13.8
2005	8	12.3
2006	10	15.4
2007	12	18.5
Total	65	100.0

A continuación se analizan las tesis a partir de los componentes descritos por Wild y Pfannkuch (1999) en el ciclo de investigación de su modelo de razonamiento estadístico.

3.4. ANÁLISIS DE LAS TESIS. RESULTADOS

3.4.1 PLANTEAMIENTO DEL PROBLEMA

Como indicamos en la sección 1.4.1 la definición del problema es crítica, pues condiciona todo el proyecto. Hemos tenido en cuenta la explicitación de variables, el tipo de objetivos que se plantea y si las hipótesis están o no bien formuladas. A continuación se describen estos aspectos.

Variables

Un punto importante en la definición del problema es la explicitación de las variables, que son un concepto fundamental dentro de un proyecto de investigación pues son necesarias para la formulación de las hipótesis o los objetivos de la investigación. En una investigación una variable representa una propiedad o atributo que tiene distintos valores y es factible medir u observarse. Los científicos utilizan variables para operacionalizar los constructos o conceptos que les interesa estudiar en el desarrollo de una investigación (Kerlinger y Lee, 2002). Por tal razón es indispensable definir conceptual y operacionalmente las variables de una investigación. De este modo, se garantiza que podrán ser medidas, evaluadas, inducidas o inferidas; con ello será claro para los demás lo que se pretende hacer. Definir operacionalmente implica mencionar las actividades que un investigador deberá hacer para medir o recolectar datos, así como las técnicas o instrumentos que se utilizarán (Kerlinger y Lee, 2002; Hernández y cols., 2007). La correcta definición operacional debe considerar la adecuación al contexto y capacidad para evaluar los componentes de la variable. Existen dos clases de definiciones operacionales: (a) especificar cómo será medida y (b) describir la manera en que el investigador manipulará los valores de la variable. Finalmente es importante mencionar la clasificación distinción entre independientes y dependientes, activas y atributo y continuas o categóricas.

En el presente trabajo se considerará que una variable fue correctamente definida si en el documento se describe correctamente, siguiendo los criterios anteriores, cómo se va a medir o manipular y se clasifica de manera explícita, en alguna de las categorías mencionadas. En la tabla 3.4.1.1 se clasifican las tesis según la definición hecha de las variables. Se han considerado las siguientes categorías:

- *No se menciona*: Aunque se han utilizado variables, que aparecen en el análisis estadístico, explícitamente en el informe no se hace ninguna referencia a cuáles son

las variables, ni el tipo de relación que van a establecer o cómo las van a medir.

- *Se clasifican.* El egresado menciona las variables de manera clara y las ubica en alguna categoría (dependiente, independiente, etc.), pero no indica cómo pretende medirlas o, si es el caso, manipularlas.
- *Se operacionalizan.* Existe la descripción de cómo se hará la medición de las variables, o cómo las manipulará. Sin embargo, se omite mencionar el tipo de variable, esto es, no se establece cuál(es) es(son) la(s) independiente(s) y la(s) dependiente(s).
- *Se clasifican y se operacionalizan.* El autor de la tesis establece cuál es el tipo de variables que utilizará en el estudio y describe cómo se medirá(n). Esta clasificación y operacionalización es consistente en el análisis estadístico que realiza posteriormente.

Se observa en la tabla 3.4.1.1 que, considerando las tres últimas, en el 86% de los trabajos, los autores asumen mencionar, de alguna manera, las variables de sus trabajos. Las reconocen elementos fundamentales para definir de forma más clara los objetivos del estudio. La principal omisión fue no hacer su clasificación (3 de cada 10 de los autores de las tesis). En otros casos omiten explicitar cómo harán la medición de las mismas.

Tabla 3.4.1.1. Definición de variables

Unidad de análisis	Frecuencia	Porcentaje
No se menciona	9	13.8
Se clasifican	18	27.7
Se operacionalizan	22	33.8
Se clasifican y operacionalizan	16	24.6
Total	65	100.0

En una tercera parte de los trabajos aunque se menciona que uno de los grupos tendrá un tratamiento y el otro no, se olvida indicar cuál es variable independiente. La operacionalización más evidente se observa cuando en los trabajos se construyó un instrumento, sobre todo si el instrumento construido es una escala, porque los autores hacen una explicación sobre cómo calificar y obtener los puntajes de la variable dependiente. Uno de cada cuatro trabajos hace la definición conceptual y operacional de las variables. En contraparte, un reducido porcentaje omitió mencionar las variables de su estudio, aunque las consideran en el análisis de datos. Pimenta (2005) analiza el

número y tipo de variables en los proyectos fin de carrera, pero no la forma en que los egresados las operacionalizan o clasifican.

Objetivos

Kerlinger y Lee (2002) y Hernández y cols. (2007) resaltan la importancia de precisar el problema de investigación y definir con claridad sus objetivos. Al respecto Wilkinson (1999) hace énfasis que un informe debe establecer de manera clara y precisa el objetivo del estudio, sobre todo definirlo y establecer su prioridad. Según la finalidad del estudio, los objetivos pueden ser (Fortín, 1999):

- *Objetivos exploratorios o descriptivos:* Cuando la finalidad del estudio es la descripción de un aspecto poco conocido, por ejemplo: “Identificar los factores asociados al aprendizaje...”
- *Objetivos relacionales:* Cuando existe un conocimiento previo del campo de estudio y el objeto de la investigación es descubrir relaciones existentes y describirlas: “Describir la relación entre las actitudes hacia las matemáticas y el aprendizaje del cálculo”,
- *Objetivos analíticos:* Cuando los conocimientos que existen sobre un campo de estudio determinado, permiten predecir los resultados de la investigación. Se formulan entonces una hipótesis que supone la existencia de relaciones entre variables. La variable independiente, introducida por el investigador en el estudio, producirá un efecto sobre la variable dependiente, modificándola: “Determinar la influencia del uso de la tecnología en la mejora de la capacidad de resolución de problemas”.

En el presente estudio, debido al tamaño reducido de la muestra y dada la diversidad de contenidos abordados, metodologías seguidas, alcance y limitantes, no se va a diferenciar entre tipos de objetivos abordados, aunque esto puede ser tema de un trabajo posterior. Se considerará, además, indistintamente el planteamiento del problema en forma de preguntas de investigación, y el establecimiento formal de objetivos. De ese modo, en las tesis se calificó este rubro con tres niveles distintos:

- *No se plantean.* En el trabajo no se mencionan preguntas de investigación ni los objetivos que se persigue. Es por tanto difícil juzgar la utilidad de los resultados.
- *Se plantean de manera incorrecta.* El autor plantea de manera poco clara las preguntas o los objetivos que se pretende estudiar o bien, son inalcanzables, no

pertinentes o incongruentes con el estudio realizado y sobre todo con los análisis.

- *Se plantean correctamente.* Se hace explícito el planteamiento del problema en el que se menciona claramente el tipo de relación que se intenta estudiar o bien, los objetivos permiten saber qué se pretende. Son congruentes con el estudio realizado.

De acuerdo con la tabla 3.4.1.2, la mayoría (55.4%) plantea correctamente sus objetivos y son congruentes con el estudio que realiza posteriormente, por lo cual, puede suponerse que los egresados han conseguido una buena formación al respecto. Un 40% plantea objetivos que luego no llevan a cabo o los describen en forma poco clara y el resto no los plantea. Se observan porcentajes menores que los informados en la investigación de Torralbo y cols. (2004), donde el 85% de los trabajos enuncia al menos objetivos generales, aunque sólo el 58% menciona objetivos específicos. En su estudio longitudinal, Vallejo (2005) indica una constancia en la definición de objetivos generales y algo menor los específicos a lo largo del tiempo, con una tendencia a disminuir en el caso de los específicos a partir de los años 90 debido a mayor uso de paradigmas cualitativos. Estos autores no indican si el planteamiento de objetivos es o no correcto.

Tabla 3.4.1.2. Formulación de objetivos

Unidad de análisis	Frecuencia	Porcentaje
No se plantean	3	4.6
Se plantean de manera incorrecta	26	40.0
Se plantean correctamente	36	55.4
Total	65	100.0

Hipótesis

Una hipótesis es una afirmación acerca de una explicación posible del fenómeno que se analiza, se formula con base en el conocimiento teórico o empírico que el investigador tenga acerca de dicha relación (León y Montero, 1997; Kerlinger y Lee, 2002). Una hipótesis debe plantearse en forma contextualizada y específica, de manera comprensible, precisa y concreta y plantear una relación verosímil y medible de las variables, para que sea factible verificarla. La actividad que se realiza al confrontar hipótesis con fenómenos empíricos, es una de las más importantes de todo el sistema científico (Hernández y cols., 2007). En un estudio, según su tipo y su alcance, se podrá plantear una o más hipótesis. Hernández y cols. (2007); León y Montero (1997) indican

condiciones sobre las hipótesis que resumimos a continuación:

- *Coherencia*. Debe ser coherente en términos de una explicación razonable que resista un análisis crítico.
- *Generalidad*. La explicación es de carácter general y trasciende a una explicación o conjetura de hechos singulares.
- *Referencia empírica*. Ha de tener referentes empíricos, que posibiliten encontrar hechos concretos sobre los cuales se podrá luego corroborarla o refutarla.
- *Ser verificable*. La lógica científica afirma que lo que da valor a cierta hipótesis es permitir ser falseada. Es decir si luego de ser puesta a prueba en reiteradas situaciones para rechazarla, logra salir adelante sin objeciones.
- *Referencia teórica*. Es preciso que se inserte en un cuerpo de teoría en forma explícita, a fin de procurar incrementar el acervo científico. La ciencia es acumulativa y una hipótesis aislada no aporta nada. La hipótesis se fundamenta en teorías previas.
- *Clara conceptualización*, que permita identificar sin lugar a dudas cada uno de los términos que involucra y que destierre, hasta donde sea posible, toda vaguedad en el enunciado.

Como se indicó en el capítulo 2, de acuerdo a Chow (1996) hay también que diferenciar entre hipótesis de investigación y estadísticas. En nuestro trabajo sólo analizamos el planteamiento de hipótesis estadísticas. Las características del estudio determinan el tipo de hipótesis; así, el tipo de relación que propone puede ser correlacional o causa-efecto y puede implicar: una variable independiente (VI) y una variable dependiente (VD); una VI y varias VD; varias VI y una VD o varias VI y varias VD. En el presente estudio se consideró cinco categorías:

- *No se requiere*. Por las características del estudio, no es necesario ni pertinente plantear hipótesis. Se trataría de estudios exploratorios o descriptivos, estudios de caso o estudios de censo en que no hay interés en generalizar.
- *Requerido y no formulado*. Estudio que involucra algunas variables y debido al uso de la estadística inferencial, se requiera hacer uno o más contrastes de hipótesis. Pero el autor realiza el contraste y no menciona o explicita las hipótesis.
- *Planteamiento confuso*. Aunque se menciona la variable que se quiere contrastar y cuál es el parámetro medido, no se establece la relación que se espera entre las

variables. Por ejemplo, no se indica si se espera una correlación positiva o negativa o la intensidad aproximada o bien si la regresión es de tipo lineal.

- *Bien planteadas y no contrastadas.* Se plantea al menos una hipótesis en forma explícita, pero en ningún momento del análisis estadístico se realiza el contraste de hipótesis de las variables relacionadas en la misma.
- *Se formulan y se hace contraste correctamente.* En el trabajo existe el planteamiento adecuado de hipótesis experimental en la descripción metodológica, así como de las hipótesis estadísticas nula y alternativa al realizar el contraste correspondiente y, además, el contraste es correcto.

Tabla 3.4.1.3. Planteamiento de hipótesis

Unidad de análisis	Frecuencia	Porcentaje
No se requiere	3	4.6
Requerido y no formulado	12	18.5
Planteamiento confuso	24	36.9
Bien planteadas, no contrastadas	0	0.0
Bien planteadas y contrastadas	26	40.0
Total	65	100.0

En cuatro de cada diez trabajos (tabla 3.4.1.3) se plantea una hipótesis de manera correcta y después se hace el contraste o análisis correspondiente, indicando coherencia en la organización del trabajo. Esto se corrobora al observar que en ningún caso se planteó una hipótesis que no fuese contrastada posteriormente. Un alto porcentaje de trabajos plantea confusamente la hipótesis, pues no se establece de manera clara la relación que se espera encontrar en los resultados. Por ejemplo, se menciona sólo que habrá diferencias entre los grupos de participantes, o bien, que se encontrará una correlación entre las variables sin especificar la magnitud o dirección del efecto. Finalmente, el porcentaje de trabajos en los que no se plantea hipótesis porque no se requiere, debió ser mayor, pues cinco de los trabajos fueron de tipo exploratorio, pero algunos de estos autores plantearon y contrastaron hipótesis. Las hipótesis son planteadas en el 71% de las tesis analizadas por Torralbo y cols. (2004), quienes mencionan que no siempre el planteamiento es adecuado, pero no indican el porcentaje y tampoco si son correctamente planteadas o requeridas. Vallejo (2005) por su parte indica que con el transcurrir de los años se va igualando la proporción de trabajos que plantean o no hipótesis iniciales. Pimenta (2005) encuentra que el 85% de sus proyectos plantean correctamente las hipótesis. Esto es el doble que en nuestro estudio.

3.4.2 PLANIFICACIÓN

El mismo problema puede abordarse desde diferentes perspectivas. El diseño de un estudio es la búsqueda de una estrategia directamente relacionada con el problema, para abordar los objetivos y el estudio de la hipótesis. Son múltiples las clasificaciones de los diseños que aparecen en los libros de metodología. Para este trabajo sólo tendremos en cuenta la división del tipo de diseño respecto a la manipulación de las variables, considerando los siguientes tipos (León y Montero, 1997).

1. *Observación*: Las variables no se manipulan, sino que se observa el fenómeno tal como ocurre.
2. *Experimental*: Se manipulan deliberadamente una o más variables independientes (supuestas efectos), dentro de una situación de control para el investigador. Las muestras son aleatorias o se usa el emparejamiento al azar (aunque este método es menos preciso). Si se lleva a cabo con rigor, se tienen grupos grandes y se posee información que indica que los grupos no son diferentes, se puede lograr un alto grado de generalidad.
3. *Cuasi-experimental*. Los diseños cuasi experimentales también manipulan deliberadamente al menos una variable independiente, solamente que difieren de los experimentos "verdaderos" en la asignación aleatoria de tratamientos, por tanto en el grado de seguridad o confiabilidad que pueda tenerse sobre la equivalencia inicial de los grupos. En los diseños cuasi experimentales los sujetos no son asignados al azar a los grupos ni emparejados, sino que dichos grupos ya estaban formados antes del experimento, son grupos intactos.
4. *Descriptivo*: No se plantea estudio de tipo causa-efecto ni estudio inferencial. La finalidad es simplemente describir los datos de una muestra, incluidos en algunos casos la determinación de correlaciones.
5. *Ex post facto*. En este tipo de diseño, tanto los sucesos que produjeron (valores de la variable independiente) como los resultados que se tienen (efectos), ya acontecieron. En ese sentido, no es factible, de modo alguno, realizar alguna manipulación, cambio o ajuste ni a los valores de las variables independientes (si así puede llamarse), ni en las de las variables dependientes (no exactamente así).

La mayor parte de trabajos tienen diseño cuasi experimental (tabla 3.4.2.1), seguido del descriptivo, pues, en los contextos escolares en los que se llevó a cabo, es difícil tomar una muestra aleatoria. En ocasiones es posible hacer una asignación al azar

de dos o más subgrupos, lo cual incrementa el control de variables, pero el estudio permanece como cuasi-experimental. Una parte importante no menciona el diseño. Debe mencionarse que no se informa si hubo algún procedimiento para controlar las variables extrañas en el estudio, ni se dice cómo se atendió el posible efecto del experimentador, que en muchos de los casos fue el profesor del grupo o el egresado quienes atendieron al grupo experimental. Tal como mencionan Hill y Shih (2009), es común que estos aspectos no se hagan explícitos, en consecuencia los hallazgos del estudio adolecen de validez interna y externa.

Tabla 3.4.2.1. Tipo de diseño utilizado

Unidad de análisis	Frecuencia	Porcentaje
No se menciona	14	21.5
Cuasiexperimental	35	53.9
Descriptivo	15	23.1
Ex post facto	1	1.5
Observacional	0	0
Experimental	0	0
Total	65	100.00

Torralbo y cols. (2004) consideran el diseño general y el específico. En el 30% de los trabajos analizados no se indica el diseño general, y el más frecuente fue el pre-experimental, seguido del descriptivo y etnográfico. Son pocos los estudios que indican el diseño específico y dentro de ellos la frecuencia de los considerados en nuestro trabajo fue pequeña. El más frecuente fue también el cuasi-experimental, pero no llega al 8%. Conclusiones similares se obtienen en Vallejo (2005) y la tendencia se mantiene en el tiempo.

Se ha analizado también el modo en que los estudiantes definen el diseño, considerándose las siguientes categorías:

- *No se define.* El egresado no hace mención al tipo de diseño que se utilizará en el desarrollo del estudio.
- *Definición incorrecta.* En el estudio se menciona que se utilizará un tipo de diseño, sin embargo, el trabajo se realiza con uno diferente.
- *Definición correcta.* Se establece de manera clara y precisa la forma en la que se organizará a los participantes, la forma de asignarlos y los momentos en los que se realizará la medición.

La mayoría (casi un 70%) define correctamente su diseño, por lo cual puede suponerse la comprensión de la manera en la que organizarán a los participantes en las

distintas condiciones del estudio, así como en la forma en que se obtendrán los datos del trabajo (ver tabla 3.4.2.2). Sin embargo, en tres de cada diez trabajos o no se enuncia o bien no se menciona correctamente, aunque aún en estos casos, los autores de las tesis hicieron en general, una compilación de datos adecuada.

Tabla 3.4.2.2. Diseño descrito en las Tesis

<u>Unidad de análisis</u>	<u>Frecuencia</u>	<u>Porcentaje</u>
No se define	14	21.5
Definición incorrecta	7	10.8
Definición correcta	44	67.7
Total	65	100.0

Esto indica que la dificultad radica, quizá, en la percepción de la necesidad de informar acerca del tipo específico del diseño, o bien de tener una confusión acerca de la terminología. Esto contrasta con lo informado por Torralbo y cols. (2004) en donde en el 72.6% de las tesis de doctorado, por ellos analizadas, no se menciona el diseño específico. Torralbo y cols. (2004) estudian la descripción del tipo de diseño general, que nosotros no hemos considerado.

3.4.3. DATOS

Poblaciones y muestras

La mayoría de las investigaciones usan una muestra de sujetos o contextos y tratan de generalizar sus conclusiones. Lo que se busca al emplear una muestra es lograr que, observando una porción relativamente reducida de unidades, podamos obtener conclusiones semejantes a las que lograríamos si estudiáramos el universo total. Cuando una muestra cumple con esta condición, es decir, cuando nos refleja en sus unidades lo que ocurre en el universo, la llamamos *muestra representativa*. Sus conclusiones son susceptibles de ser generalizadas al conjunto del universo (Goetz y Lecompte; 1988). Será importante que el procedimiento de muestreo evite el sesgo para incrementar la posibilidad de generalizar los resultados. Una primera división que suele hacerse entre las muestras consiste en separarlas en muestras probabilísticas y no probabilísticas.

A veces no se puede tomar una muestra probabilística, y la muestra es intencionada, lo cual puede ser útil, aunque la posibilidad de generalizar es más limitada. Por ello el investigador debe describir las características del muestreo que

incluye: (a) La población objetivo, que es la teóricamente alcanzable y la población alcanzada: la que realmente se puede muestrear (b) Las unidades de muestreo: elementos, conglomerados, estratos y el tipo (aleatorio o no, estratificado, etc.). Sobre esta variable se han tenido en cuenta las siguientes categorías:

- *No se determinan*: En el trabajo no se menciona cuál es la población, ni tampoco se menciona cómo se seleccionó la muestra de participantes.
- *Determinación incorrecta*: Se observa confusión en la definición de la población y la muestra, en particular se manejan como sinónimos.
- *Determinación correcta*: Se hace la caracterización de la población y se menciona cómo se seleccionó la muestra.

En los datos de la tabla 3.4.3.1 se observa que la mayoría de los trabajos tiene una descripción adecuada de los participantes, así como de la manera en la que fueron seleccionados. Esto se relaciona con el número de trabajos en los que el diseño es cuasi experimental, en consecuencia, los autores sólo deben mencionar las características de la escuela, cómo se organiza los grupos y la forma en que se eligió los grupos. Empero, no en todos los casos cuando se utilizó este tipo de diseño, se hizo la descripción. Los estudios en los que no se determinan las características de la muestra, son aquellos en los que el número de participantes fue mayor, generalmente en los estudios de tipo descriptivo.

Tabla 3.4.3.1. Muestra

Unidad de análisis	Frecuencia	Porcentaje
No se determinan	18	27.7
Determinación incorrecta	11	16.9
Determinación correcta	36	55.4
Total	65	100.0

Puede suponerse que los egresados asumen que al ser un estudio que no se propone generalizar, es innecesario describir claramente la muestra con la que se trabajó. Torralbo y cols. (2004) y Vallejo (2005) separan estos dos puntos. En un porcentaje ligeramente superior de los trabajos revisados por Torralbo y cols. (2004), se hace la descripción de la población de estudio. La diferencia mayor radica en la cantidad de trabajos en la que no se identifica, ni se describe la población, 27.7% en nuestro caso, contra sólo el 3.7% en el caso de las tesis de doctorado en educación matemática. Los autores estudian el tipo de muestreo y tamaño muestral, aspectos que

no hemos considerado en este trabajo.

Instrumentos de recogida de datos

Un elemento fundamental es la identificación de instrumentos o técnicas para la generación de los datos con los que se podrá alcanzar los objetivos del estudio realizado. En todos los trabajos se encontró el uso de al menos un instrumento. Debe mencionarse que en la tabla 3.4.3.2 se contabiliza el número de trabajos en el que se informa haber utilizado ese tipo de instrumento y no la cantidad de instrumentos usados. En varios de los trabajos se utilizó más de uno y en un caso los egresados utilizaron ocho distintos tipos.

En la mitad de las tesis de la licenciatura revisadas para este trabajo, se informa haber utilizado un cuestionario de evaluación de conocimiento. Puesto que se trata de una licenciatura en psicología educativa, un número importante de tesis fueron estudios de tipo cuasiexperimental en el que se hizo una intervención psicopedagógica a un grupo escolar. La frecuencia alta del uso de la escala tipo Likert refleja el interés por estudiar la relación de las variables psicosociales (autoconcepto, autoestima, motivación de logro, asertividad), los estilos educativos y los hábitos de estudio con el aprendizaje. Un porcentaje reducido de trabajos utiliza alguna prueba psicológica.

Tabla 3.4.3.2. Tipo de instrumentos utilizados

Tipo de instrumento	Frecuencia	Porcentaje
Cuestionario de conocimientos	33	50.8
Escala tipo Likert	28	43.1
Cuestionario de opinión	12	18.5
Mapa conceptual	6	9.2
Prueba psicológica	4	6.2
Escala de desarrollo psicológico	3	4.6
Teorías implícitas	1	1.5

Torralbo y cols. (2004) indican que los autores prefieren construir sus propios instrumentos para las investigaciones. El instrumento más utilizado por la mayoría de los autores son las pruebas ad hoc, que han sido utilizadas en 66,7 % de las tesis doctorales, la entrevista ha sido utilizada en 43.7%, la observación en 31.9% De estos datos deducen una fuerte presencia de todos aquellos instrumentos que se utilizan en la metodología cualitativa. Vallejo (2005) destaca el uso de la encuesta (pensamos se refiere al cuestionario) indicando una caída del uso de pruebas no estandarizadas

(denominadas ad-hoc por Torralbo y cols.). Al comparar con nuestro se observa, en esos estudios, un porcentaje mayor de trabajos que utiliza lo que ellos denominan pruebas ad hoc.

Construcción de instrumentos

Se ha tenido en cuenta también si el egresado construye su propio instrumento o usa uno previamente construido. En la tabla 3.4.3.3 hemos contabilizado el total de instrumentos por categoría en las 65 tesis. El total es mayor de 65 porque algunos egresados usan instrumentos de más de una categoría. Las categorías consideradas son las siguientes:

- *Construido*. El egresado lleva a cabo el diseño y la construcción de un instrumento específico para su estudio.
- *Existente*. Se utiliza un instrumento que otros autores han construido y se informa de ello en la literatura.
- *Existente modificado*. Se usa un instrumento existente al cual se le hace adecuaciones para poder ser usado en la muestra de estudio.
- *Construido y existente*. Se usa un instrumento que se construye de manera específica y también su utiliza uno existente.
- *Técnica*. Trabajos que utilizaron la elaboración de un mapa conceptual para evaluar el conocimiento de los participantes.

Tabla 3.4.3.3. Construcción del instrumentos utilizado

<u>Unidad de análisis</u>	<u>Frecuencia</u>	<u>Porcentaje</u>
Construido	42	64.6
Existente	23	35.4
Existente modificado	11	16.9
Construyen y usan uno existente	11	16.9
Técnica	6	9.2

En la mayoría de las tesis el egresado construye el instrumento con el que obtendrá datos, lo que significa un proceso laborioso para el cual se requiere un periodo extra y, en varios casos resultó más complejo y extenuante que el estudio mismo. Esto se incrementa si, como sucedió en algunos trabajos, se construyó un instrumento y se modificó alguno existente. No obstante debe mencionarse que la mitad de esas modificaciones consistió en la traducción literal o en el cambio de palabras con un significado distinto en la cultura o contexto en el que el instrumento fue elaborado;

característica ésta de un alto porcentaje de estudios en psicología (Pelechano, 2002).

Fiabilidad y validez del instrumento

En todo trabajo es importante especificar su precisión (Fox, 1987). Las fuentes de error en la investigación pueden ser de naturaleza determinista y aleatoria. Los sesgos, de naturaleza determinista y magnitud desconocida, se derivan de nuestros procedimientos de investigación, selección de la muestra, elaboración de instrumentos y toma de datos. No disminuyen al aumentar el tamaño de la muestra y pueden ser evitados cambiando los métodos utilizados. La ausencia de sesgo se conoce como *validez* y es la principal característica deseable en un instrumento de recogida de datos. Los errores aleatorios son debidos a la variabilidad del material experimental y al aumentar el tamaño de la muestra pueden disminuirse y también mediante la precisión o *fiabilidad* del instrumento (Batanero y Díaz, 2009).

La validez y fiabilidad son conceptos teóricos, pero la psicometría ha desarrollado diferentes métodos para obtener una estimación. La fiabilidad se suele estimar mediante un *coeficiente de fiabilidad*, que es un indicador de la consistencia teórica de las puntuaciones observadas, en el sentido de proporcionar un valor numérico que indica el *grado de confianza* que podíamos tener en dichas puntuaciones como estimadores de las puntuaciones verdaderas de los sujetos (Batanero y Díaz, 2006a). El coeficiente más conocido es el de consistencia interna Alfa, pero también se usa la fiabilidad de la prueba repetida o test-retest, entre otras (Martínez, 1995).

Messick (1998) indica que la validación es un proceso continuo, puesto que cualquier cuestionario puede mejorarse en el futuro. Entre otros métodos, los habituales son el estudio de la validez de contenido, de constructo y discriminante. La validez de contenido, es decir, el grado en que el instrumento de evaluación refleja el dominio que nos interesa en forma satisfactoria. Mide la adecuación de los ítems de un test como muestra de un universo más amplio de ítems representativos del contenido (Martínez, 1995). Según esta autora la validez discriminante se define como la capacidad de discriminación del cuestionario con respecto a una variable independiente que esté relacionada con la variable que se trata de evaluar. La validez de constructo evalúa el grado en que la estructura de las respuestas refleja una estructura hipotética del constructo, generalmente se evalúa mediante el análisis factorial.

Hemos estudiado si el investigador informa sobre la validez y fiabilidad (con alguno de los métodos estándares). Se han encontrado las siguientes categorías:

- No se menciona. El egresado no menciona proceso de validación y de fiabilidad; por tanto no queda claro la naturaleza de la variable medida ni el grado de precisión de las conclusiones.
- Fiabilidad sin validez. Hay proceso de fiabilidad pero no de validez. Esta categoría es algo mejor, porque al menos informa sobre el control del error aleatorio. Pero no nos indica la presencia de sesgo ni tampoco la naturaleza de la variable que se mide
- Validez sin fiabilidad. Hay proceso de validez pero no de fiabilidad. En este caso, no hay información sobre la naturaleza de la variable ni sobre el error aleatorio, pero sí del control del sesgo en la medición.
- Fiabilidad y validez. Se define las características métricas del instrumento, hay proceso de fiabilidad y de validez. Se proporciona toda la información necesaria para hacer inferencias a partir del instrumento.

Casi a partes iguales encontramos que las categorías más frecuentes son o no mencionar ni validez ni fiabilidad o bien llevar a cabo sólo el estudio de validez. En estos últimos la mayoría de las veces por el método de validez de contenido (excepto en tres estudios en los que se usa validez de constructo), en algunos casos apoyado con un estudio piloto (ver tabla 3.4.3.4). Esta aproximación a la validez no requiere tanto estudio estadístico pues puede hacerse en forma cualitativa o bien mediante juicio de expertos que fue un método muy usado.

Tabla 3.4.3.4. Validación y fiabilidad

Unidad de análisis	Frecuencia	Porcentaje
No se menciona	26	40.0
Fiabilidad sin validez	4	6.2
Validez sin fiabilidad	25	38.5
Fiabilidad y validez	10	15.4
Total	65	100.0

Deducimos que el tema de fiabilidad fue difícil, bien en su comprensión o en el cálculo, puesto que requiere conocer los procedimientos habituales. Es probable también que haya influido el hecho de utilizar instrumentos contruidos, de los cuales se puede asumir que son por sí mismos válidos y fiables, por lo cual no es necesario mencionar estos aspectos de manera específica. Esto permite suponer que para los egresados de la licenciatura en psicología educativa, el concepto de fiabilidad es menos fácil de conceptualizar, o bien, que existe una confusión entre este concepto y el de

validez. Una explicación más es la asunción por parte de los egresados, que la validez de un instrumento es garantía de su fiabilidad.

En el estudio de Torralbo y cols. (2004) el porcentaje de tesis doctorales que no informa de algún proceso de validez es muy semejante (37%) y los que no mencionan el uso de un procedimiento de fiabilidad es aún mayor (60%). Como en nuestro caso, lo más frecuente fue el estudio de la validez de contenido. Sin embargo, el autor señala que, muchos autores usan dos métodos de validación, lo que no ocurre en nuestro caso. Solo el 40% de los casos analiza la fiabilidad, en general, mediante el método de consistencia interna. Vallejo (2005) indica que la preocupación por la validez crece, pero el método de validación sigue siendo muy pobre. Se confirma también la escasa preocupación por la fiabilidad.

3.4.4 ANALISIS

Número de métodos

El número de métodos estadísticos utilizados en los trabajos se presenta en la tabla 3.4.4.1, siendo lo más habitual dos o tres diferentes. Para los trabajos con dos más procedimientos, uno de los métodos estadísticos empleados es alguno de estadística descriptiva.

Tabla 3.4.4.1. Número de métodos estadísticos empleados

Unidad de análisis	Frecuencia	Porcentaje
1	0	0.0
2	41	63.1
3	15	23.1
4	4	6.2
5	3	4.6
6	2	3.1
Total	65	100.0

Tipos de métodos

En la tabla 3.4.4.2 se describen los métodos específicos utilizados en los proyectos. Lo más frecuente es el empleo de tablas estadísticas y cálculo de frecuencias y porcentajes. Se observa que en algunos trabajos se utilizó el cálculo de promedios, siempre asociado al uso de la prueba t, por lo cual puede asumirse que en el 100.0% de los trabajos se usa algún procedimiento de la estadística descriptiva.

En Torralbo y cols (2004) el 85.9% de los trabajos calcula estadístico

descriptivo, se usa la prueba de chi-cuadrado 34.8%; la *t* de Student en 20%; la bondad de ajuste en 12.6%; la *z* normal en 11.1%; procedimientos de comparaciones múltiples en 10.4%; y estadísticos no paramétricos un 5.2%, correlacionales bivariadas (36.3%), las propias de los índices de fiabilidad (29.6%) y el análisis factorial (23.7%). No analiza si se aplica correctamente o no. Vallejo (2005) indica un aumento de las técnicas descriptivas, inferenciales y correlacionales a partir de los años 90, señala el escaso uso de técnicas inferenciales multivariantes como MANOVA frente a las bivariantes, aunque sí se usan técnicas multivariantes correlacionales. En el estudio de Pimenta (2005) el 97% usa métodos descriptivos, el 20% el test de diferencia de muestras *t*; el 22% el test de Wilcoxon y el 14% el de Kruskal Wallis, el resto de los métodos tiene un uso puntual.

Tabla 3.4.4.2. Métodos estadísticos empleados

Unidad de análisis	Frecuencia	Porcentaje
Descriptivos, tablas y porcentajes	55	84.6
Contraste <i>t</i> de student	33	50.8
Test de Wilcoxon	13	20.0
Alfa de Cronbach	14	21.5
Coefficiente correlación Pearson	10	15.4
Cálculo promedios	10	15.4
Coefficiente correlación Spearman	8	12.3
Análisis factorial	6	9.2
ANOVA	5	7.7
Chi cuadrado	5	7.7
Tests U de Mann-Whitney	4	6.2
Test de Kruskal-Wallis	3	4.6
Tests Mc Nemar	1	1.5
Regresión logística	1	1.5

Se observa también en la tabla la frecuencia alta de pruebas de contraste entre dos o más muestras. Este hecho es consecuente con el mayor número de trabajos con diseño cuasiexperimental, que requieren al menos dos mediciones (previa y posterior) en una o más muestras. En cuanto al tipo de procedimiento estadístico, la lista observada en el presente se asemeja al informada en Pimenta (2005), aunque este autor no hace un estudio de la frecuencia. En el trabajo de Torralbo y cols. (2004), se usa con mayor frecuencia el ANOVA (48.1%), la prueba chi cuadrado (34.8%), análisis de factores y pruebas de fiabilidad y, con menor porcentaje, la *t* de student (20%). En los trabajos de licenciatura no aparece la bondad de ajuste.

En relación con lo que los egresados revisaron durante su formación profesional, tal como lo informa Pimenta (2005), resulta interesante mencionar la presencia de

procedimientos estadísticos que no fueron revisados en los cursos de estadística, ni de metodología, tal es el caso del análisis factorial y la regresión logística. Si los egresados han hecho uso de estos procedimientos se debe, como lo concluyen Harraway et al (2001), a la participación directa del asesor de la tesis, quien sí hace uso de esos procedimientos en sus trabajos y publicaciones. Es un caso similar en aquellos trabajos en donde se explicita el uso de SPSS para analizar los datos, puesto que en los cursos no se enseña el manejo de ese programa. En otro orden de ideas, algunos de los trabajos tienen variables cualitativas u ordinales o bien las muestras son pequeñas o no se cumplen los supuestos de los métodos paramétricos. Entonces recurren a métodos no paramétricos. En nuestro caso el uso de no paramétricos es inferior a los mencionados por Pimenta (2005), en oposición el uso de correlación y multivariados es bastante mayor.

Tabla 3.4.4.3. Categorías de los métodos estadísticos empleados

<u>Unidad de análisis</u>	<u>Frecuencia</u>	<u>Porcentaje</u>
Estadística descriptiva	63	96.9
Contraste t	33	50.8
No paramétricos	26	40.0
Correlación regresion	19	29.2
Multivariados	18	27.7
Confiabilidad	14	21.5
ANOVA	5	7.7

Corrección en la elección del método

Pimenta (2005) sugiere que la elección incorrecta de un método estadístico puede ser debida a varios posibles errores:

1. Adecuación del método al tipo de variables (según si son cuantitativas, cualitativas y ordinales). En ocasiones los licenciados usan un método que sólo es válido, por ejemplo para variables numéricas, con variables cualitativas u ordinales.
2. Adecuación del método al tipo de problema. Por ejemplo, si se quiere estudiar la relación entre dos variables. Se realiza un contraste de diferencia de medias, en vez de la correlación.
3. El método puede ser adecuado al tipo de problema y variable, pero no se cumplen los supuestos. Por ejemplo, para utilizar el test Chi-cuadrado en variables cualitativas la frecuencia en cada celda ha de ser mayor que 1 y al menos el 80% de las celdas tiene que tener frecuencia superior a 5.

En la tabla 3.4.4.4 presentamos la distribución de tesis, según elección y

aplicación del método, donde se han tenido en cuenta las siguientes categorías.

- Elección incorrecta: Sería los casos (1) y (2) descritos por Pimenta, bien que el método no es adecuado al tipo de variable o al tipo de problema.
- Elección correcta sin aplicación: El método de elige correctamente en función de la variable y tipo de problema. Aunque no los tuvimos en cuenta en la tabla anterior, como método usado, los citamos acá pues el estudiante indica que debe usarlo pero no lo aplica, posiblemente por no disponer de conocimientos de manejo de la paquetería. Finaliza aplicando un método incorrecto.
- Elección correcta y aplicación incorrecta, sería el caso (3) de Pimenta, es decir, el estudiante elige un método adecuado al tipo de problema y variable pero no se cumplen los requisitos, por ejemplo, no cumple la normalidad.
- Elección y aplicación correcta del método, en función del problema y las variables.

Tabla 3.4.4.4. Método elegido

Unidad de análisis	Frecuencia	Porcentaje
Elección incorrecta	13	20.0
Elección correcta sin aplicación	0	0.0
Elección correcta y aplicación incorrecta	3	4.6
Elección y aplicación correcta	49	75.4
Total	65	100.0

Los resultados indican que la mayoría hacen una elección y aplicación correctas, seguido de un 20% de elección incorrecta. Todo ello indica un buen aprendizaje por parte de los egresados. En el trabajo de Pimenta (2005) todos los métodos fueron adecuados al tipo de variable y el 80% eran adecuados al tipo de problema planteado.

3.4.5 INTERPRETACIÓN Y CONCLUSIONES

En la elaboración de la tesis los estudiantes recorren todos los pasos del método estadístico, desde el planteamiento del problema, la definición de las preguntas, recogida y análisis de datos y obtención de conclusiones. Recogidos sus datos, además de trabajar con modelos estadísticos (como contrastes de hipótesis), han de interpretar los resultados del trabajo con el modelo en el contexto del problema. Hemos analizado este último paso, utilizando las siguientes categorías:

- *No interpretar datos o interpretación incorrecta.* Cuando el tesista presenta resultados estadísticos, y no realiza una interpretación posterior de los resultados o bien la interpretación de es incorrecta, pues no se apoya en la evidencia empírica

obtenida.

- *Interpretación incorrecta del significado de los valores obtenidos:* Cuando se confunden valores de los estadísticos calculados o resultados obtenidos de los métodos estadísticos, por ejemplo, cuando se confunde el p-valor con el valor t o z de la prueba estadística. Por ejemplo, si un resultado significativo se toma como no significativo o viceversa o que se estima incorrectamente el valor del efecto.
- *Confusión de conceptos estadísticos.* Algunos tesisistas confunden conceptos estadísticos y en consecuencia los interpretan en forma incorrecta en sus informes de investigación. Por ejemplo, calculan un coeficiente de correlación e interpretan el resultado significativo indicando que hay diferencia entre las medias de las dos variables (en lugar de indicar una relación creciente o decreciente).
- *Interpretación de manera correcta.* Cuando muestran conocimiento de los conceptos e interpretan correctamente en el problema los resultados obtenidos.

Tabla 3.4.5.1. Interpretación de los resultados de las pruebas estadísticas

Unidad de análisis	Frecuencia	Porcentaje
Interpretación incorrecta o no interpreta	11	16.9
Interpretación incorrecta de los valores obtenidos	5	7.7
Confusión de conceptos estadísticos	17	26.2
Interpretación correcta	32	49.2
Total	65	100.0

En la tabla 3.4.5.1 se presentan los resultados, que de nuevo, mayoritariamente son correctos, evidenciando un buen aprendizaje y capacidad de interpretación de los resultados estadísticos. Hay un 26% de estudiantes que confunden conceptos; en un trabajo posterior trataremos de determinar cuáles son las confusiones confundidos comunes. Un 17% presenta los cálculos sin interpretar. Torralbo y cols. y Vallejo (2005) no estudian si la interpretación de los métodos estadísticos es correcta, sólo se revisa si se hace exposición de conclusiones. Al respecto, indican que casi la totalidad de las investigaciones describieron sus hallazgos o logros (98.5%). Y sólo en dos de ellas no han descrito dichos logros (1.5%). En Pimenta, sólo el 60% de los estudiantes realiza interpretación correcta de sus análisis estadísticos. El 25% de los estudiantes en el trabajo de Arteaga (2008) se limita a producir gráficos sin interpretarlos y el 70% aproximadamente, no llega a obtener una conclusión sobre el problema planteado o bien obtiene una conclusión incorrecta.

3.4.6 MODOS FUNDAMENTALES DE RAZONAMIENTO ESTADÍSTICO

Las variables analizadas hasta este punto nos proporcionan indicadores del razonamiento estadístico de los egresados en el primer componente del modelo de Wild y Pfannkuch (1999). En este apartado se estudian algunos indicadores del razonamiento estadístico del segundo componente de dicho modelo, comparando los resultados con los de Pimenta (2005). Más concretamente, nos centramos en la evaluación de los cinco modos de razonamiento estadístico fundamentales incluidos en este segundo componente, que se describieron en la sección 1.4.2.

Reconocer la necesidad de los datos

Sobre este componente se han considerado cuatro indicadores: (a) si el alumno recoge datos personalmente para su tesis; (b) si caracteriza correctamente la población y muestra utilizadas en el estudio; (c) si construye o (d) si usa instrumentos construidos de recogida de datos. Todas estas categorías se relacionan con la percepción de la necesidad de los datos y fueron usadas por Pimenta (2005).

Tabla 3.4.6.1. Frecuencia y porcentaje de reconocimiento de la necesidad de obtener datos

Necesidad de datos	Sí	Porcentaje
Recogen datos	65	100
Caracteriza correctamente la población y muestra	36	55.4
Construyen instrumentos	53	81.5
Usan instrumentos construidos	45	69.2

En la totalidad de las tesis analizadas se recogen personalmente datos adecuados para resolver el problema (Tabla 3.4.6.1) al igual que en el caso de Pimenta (2005). Un 55% también describe adecuadamente la población muestra, con objeto de reconocer el alcance de su estudio, frente al 85% en Pimenta. La mayoría construye algún instrumento para obtener sus datos, frente al 23% en el estudio de Pimenta, o usan uno ya construido. Todo ello es indicio de que los egresados reconocen la necesidad de recoger datos en sus investigaciones.

Transnumeración

Consideramos que ha habido transnumeración si los alumnos transforman los datos brutos y los representan en una tabla o gráfico estadístico, o si presentan un resumen de los datos adecuado al tipo de problema, bien construido (Tabla 3.4.6.2). Todos estos procedimientos permiten extraer una conclusión no claramente visible de los datos originales. Casi la totalidad de nuestros egresados utiliza tanto tablas, como gráficos

para representar sus datos. Hay también un predominio de medidas de posición central, lo que parece indicar que esta idea se comprende mejor que las de dispersión o correlación que se usa con menor frecuencia. Los porcentajes informados por Pimenta (2005) en cuanto al uso de las medidas de tendencia central (70.6) son semejantes a las aquí informadas; no es el caso de las medidas de dispersión (55.6%). Debe mencionarse que Pimenta hizo la evaluación acerca de la interpretación correcta tanto de las tablas y los gráficos como de los estadísticos.

Tabla 3.4.6.2. Frecuencia y porcentaje de transnumeración

Trasnumeración	Sí	Porcentaje
Construyen tablas	61	93.9
Calculan correctamente estadísticos de tendencia central	46	70.8
Calculan correctamente estadísticos de dispersión o correlación	19	29.2
Representan los datos a través de gráficos	53	81.5

Percepción de la variación

Como indicadores en este punto se consideran la clasificación y operacionalización correcta de las variables (implica reconocer variabilidad de las medidas), cálculo de fiabilidad y validez (que implica reconocer variabilidad de los datos) y uso de métodos inferenciales (variabilidad del muestreo). Los egresados reconocen la variabilidad en la muestra pues usan métodos inferenciales en su mayoría (Tabla 3.4.6.3). Sin embargo, son pocos los que explícitamente operacionalizan las variables del estudio, por lo que este punto debe ser considerado en la propuesta educativa.

Mientras muchos hacen consideraciones respecto a la validez de los instrumentos, son pocos los que hacen el estudio de fiabilidad, por lo que parece no reconocerse la variabilidad aleatoria en los datos y sí el sesgo constante. Todo ello sugiere el predominio de la formación determinista del estudiante y la necesidad de reforzar su formación estocástica. En el trabajo de Pimenta (2005) se informa de resultados semejantes, el 91.2% de los estudiantes perciben la variación con alguno de estos indicadores, y no todos tienen claro la incertidumbre de sus resultados.

Tabla 3.4.6.3. Frecuencia y porcentaje de percepción de la variación

Percepción de la variación	Sí	Porcentaje
Se clasifica y operacionaliza las variables	16	24.6
Calculan Fiabilidad	14	21.5
Calculan Validez	35	53.8
Uso de métodos inferenciales	64	98.5

Razonamiento con modelos estadísticos

Hemos considerado al igual que Wild y Pfannkuch (1999) como modelo cualquier método estadístico, considerando como indicadores, el uso de los mismos, en general y particularizados, así como su elección y aplicación correcta. Puesto que establecer una hipótesis supone trabajar con un modelo y contrastarlas, implica el uso del modelo de contraste, hemos considerado como otro indicador el establecimiento y contraste adecuado de hipótesis.

En la tabla 3.4.6.4 se observa que en la totalidad de los trabajos se utilizan los modelos estadísticos. Al clasificarlos por tipo, casi la totalidad son de tipo descriptivo o inferencial, y en uno de cada tres un modelo correlacional o multivariado. Al estudiar la corrección, observamos que el 75%, elige y aplica el modelo estadístico adecuado. En el caso de Pimenta (2005), el 97.1% aplica procedimientos de estadística descriptiva; asimismo, en uno de cada cinco casos se usó la prueba no paramétrica de Wilcoxon (22%), y un porcentaje similar (20%) la t de student para muestras pareadas. En casos aislados se usan las pruebas de Kruskal-Wallis, Mann-Whitney, y coeficientes de correlación de Pearson y Spearman. Los estudiantes de Pimenta no consideraron, en general, los supuestos de los métodos aplicados. Esto se hizo notorio al realizar el contraste de hipótesis e incluso al establecer correctamente las hipótesis del estudio, tal como se observó en el 25% los casos del presente estudio.

Tabla 3.4.6.4. Frecuencia y porcentaje de razonamiento con modelos estadísticos

Razonamiento con modelos estadísticos	Sí	Porcentaje
Uso métodos estadísticos	65	100
Hipótesis correctamente establecidas y contrastadas	26	40.0
Modelos descriptivos	63	96.9
Modelos inferenciales	64	98.5
Correlacionales	19	29.2
Multivariados	18	27.7
Elección y aplicación correcta del modelo	49	75.4

Integración de la estadística y el contexto

Finalmente analizamos si los egresados son capaces de integrar los resultados estadísticos obtenidos en el contexto de su tesis (Tabla 3.4.6.5). En la mayoría de los trabajos se presenta esta integración en mayor o menor medida, puesto que en las tesis se plantean correctamente los objetivos en el apartado de método. El diseño también está especificado y en un gran porcentaje correctamente elegido y descrito. El punto más débil es la correcta interpretación; este resultado coincide con lo obtenido por

Arteaga (2008). El autor sugiere que la enseñanza de la estadística prepara al estudiante para aplicar los métodos, pero no para obtener conclusiones a partir de ellos, lo cuál podría ser un punto para mejorar la enseñanza en nuestra licenciatura.

Tabla 3.4.6.5. Frecuencia (y porcentaje) de integración de la estadística en el contexto

<u>Integración de la estadística en el contexto</u>	<u>Sí</u>	<u>Porcentaje</u>
Objetivos bien planteados	36	55.4
Indican el diseño	52	80.0
Diseño correcto	44	67.69
Interpretación correcta	32	49.2

En el trabajo de Pimenta (2005) la integración de la estadística en el contexto se evalúa a través del uso de graficas, medidas de tendencia central y medidas de dispersión. El autor menciona que en general los estudiantes sí hacen una integración adecuada, aunque como ya se comentó antes, el problema principal radica en la elección adecuada del procedimiento estadístico y su interpretación, como en el presente estudio.

4. CONCLUSIONES

Como se indicó en el primer capítulo, en este trabajo hemos iniciado una investigación con una componente metodológica, a la vez que con una componente didáctica. Desde el punto de vista metodológico, el trabajo se orienta a iniciar una evaluación del uso de las herramientas que los estudiantes han adquirido en su formación del área metodológica en los trabajos de Tesis de licenciatura de la titulación en Psicología Educativa Universidad Pedagógica Nacional en México, D.F. Desde el punto de vista didáctico, se trata de encontrar criterios para mejorar dicha formación. A continuación exponemos nuestras conclusiones y algunas ideas para seguir la investigación.

Conclusiones respecto a objetivos y expectativas

Para este trabajo señalamos dos objetivos principales:

- El primer objetivo fue *iniciar la elaboración de un estado de la cuestión* sobre la problemática del uso adecuado de métodos estadísticos en la investigación (con particular énfasis en la investigación psicológica).

Este objetivo se abordó en el Capítulo 2, donde iniciamos una síntesis sobre esta problemática que será ampliada en el futuro. Con base en la revisión de la literatura es posible plantear algunos aspectos. En primer lugar, puede mencionarse que el uso que los profesionales hacen de la estadística, es un tema ampliamente discutido en la literatura, en distintos campos disciplinares, entre ellos, el campo de la educación matemática y el de la psicología.

Resulta significativamente relevante encontrar trabajos que elaboran análisis del uso de la estadística en los artículos publicados en revistas de difusión o divulgación científica. Aún más significativo es encontrar que algunos investigadores enfrentan dificultades semejantes a los hallados en trabajos de profesionales con un menor nivel de experiencia en la difusión de los resultados de la investigación. En ese sentido, se puede argumentar que éste es un tema que involucra a los profesores de estadística y métodos de investigación.

En el campo de la educación matemática, se ha hecho estudios para valorar aspectos metodológicos de las tesis doctorales o de artículos en revistas de investigación, así como proyectos estadísticos de futuros profesores. Sin embargo pocos

de estos trabajos han tenido una finalidad de mejora de las asignaturas metodológicas. En este sentido, pensamos que se abre una línea de investigación al respecto.

- *El segundo objetivo fue iniciar un estudio limitado* de evaluación del uso de las herramientas metodológicas y estadísticas en las Tesis de Licenciatura.

Este objetivo se aborda en el capítulo 3. Aunque los resultados no son generalizables, debido al tamaño de la muestra, con los hallazgos mencionados sí es posible elaborar algunas consideraciones, si bien son limitadas, servirán como base para continuar una posterior investigación. Respecto a este objetivo se plantearon varias expectativas iniciales:

Una primera expectativa era que los métodos estadísticos que se utilizan en las tesis serán básicamente aquellos que los egresados aprendieron durante su formación profesional, en coincidencia con lo señalado por Santisteban (1990) y lo encontrado por Harraway y cols. (2001) y Pimenta (2005). Esta expectativa se ha confirmado con los datos pues fueron pocos los estudiantes que utilizaron algún método no explicado en clase, como, por ejemplo, el análisis factorial. Por un lado, estos resultados indican que los métodos estadísticos enseñados fueron útiles para los alumnos. Se podría pensar también en ofrecer servicio de consultoría estadística a aquellos estudiantes que deseen aplicar un método no enseñado durante la licenciatura.

Coincidiendo con Pimenta (2005), esperábamos encontrar dificultades en la aplicación de los procedimientos estadísticos, en la elección correcta de una prueba de hipótesis y la interpretación de los resultados. Se observaron resultados similares a los informados en Pimenta en relación con el uso de los procedimientos estadísticos. Los estudiantes utilizaron la mayoría de los revisados en los cursos, aunque aparecen unos que no se estudiaron en clase. Empero, la parte débil de este aspecto es la interpretación de los resultados. En ese caso se observa una fuerte confusión de conceptos estadísticos; también aparece el error al interpretar el resultado del valor de la prueba y su probabilidad asociada y, en cuanto al proceso, confundir el tipo de muestras (relacionadas o independientes) que se comparan.

Otra expectativa se refería a las variables que hemos descrito sobre el planteamiento del problema, la planificación y recogida de datos. Esperábamos que nuestros resultados fuesen inferiores que los reportados para variables similares por Torralbo y sus colaboradores, pues en nuestro caso se trata de estudiantes de grado (en

vez de doctorado) y la asesoría que reciben es menor que en aquél caso.

Efectivamente, el aspecto de planteamiento del problema es el de menor nivel de concreción entre los estudiantes de psicología, pues los aspectos en los que se observa mayor cantidad de dificultades son las variables, los objetivos y las hipótesis. Debe mencionarse que en términos generales, los trabajos aquí analizados están por debajo en la mayoría de los aspectos de las tesis doctorales revisadas por Torralbo y cols. (2004) y Vallejo (2005). Al respecto debe mencionarse que las diferencias son, sin duda, producto del distinto nivel de concreción del trabajo, de la profundidad del estudio y, por supuesto, de la diferencia en la experiencia realizando investigación de los egresados de una licenciatura y de un doctorado. Podría ser interesante utilizar el mismo instrumento utilizado en esos trabajos y analizar las tesis de la licenciatura, de ese modo se tendrá un mejor parámetro de comparación.

Por otro lado, también es debido al tipo de estudio que se realiza. Al ser un estudio cuasiexperimental en el que se hace una intervención psicopedagógica a un alumno o grupo de alumnos o profesores o padres, no parece necesario definir de manera explícita y precisa ninguno de esos aspectos del ciclo de investigación, en particular las variables. Por otro lado, es importante mencionar que una buena cantidad de dificultades en el planteamiento de hipótesis fue cuando no era necesario elaborarlás pues el estudio no lo requería. Los egresados lo hacen porque perciben que es necesario, sin embargo, logran un planteamiento confuso y mal elaborado. Con base en estos resultados se puede suponer un mayor énfasis en la revisión de los distintos tipos de estudio, sus características y su relación con el ciclo de la investigación.

Finalmente, esperábamos mejor aplicación de los modos fundamentales de razonamiento estadístico que lo encontrado por Pimenta (2005) pues nuestros alumnos reciben mayor número de cursos estadísticos y metodológicos. Respecto a esta expectativa, observamos en general que el razonamiento estadístico de los egresados en nuestro estudio es mayor de lo supuesto inicialmente, pues en comparación con el estudio de Pimenta (2005), nuestros egresados describen más adecuadamente el tipo de diseño que utilizaron para hacer su trabajo.

Asimismo, hacen mayor uso de instrumentos específicos para obtener sus datos. Esta diferencia es mayor si se compara la cantidad de trabajos en los que se explicita la construcción de instrumentos. En consecuencia, si bien se reconoce la variabilidad de los resultados en porcentajes semejantes, el uso de validez en los trabajos del presente estudio es mayor. Una explicación es que nuestros graduados tuvieron, durante su

formación, una cantidad mayor de cursos de metodología que los estudiantes del trabajo de Pimenta (2005); aunque debe mencionarse que los cursos de estadística fueron similares en número, duración y temas abordados. Otro elemento que puede producir la diferencia es el tiempo de trabajo para elaborar la tesis. En el caso de los egresados de psicología, el promedio es un año con tres meses para concluir el trabajo una vez registrado como proyecto de tesis.

Líneas de investigación futura

Los resultados obtenidos sugieren el interés de continuar esta investigación, ampliando la muestra de tesis revisadas y el tipo de variables consideradas, así como profundizando en el análisis. Algunos puntos en los que se podría continuar son los siguientes:

- Realizar un análisis más profundo del uso procedimientos estadísticos, considerando si cuando se usa un método se respetan los supuestos para su uso, analizando la interpretación de los estudios de inferencia (por ejemplo del valor p , intervalos de confianza, etc.), y los errores en esta interpretación.
- Analizar otras variables utilizadas por Arteaga (2008), Pimenta (2005) y Torralbo y colaboradores y que no se han considerado en el presente trabajo. Por ejemplo, estudiar los niveles de complejidad semiótica en la elaboración e interpretación de gráficas y tablas; analizar la evolución temporal de las variables; comparar resultados según la temática de estudio.
- En función de las principales carencias de formación que se encuentren en un estudio más completo, se podría proponer una mejora de los cursos de estadística en la licenciatura. Una posibilidad sería incorporar actividades didácticas basadas en la lectura crítica de los informes estadísticos en la literatura disciplinar, bien profesional o de investigación. Dichas actividades podrían ser experimentadas y evaluadas en el curso Seminario de Tesis I y II y en Seminario taller de concentración I y II.

REFERENCIAS

- Abelson, R. P. (1997). On the surprising longevity of flogged horses: Why there is a case for the significance test? *Psychological Science*, 8(1), 12-14.
- Alatorre S., De Bengoechea N., López L., Mendiola E., y Villarreal A. (2003a). *Antología 1 de Estadística*. México: Universidad Pedagógica Nacional.
- Alatorre S., De Bengoechea N., López L., Mendiola E., y Villarreal A. (2003b). *Antología 2 de Estadística*. México: Universidad Pedagógica Nacional.
- Alatorre S., De Bengoechea N., López L., Mendiola E., y Villarreal A. (2003c). *Antología 3 de Estadística*. México: Universidad Pedagógica Nacional.
- Alatorre S., De Bengoechea N., López L., Mendiola E., y Villarreal A. (2003d). *Antología 4 de Estadística*. México: Universidad Pedagógica Nacional.
- Alvarado, H. (2007). Significados del teorema central del límite en la enseñanza de la estadística en ingeniería. Tesis Doctoral. Universidad de Granada.
- Álvarez-Gayou, J. J. L. (2003). *Cómo hacer investigación cualitativa. Fundamentos y metodología*. México: Paidós Educador.
- Arteaga, P. (2008). *Análisis de gráficos estadísticos elaborados en un proyecto de análisis de datos*. Memoria de Master. Universidad de Granada.
- Arteaga, P., Batanero, C. Cañadas, G. y Contreras, J. M. (2009). La componente social y cultural de las tablas y los gráficos estadísticos. En C. Cañadas, J. M. Contreras (Eds). *Actas de las XV Jornadas de Investigación en el Aula de Matemáticas: Dimensión Histórica, Social y Cultural de las Matemáticas*. Granada: Sociedad Thales de Educación Matemática. CD-ROM.
- Arteaga, P., Batanero, C., Díaz, C. y Contreras, J. M. (2009). El lenguaje de los gráficos estadísticos. *UNION*, 18, 93-104.
- Asociación Nacional de Universidades e Institutos de Educación Superior, ANUIES. (2008). *Anuarios estadísticos*. Online: www.anuies.mx/servicios/e_educacion/index2.php
- Batanero, C. (2000). Controversies around the role of statistical tests in experimental research, *Mathematical Thinking and Learning*, 2(1-2), 75-98.
- Batanero, C. (2000) Dificultades de los estudiantes en los conceptos estadísticos elementales: El caso de las medidas de posición central. In C. Loureiro, F. Oliveira, and L. Brunheira (Eds.), *Ensino e Aprendizagem da Estatística* (pp. 31-48). Lisboa: S.P.E.

- Batanero, C. (2000). Significado y comprensión de las medidas de posición central. *Uno: Revista de Didáctica de las Matemáticas*, 7(25), 41-58.
- Batanero, C. (2002). *El proceso de investigación educativa. Etapas principales*. Documento de Trabajo para el curso de doctorado Diseño de investigaciones educativas en ciencias e matemáticas, Universidad de Santiago de Compostela
- Batanero, C. (2004). ¿Hacia dónde va la educación estadística? *Biaix*, 5, 2-3.
- Batanero, C., Arteaga, P. y Ruiz, B. (2010). Análisis de la complejidad semiótica de los gráficos producidos por futuros profesores de educación primaria en una tarea de comparación de dos variables estadísticas. *Enseñanza de las Ciencias*, 28(1), 141-154
- Batanero, C. y Díaz, C. (2006). Análisis del proceso de construcción de un cuestionario sobre probabilidad condicional. Reflexiones desde el marco de la TFS En A Contreras (Ed.), *Investigación en Didáctica de las Matemáticas* (pp. 13-36). Universidad de Granada.
- Batanero, C. y Díaz, M. C. (2007). ¿Cómo puede el método bayesiano contribuir a la investigación en psicología y educación? *Paradigma* 27(2), 35-54.
- Batanero, C. y Díaz, C. (2009). *El proceso de inferencia causal. Valoración de los resultados de la investigación*. Documento de Trabajo para el curso de doctorado Diseño de investigaciones educativas en ciencias e matemáticas, Universidad de Santiago de Compostela
- Batanero, C., Díaz, C. y Wilhelmi, M. (2008). Errores frecuentes en el análisis de datos en psicología y educación. *Publicaciones*, 35, 109-123.
- Batanero, C., Godino, J. D. y Vallecillos, A. (1992). El análisis de datos como útil y objeto de la didáctica de la matemática. *Educación Matemática*, 4(1), 46-53.
- Batanero, C. y Sánchez, E. (2005). What is the nature of high school student's conceptions and misconceptions about probability? En G. Jones (Ed.), *Exploring probability in school: challenges for teaching and learning* (pp. 241-266). New York: Springer.
- Batanero, C. & Serrano, L. (1994). La aleatoriedad, sus significados e implicaciones educativas. *UNO*, 15-28.
- Barnett, V. (1988). Statistical consultancy. A basis lar teaching and research. En R. Davison y J. Swift (Eds.), *Proceedings of the Second International Conference en Teaching Statistics* (pp. 303-307). Victoria: University of Victoria.
- Behar, R. (2001). *Aportaciones para la mejora del proceso de enseñanza-aprendizaje*

- de la estadística*. Tesis doctoral. Universidad Politécnica de Cataluña.
- Bertin, J. (1967). *Semiologie graphique*. Paris: Gauthier-Villars.
- Bisquerra, R. (1989). *Metodos de investigacion educativa*. Grupo Editorial CEAC SA. Barcelona. Espanha.
- Bollás, P., Ruiz, A., Castillo, A. y Pedraza, H. (2005). *Instructivo de titulación*. Universidad Pedagógica Nacional, Licenciatura en Psicología Educativa: Documento Interno.
- Carpintero, H. (2005). El psicólogo que se busca. *Análisis y Modificación de Conducta*, 31(136-137), 189-207.
- Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, 49(12), 997-1003.
- Coll, C. y Valls, E. (1992). El aprendizaje y la enseñanza de procedimientos. En C. Coll, J. I. Pozo, B. Sarabia y E. Valls. *Los contenidos en la Reforma* (pp. 81-132). Madrid: Santillana.
- Cumming, G., Williams, J. y Fidler, F. (2004). Replication, and researchers' understanding of confidence intervals and standard error bars. *Understanding Statistics*, 3, 299-311.
- De Vega, M. (2005). Psicología científica y práctica profesional del psicólogo. *Análisis y Modificación de Conducta*, 31(136-137), 155-174.
- Díaz, C. (2007). *Viabilidad de la enseñanza de la inferencia bayesiana en el análisis de datos en psicología*. Tesis Doctoral. Universidad de Granada.
- Díaz, C. y Batanero, C. (2006). ¿Cómo puede el método bayesiano contribuir a la investigación en psicología y educación? *Paradigma*, 27(2), 35-53
- Díaz, C., Batanero, C. y Wilhelmi, M. (2008). Errores frecuentes en el análisis de datos en educación y psicología. *Publicaciones*. 35, 109-133.
- Díaz, C. y de la Fuente, I. (2004). Controversias en el uso de la inferencia en la investigación experimental. *Metodología de las Ciencias del Comportamiento*, Volumen especial 2004, 161-167
- Estes, W. K. (1997). Significance testing in psychological research: Some persisting issues. *Psychological Science*, 8(1), 18-20.
- Estrada A. (2000). *Actitudes hacia la estadística e instrumentos de evaluación*. Lleida: Departamento de Matemática.
- Estrada, A. y Batanero, C. (2008). Explaining teachers' attitudes towards statistics. En C. Batanero, G. Burrill, C. Reading y A. Rossman (Eds.). *Joint ICMI/ IASE Study: Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher*

- Education. Proceedings of the ICMI Study 18 Conference and IASE 2008 Round Table Conference.* Monterrey: International Commission on Mathematical Instruction e International Association for Statistical Education. CD- ROM
- Fernández, J. A., Batanero, C., Contreras, J. M. y Díaz, C. (2009). A simulação em Probabilidades e Estatística: potencialidades e limitações. *Quadrante*, XVIII (1-2), 161-183
- Fidler, F. y Cumming (2005): «Interval estimates for statistical communication: problems and possible solutions» en Trabajo presentado en la *IASE Satellite Conference on Communication of Statistics*. Sydney: International Association for Statistical Education.
- Fortín M. F. (1999). *El proceso de investigación: de la concepción a la realización*. Madrid: McGraw-Hill,
- Fox, D. J. (1987). *El proceso de investigación en la educación*. Pamplona: Eunsa. Segunda ed.
- Gephart, R. P. (1988). *Ethnostatistics :Qualitative foundations br quantitative research*. Sage University Paper series in Qualitative Research Methods (Vol.2). Beverly Hills, CA: Sage.
- Gingerenzer, G. (1993). The superego, the ego and the id in statistical reasoning. En G. Keren y C. Lewis (Eds.), *A handbook for data analysis in the behavioural sciences: Methodological issues* (pp. 311–339). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Goetz, L. P. y Lecompte, M. D. (1988). *Etnografía y diseño cualitativo en la educación*. Madrid: MacGraw Hill.
- Harris, R. J. (1997). Significance tests have their place. *Psychological Science*, 8(1), 8-11.
- Harris, R. R. y Kanji, G. K. (1988). The challenges of statistical training. En: R. Davidson y J. Swift (Eds.). *Proceeding of the Second International Conference en Teaching Statistics* (pp. 498-501). Victoria: University of Victoria.
- Hernández, G. J., Sánchez, H. S., Ortega, S. M. C. y Millán, P. (1990). *Propuesta general de reestructuración curricular de la Licenciatura en Psicología Educativa de la UPN*. México: Universidad Pedagógica Nacional, Academia de Psicología Educativa: Documento Interno.
- Hernández, R.; Fernández, C. y Baptista, P. (2007). *Fundamentos de metodología de la investigación educativa*. Madrid: McGraw Hill.

- Hil, H.C. y Shih, J. (2009). Examining the quality of statistical mathematics education research. *Journal for Research in Mathematics Education*, 40(3), 241-250.
- Hunter, J. E. (1997). Needed: A ban on the significance test. *Psychological Science*, 8 (1), 3 – 7.
- Kerlinger, F. N. y Lee, H. B. (2002). *Investigación del comportamiento. Métodos de investigación en ciencias sociales*. México: McGraw-Hill.
- Krippendorff, K. (1991). *Metodología del análisis de contenido*. Buenos Aires/Barcelona: Paidós-Comunicación.
- León, O. G. y Montero, I. (1997). *Diseño de investigaciones*. Madrid: McGrawHill
- López-Portillo, J. L. (1978). Decreto de creación de la universidad pedagógica nacional. *diario Oficial*, agosto 25. México: Gobierno de la República.
- Lozano, A. y Rodríguez, M. (2006). *Perfil de ingreso. Serie histórica 1995-2003: Estudios sobre la universidad*. México: UPN.
- Martínez Arias, R. (1995). *Psicometría: teoría de los tests psicológicos y educativos*. Madrid: Síntesis
- Miles, M. B. y Huberman, A. M. (1984). *Qualitative data analysis*. Beverly Hills, CA: Sage.
- Morrison, D. E., y Henkel, R. E. (Eds.). (1970). *The significance tests controversy. A reader*. Chicago: Aldine.
- Olivo, E. (2008). *Significados de los intervalos de confianza para los estudiantes de ingeniería en México*. Tesis doctoral. Universidad de Granada.
- Olivo, E., Batanero, C. y Díaz, C. (2008). Dificultades de comprensión del intervalo de confianza en estudiantes universitarios. *Educación Matemática*, 20(3), 55-82.
- Tomcho, T. J. y Foels, R. (2010). The power of teaching activities: statistical and methodological recommendations. *Teaching of Psychology*, 36(2), 96-101.
- Torralbo, M., Fernández Cano, A., Rico, L., Maz, A. y Gutiérrez, M. P. (2003). Tesis doctorales españolas en educación matemática. *Enseñanza de las Ciencias*, 21(2), 295-305
- Sánchez, C., y Estepa, A. (2000). Correlación y regresión en los primeros cursos universitarios. En M. Beltrán (Ed.), *Actas de las Jornadas Europeas de Estadística: la enseñanza y la difusión de la estadística*. Palma de Mallorca: Instituto Balear de Estadística. Online: www.caib.es/ibae/esdeveniment/jornadas_10_0

- Santisteban, R. C. (1990). La estadística en psicología. *Estadística española*, 31(122), 461-501.
- Schuyten, G. (1990). Statistical thinking in psychology and education. Presentado en el ICOTS III. University of Otago.
- Seoane, J. (2005). Psicólogos y psicólogos. *Análisis y Modificación de Conducta*, 31(136-137), 139-153.
- Universidad Pedagógica Nacional. (2002). *Reglamento general para la titulación profesional de licenciatura de la Universidad Pedagógica Nacional*. Universidad Pedagógica Nacional, Secretaría Académica: Documento Interno.
- Valera, A. Sánchez Meca, J. y Marín. F. (2000). Contraste de hipótesis e investigación psicológica española: análisis y propuestas. *Psicothema* 12 (2), 549-552
- Vallecillos, A (1994). *Estudio teórico experimental de errores y concepciones sobre el contraste de hipótesis en estudiantes universitarios*. Tesis doctoral. Universidad de Granada.
- Vallecillos, A. y Batanero, C. (1997). Conceptos activados en el contraste de hipótesis estadísticas y su comprensión por estudiantes universitarios *Recherches en didactique des mathématiques* 17 (1), 29-48
- Visauta, R. (1989). *Técnicas de investigación social I. Recogida de datos*. Barcelona: P.P.U.
- Weber, R. P. (1985). *Basic content analysis*. Londres: Sage.
- White. A. L. (1980). Avoiding errors in educational research. En: R S. Shumway (Ed.). *Research in Mathematics Education*. (pp.47-65). Reston, VA: N.C.T.M.,
- Wild, C., y Pfannkuch, M. (1999). Statistical thinking in empirical enquiry (con discusión). *International Statistical Review*, 67(3), 223-265.
- Wilkinson, L. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, 54, 594-604.