

Estadística aplicada a la educación. Investigación pedagógica

Autor: Pearson Educación

Presentación del curso

La **Estadística aplicada a la educación** tiene un área específica de estudio sobre la cual trata este curso: la **Estadística** en el proceso de **investigación pedagógica** empírica. Presentamos un estudio de aproximación a la **Estadística** como ciencia y al rol que desempeña en la **investigación** dentro del área de la **educación**.

He aquí un resumen de temas que podrás encontrar en este curso: **concepto y funciones de la Estadística; tipos de Estadística (estadística descriptiva e inferencial)**; el rol de la estadística en el proceso de investigación: formulación de hipótesis, manejo de variables, objetivos, decisión estadística, etc.; la relación de la **Estadística** con las Ciencias Sociales, y además, un análisis de las **posibilidades y límites de la Estadística** en este campo.

Aprende con este curso de **Pearson**, fragmento del libro: "**Estadística aplicada a la educación**", de los autores Ramón Pérez, José Luis García, Juan Antonio Gil, Arturo Galán. Puedes descubrir y adquirir libros de **Pearson** en: www.jetlibros.com.

1. La estadística y la investigación pedagógica

La estadística en el proceso de investigación pedagógica empírica

1.1. Introducción

En este curso de Estadística, pretendemos realizar una primera aproximación a lo que es la Estadística como ciencia y, sobre todo, al importante papel que desempeña en el campo de la investigación en educación, pues nos sirve de apoyo y ayuda para abordar el estudio de diferentes problemas que pueden contribuir a la mejora de la educación y de la sociedad.

Son muchas las razones que justificarían la presencia de la Estadística en los estudios pedagógicos, desde saber interpretar algunas publicaciones científicas hasta poder llegar a diseñar y desarrollar propuestas de trabajo en el campo educativo. Siguiendo las directrices de algunos autores, que se han ocupado del tema, como Welkowitz, Van Dalen y Meyer, Ferguson, Kerlinger, Fox, etc.; estas razones se pueden agrupar en los siguientes campos:

— Comprender los trabajos que se publican en revistas científicas, libros, informes, etc.; tanto si son recogidos en papel como a través de las nuevas tecnologías de la información y comunicación. Sin tener unos conocimientos mínimos será difícil llegar a entender muchos de los trabajos publicados en el campo educativo.

— Entender los procesos implicados en la investigación educativa que hacen posible la inferencia desde la muestra a la población y conocer las garantías que nos ofrecen estas decisiones, pues desde estos trabajos se recoge una gran información que ayudará en la elaboración de conclusiones y permitirá realizar predicciones de cara al futuro.

— Facilitar el propio desarrollo de la investigación socio-educativa. Ello exige que los futuros graduados han de llevar a cabo trabajos empíricos en su campo de trabajo profesional, algo que será difícil de abordar si no se tiene un conocimiento básico de la Estadística. No se trata por tanto de formar profesionales o expertos en este campo de estudio, sino de capacitarles para entender y aplicar esos conocimientos en la propia actividad profesional.

En resumen, entendemos que es preciso poder leer y comprender las publicaciones científicas, así como diseñar y desarrollar estudios empíricos en el campo pedagógico y entender las directrices que guían el método científico como medio de formación intelectual. Estamos ante tres argumentos sólidos que apoyan el estudio de una materia de estas características en la formación de los futuros graduados en el campo pedagógico y social.

Nota: Este curso es un fragmento del libro: "**Estadística aplicada a la educación**", de los autores Ramón Pérez, José Luis García, Juan Antonio Gil, Arturo Galán, publicado por **Pearson Educación** (ISBN: 9788483226360). Puedes descubrir y adquirir libros de **Pearson** en: www.jetlibros.com.

2. Estadística. Concepto y funciones (1/2)

1.2. Concepto y funciones de la Estadística

Al término Estadística se le asignan diversas acepciones por parte de las personas vinculadas al campo socioeducativo; así se puede entender como una serie de conjuntos de números, como método de trabajo propio del campo empírico, como a la forma de resolver problemas educativos, e incluso, a considerarla como una ciencia.

Entendida como conjunto de datos numéricos propia de los censos de personas que se ubican en un determinado espacio físico, podemos reseñar las estadísticas sobre los niveles de escolarización obligatoria, sobre el acceso a los estudios superiores de los inmigrantes, etc. Desde el otro extremo, al considerarla como ciencia, dirige su atención al estudio de los fenómenos aleatorios, con este enfoque se ha conseguido avanzar en el desarrollo de métodos, técnicas y modelos que nos ayudan en la resolución de problemas pedagógicos y, sobre todo, a la toma de decisiones. La Estadística actúa como una disciplina puente entre los modelos matemáticos y los fenómenos reales, que, además, proporciona una metodología para evaluar las discrepancias entre la realidad y la teoría.

Entre los autores que han definido la Estadística, quizá una de las más completas y comprensibles nos la ofrece Kerlinger (1985) cuando la define como:

«La teoría y el método de analizar datos cuantitativos obtenidos de muestras de observaciones, para estudiar y comparar fuentes de variancia de fenómenos, ayudar a tomar decisiones sobre aceptar o rechazar relaciones hipotéticas entre los fenómenos y ayudar a hacer inferencias fidedignas de observaciones empíricas» (p. 192).

Para Vélez y otros (2006):

«la **Estadística** es la ciencia que estudia, mediante métodos cuantitativos, las poblaciones que se obtienen como síntesis de la observación de unidades estadísticas» (p. 8).

Por lo tanto, cumple un importante papel instrumental de apoyo a la investigación socioeducativa: nos permite comprender las posibilidades y limitaciones de los trabajos empíricos y desarrollar un pensamiento crítico y antidogmático en el estudio de la realidad.

McMillan y Schumacher (2005) señalan que las funciones que puede desempeñar la Estadística para los pedagogos, educadores sociales o maestros, se pueden concretar en los siguientes puntos:

- Entender y criticar artículos profesionales.
- Mejorar la evaluación del aprendizaje de los estudiantes.
- Dirigir, incluso a modo de iniciación e informal, trabajos sencillos de investigación, para ir profundizando a medida que se avanza en los estudios de los grados.
- Entender las propuestas de evaluación de programas socioeducativos, del

profesorado y de las instituciones educativas y sociales.

— Ayudar en la preparación ciudadana y como consumidor de productos, tomando decisiones bien fundamentadas en datos o argumentos cuantitativos.

— Mejorar la profesión socioeducativa, pues proporciona habilidades de comunicación, debate, discusión de investigaciones que tengan implicaciones para la práctica social y educativa.

3. Estadística. Concepto y funciones (2/2)

Así pues, la Estadística puede aportar una primera aproximación al campo de estudio a partir de los datos que facilita al investigador, con posterioridad se pueden aplicar los principios derivados de la Estadística, entendida como ciencia, para la resolución de problemas o para la toma de decisiones. Por ejemplo, puede realizarse una recopilación de datos sociodemográficos de una determinada zona residencial, para prever la construcción de escuelas infantiles a corto plazo. Esta información puede ser útil para las autoridades administrativas.

También nos puede facilitar datos sobre la puesta en marcha de un programa de intervención educativa frente a otro que se venía utilizando con anterioridad en los mismos sujetos, es decir, nos permitirá detectar si las diferencias en los resultados académicos son motivadas por el programa o se pueden explicar por efecto del azar.

Downie y Heath (1983) señalan que los profesionales de la pedagogía, sociología, psicología, en sus trabajos de estudio e investigación, suelen disponer de muchos y variados datos que necesitan ser analizados e interpretados, algo que no sería posible sin un conocimiento de la Estadística, así podemos señalar que nos ofrece las siguientes **posibilidades**:

- Permite calcular las medias aritméticas de los grupos, para comparar niveles de dominio de las variables.
- Establece el grado de dispersión o variabilidad de las observaciones y datos.
- Representa gráficamente, mediante tablas y figuras, los grupos o los sujetos individuales.
- Ayuda en la transformación de datos para compararlos (centiles, puntuaciones típicas).
- Establece relaciones entre variables y calcula el grado de relación entre las mismas.
- Facilita el cálculo de la fiabilidad y la validez de los instrumentos de recogida de datos y de medida.
- Recurre a diversas mediciones o combinación de variables para predecir el comportamiento futuro.
- Permite extrapolar los resultados de muestras representativas a grupos más amplios (poblaciones).
- Compara las actuaciones de dos o más grupos y establece la existencia o no de diferencias significativas, lo que permite comprobar qué procedimiento de trabajo ofrece mejores resultados.

De forma sintética, las principales funciones de la Estadística se pueden agrupar en torno a tres grandes apartados:

a) Facilitar el manejo de datos amplios y dispersos: se trata de una función eminentemente simplificadora y descriptiva, pues se pretenden reducir a índices o estadísticos (media, mediana, desviación típica, correlación) las características que identifican a un conjunto de datos. Así se procede a la ordenación, la

categorización, posición, variabilidad, simetría y asimetría, representaciones gráficas y estudios de relaciones entre variables.

b) Inferir desde la muestra a la población: este proceso consiste en el análisis que sigue el investigador para extrapolar los resultados obtenidos en las muestras a las poblaciones de las que se extrajeron. Estamos hablando de la generalización de los resultados de muestras, cuando éstas cumplen los requisitos de suficiencia y representatividad.

Es decir, desde los estadísticos (proceden de muestras) se pueden estimar los parámetros (medidas de población).

c) Ayudar en la toma de decisiones: tanto desde los meros análisis descriptivos de grupos como el análisis de los valores críticos que ponen de relieve la existencia de diferencias significativas entre los grupos de tratamiento. En ambos supuestos se trata de incorporar cambios en las actuaciones educativas que redunden en mejoras de las formas de actuación en el campo socioeducativo.

4. Tipos de estadística. Estadística descriptiva

1.3. Tipos de Estadística

Teniendo en cuenta las funciones y cometidos de la Estadística parece desprenderse que nos encontramos ante dos tipos diferenciados de la misma. Ahora bien, debemos ser conscientes que su presencia no nos asegura la calidad de la investigación, pero sí se acepta como lenguaje universal, aunque su significado se deriva del propio diseño de la investigación.

En el ámbito de la Estadística entendida como método de aplicación de los principios científicos para la resolución de problemas socioeducativos y la toma de decisiones, podemos identificar dos grandes tipos, según las tareas a las que debe enfrentarse, la descriptiva y la inferencial.

1.3.1. Estadística descriptiva

A esta modalidad de Estadística se le asocian los procesos de análisis que se llevan a cabo con los datos empíricos recogidos en las muestras, en líneas generales este proceso concluye con la obtención de unos valores numéricos que reciben la denominación de *estadísticos*, cada uno de ellos pone de relieve una característica representativa del grupo de sujetos que se analiza. En realidad, transforma un conjunto de números u observaciones en índices que sirven para describir o caracterizar esos datos dentro de los grupos de sujetos.

La podemos considerar como una parte de la Estadística que se ocupa del estudio de los métodos y técnicas necesarios para la descripción gráfica y numérica de los conjuntos de datos numerosos. Ello nos ofrece una visión global del grupo de sujetos que es objeto de estudio.

Esta propuesta tiene limitaciones en la interpretación de los estadísticos, pues en muchas ocasiones nos debemos centrar en una comparación entre el valor de la muestra y otros que procedan de muestras similares, por lo que no aporta suficientes argumentos científicos al investigador en la toma de decisiones sobre los grupos.

Por ejemplo, si en un grupo (A) se obtiene una media de 5,25 y en otro grupo (B) de similares características se ha obtenido una media de 5,75; en el campo de la Estadística descriptiva nos limitaríamos a afirmar que el segundo grupo (B) tiene un nivel más elevado en la variable que se está midiendo; en el caso de que el primer grupo (A) alcanzara una desviación típica ($s=2,25$) y el segundo (B) ($s=1,75$) podríamos afirmar que las puntuaciones del primer grupo son más dispersas que en el segundo, también decimos que el segundo grupo es más homogéneo y el primero más heterogéneo. En el caso de las correlaciones entre variables o grupos si tenemos un valor de 0,78 a simple vista nos puede parecer que la relación es buena, aunque esta apreciación simple necesitaría de otras referencias para proceder a una valoración más objetiva.

Por lo tanto, podemos afirmar que la *Estadística descriptiva es aquella parte de la Estadística que utiliza estadísticos procedentes de muestras o de poblaciones con una finalidad eminentemente descriptiva o informativa de las mismas*. Así pues, realiza una tarea de síntesis y descripción de las características de uno o más conjuntos de datos, lo que refleja su propia naturaleza. El uso de la misma es la forma más adecuada para resumir los datos y resulta indispensable para interpretar los resultados de los trabajos empíricos.

En esta línea de análisis sobre la Estadística descriptiva, Fox (1981) señala cinco funciones principales:

- Conocer los estadísticos o características esenciales de un conjunto de datos.
- Interpretar lo que nos dicen los estadísticos muestrales sobre los parámetros o medidas de población de la que se extrajo la muestra.
- Conocer la existencia, el sentido y la magnitud de la relación entre dos variables.
- Explorar las relaciones, manifiestas o no, en la estructura que define la relación de más de dos variables.
- Utilizar los estadísticos procedentes de muestras para conocer las tendencias en el comportamiento de nuevos grupos de sujetos.

Según la naturaleza o características de las muestras y los datos se puede diferenciar entre: la Estadística descriptiva univariada, que cubre la más simple de las funciones, pues intenta descubrir y analizar una distribución de datos que provienen de la medición de una variable en una muestra y la Estadística descriptiva bivariada, que recoge y analiza datos de dos variables, es el campo propio de las correlaciones. Cuando intervienen más de dos variables se habla de la multivariada.

5. Tipos de estadística. Estadística inferencial

1.3.2. Estadística inferencial

Pretende avanzar más en el estudio de la realidad socioeducativa, pues le corresponde decidir sobre aquellas cuestiones no resueltas por la descriptiva; así, trata de extrapolar los resultados que se han obtenido en muestras a las poblaciones respectivas de las que proceden. En este sentido podemos afirmar que se *se ocupa de los métodos estadísticos que nos sirven para realizar inferencias objetivas sobre los datos disponibles y trasladarlos a grupos más amplios*. Es decir, se emplea para realizar predicciones sobre la similitud de una muestra con la población de la que fue extraída. Por lo tanto, se ocupa de los métodos que son precisos para establecer conclusiones sobre una población a partir de una muestra de la misma.

Tomando como referencia los datos del ejemplo anterior, entre las medias 5,25 y 5,75 existe una diferencia empírica de 0,50 puntos. Ahora bien, la Estadística inferencial nos permitirá afirmar si esa diferencia es significativa, es decir, hay una diferencia real o bien se puede explicar por efecto del azar (aleatorización), para probar esta afirmación será preciso recurrir a las correspondientes pruebas estadísticas de contraste. En el caso de un coeficiente de correlación de 0,75, una vez sepamos el procedimiento de cálculo (Pearson, Spearman, Biserial-puntual, Tetracórico, etc.) tendremos que recurrir a una serie de pruebas estadísticas de contraste que nos indicarán si ese valor pone de relieve una diferencia real (significativa) o meramente aleatoria, que se puede explicar por efecto del azar; en el primer supuesto podemos afirmar que si cambiamos los valores de una variable, podemos aventurar que la otra se verá también modificada.

La Estadística inferencial tiene por finalidad la obtención de una serie de conclusiones sobre algún aspecto o variable presente en una población a partir de las observaciones de comportamientos en una o varias muestras. Es decir, los valores de la población (parámetros) nos permiten conocer el fenómeno o hecho socioeducativo en muestras amplias, además de resolver y fijar las hipótesis, plantear leyes y tomar decisiones con rigor científico, valorando en estos casos los márgenes de error con los que realizamos nuestras afirmaciones (nivel de confianza).

Hablar de nivel de confianza nos permite abordar un tema crucial, pues hemos de tener en cuenta que en el marco de la investigación socioeducativa, las decisiones que se toman no se pueden realizar en términos de certeza, sino de probabilidad, por lo que en esas estimaciones debemos fijar los márgenes de error en nuestras afirmaciones que suelen ser de 0,05 y 0,01 (en términos de porcentajes, del 5% y del 1% respectivamente); que llevan asociados los correspondientes niveles de confianza, complementarios a los anteriores, generalmente expresados en porcentajes: 95% y 99% respectivamente.

La inferencia nos ayuda en la toma de decisiones sobre la aceptación o el rechazo de las relaciones previstas en la hipótesis (posible solución del problema), aunque estas decisiones suelen matizarse con la fijación del margen de error. En cualquier caso, debemos ser prudentes en el momento de llevar a la práctica los cambios o

modificaciones que propugnaban las hipótesis, pues en el campo educativo se presentan otros factores o condiciones no suficientemente planificadas que pueden distorsionar los resultados finales.

Entendemos que son dos los grandes campos que forman la Estadística inferencial: la *estimación de parámetros* y el *contraste de hipótesis*. La estimación se puede llevar a cabo mediante la elección de un solo valor de la muestra que se transforma en parámetro (estimación puntual) o a través de unos límites entre los cuales se espera se encuentre el verdadero valor del parámetro (estimación por intervalos), en este caso debemos ser conscientes de que esos límites vienen influenciados por los errores aleatorios y los sistemáticos.

El contraste de hipótesis consiste en probar mediante datos empíricos las hipótesis que se plantean en el proceso de investigación, de tal forma que son los datos los que deben ofrecer una respuesta a los planteamientos iniciales del investigador y no al revés, pues se violaría el sentido de la ciencia. Como decía Kerlinger, demos oportunidad a los datos de mostrar que los cambios en los comportamientos son achacables a las tesis iniciales.

6. La estadística en el proceso de investigación

1.4. El papel de la Estadística en el proceso de investigación

Después de analizar los aspectos conceptuales y funcionales de la Estadística, entendida como ciencia instrumental de apoyo a la investigación, es el momento de abordar el papel que desempeña en el proceso de investigación empírica y, de forma más detallada, en cada una de las principales fases de ese proceso. Desde este punto de vista nos interesa destacar aquellas funciones que se relacionan de forma más directa con este proceso, ello no quiere decir que descartemos los demás usos. En este sentido, la Estadística descriptiva aporta información que puede servir de base para trabajos posteriores y análisis más complejos que entran en el cometido de la Estadística inferencial.

Así, los datos recogidos deben responder a unas hipótesis u objetivos previamente planteados, pues las tablas y gráficos no significan nada en sí mismas, si no existen garantías de su representatividad, del control de variables que es preciso tener en cuenta y la elección de los instrumentos más adecuados. Por ello, en las páginas que siguen queremos poner de relieve el papel de la Estadística en cada una de las fases del proceso (que se verá con más detalle en el siguiente capítulo), pues así quedarán delimitadas las posibilidades y limitaciones de cada una de las propuestas de estudio en el campo socioeducativo. Hemos de señalar que en alguna de las fases va a estar presente de forma explícita, mientras en otras ofrecerá pautas para entender y desarrollar esos apartados.

1.4.1. En el problema de investigación

El punto de arranque de toda propuesta de investigación se ubica en la identificación y selección del problema, en estos momentos debe estar presente la Estadística, no como cálculo de estadísticos, sino como garantía para poder establecer relaciones entre las características que se analizan; ello se lleva a cabo mediante diversos modelos estadísticos, desde ellos buscamos si existe alguno que permita la resolución del problema o bien, si se pudieran aplicar varios, seleccionar aquél que es más adecuado en función de unos criterios fijados.

Así, cuando procedemos a analizar el problema y sus características identificativas se pondrán de manifiesto las variables que intervienen en esos procesos, cómo actúan en el contexto de la investigación, qué datos de la variable dependiente necesitan ser recogidos, qué calidad tiene la información, qué pruebas estadísticas es preciso aplicar, en suma se trata de identificar si existe un modelo estadístico capaz de ofrecer una respuesta adecuada a ese problema.

Si el investigador no contempla estas sugerencias se puede encontrar más adelante con un camino sin salida, donde no puede avanzar más y, en ocasiones, se puede ver obligado a abandonar el trabajo. A veces se sigue adelante, pero a riesgo de alcanzar unos resultados que no pueden ser considerados válidos, por su incoherencia entre el problema y el modelo estadístico generado para su resolución.

En síntesis podemos afirmar que es la Estadística la que nos permitirá afirmar que el problema cumple una condición inexcusable en su formulación *que sea resoluble*, es decir, que con los datos que se puedan recoger en un futuro sea posible alcanzar las respuestas esperadas.

7. La estadística en el proceso de investigación. Hipótesis, objetivos, variables

1.4.2. En la formulación de hipótesis y objetivos

Cuando el investigador formula su hipótesis o enumera sus objetivos se le exige que las hipótesis sean *contrastables* y los objetivos *comprobables*. Ello será posible si disponemos de una serie de instrumentos de recogida de datos que nos ofrezcan una información nítida y detallada sobre ese problema que nos ocupa y cuya solución se adelanta en la hipótesis.

Referido al tipo de hipótesis nos permite conocer qué estadísticos deben ser calculados, qué análisis son necesarios en ese problema. También nos ayuda a determinar si la hipótesis que se plantea ha de ser unilateral o bilateral, es decir, el sentido de la hipótesis de investigación. Además nos ofrece argumentos para seleccionar la modalidad de análisis estadístico que es el más adecuado para la validación o comprobación empírica de esa hipótesis.

Si queremos averiguar la influencia de un método para el aprendizaje de la lengua inglesa, es lógico suponer que al trabajar con dos o más grupos de sujetos, el dato que realmente vamos a necesitar será la media de cada grupo, que nos indica el nivel alcanzado en esta asignatura. En cambio, si nos planteamos las características de los grupos y la eficacia de una prueba objetiva, puede ser de interés conocer la desviación típica (s) o la varianza (s^2) para determinar la mayor o menor dispersión de los datos.

En el supuesto de las hipótesis planteadas sobre la existencia o no de *diferencias* entre los grupos, es preciso recurrir a las pruebas estadísticas de contraste que nos permitan decidir sobre la significación estadística o no de esas diferencias; pueden referirse a las medias, las varianzas, los porcentajes, las relaciones, etc. Esta elección viene condicionada por los conocimientos que sobre el tema de estudio posee el investigador; así en la gran mayoría de los casos nos decantamos por plantear hipótesis bilaterales (existirán diferencias entre los grupos); mientras que cuando existen otros trabajos debidamente contrastados, o bien que responden a teorías previas se pueden formular hipótesis unilaterales (se decantan por una de las opciones). Una elección u otra de la hipótesis van a tener repercusiones en la búsqueda de los valores teóricos o críticos de distribución del correspondiente estadístico. Con ello buscamos la existencia de diferencias significativas y no meramente aleatorias.

Si la hipótesis se decanta por establecer el valor predictivo de una prueba sobre otra, nos estamos refiriendo a los valores que alcanzará la correlación, medidos en tiempos diferentes (validez predictiva). Desde esta formulación será preciso determinar la función predictiva del coeficiente que se traslada a las correspondientes fórmulas derivadas de la regresión. Se trata de ofrecer el peso relativo de las variables independientes sobre la dependiente mediante las ecuaciones de regresión lineal.

1.4.3. En el control de variables extrañas

En la investigación de naturaleza cuantitativa, uno de los elementos clave consiste en alcanzar un buen control como medio que nos permita asegurar la validez de los resultados alcanzados. Se trata de llegar hasta donde sea posible, respetando al objeto de estudio (la persona), para evitar la contaminación de los resultados. Debemos partir de la base que este control no será equiparable al que se puede lograr en el ámbito de las ciencias físico-naturales y, en general, en las ciencias experimentales, pues las condiciones a que se someten pueden ser muy estrictas y rígidas; en cambio, en el marco de la investigación socioeducativa el sujeto de estudio es la persona, lo que lleva consigo connotaciones morales o éticas, ello limita el grado de control.

Entre las diversas formas de control que tenemos en la investigación en educación hay algunas que apenas guardan relación con la Estadística (trabajar solamente con sujetos de un solo sexo, elegir un diseño determinado); pero hay otras de naturaleza estadística como pueden ser el tomar unos determinados valores en la variable independiente, formar parejas o bloques de sujetos que puntúan de forma similar en una prueba previa. En estos supuestos la Estadística nos aporta los análisis pertinentes para la selección de sujetos en función de esas características de control.

Así se puede determinar el valor del cociente intelectual, del nivel previo de conocimientos en un determinado campo del saber, del influjo del entorno social, de la edad y maduración de los sujetos. Con todos estos datos se podrán ocupar antes del tratamiento o intervención del investigador sobre la variable independiente, con lo que podemos garantizar que la existencia de diferencias posteriores se debe a la intervención sobre la variable independiente y no a otras causas ajenas al proceso investigador.

En síntesis garantizar el control en los procesos de investigación empírica supone que se aíslan o minimizan los efectos de las covariaciones y la influencia de las variables extrañas que pudieran llegar a ofrecer explicaciones alternativas a las buscadas por el investigador. El hecho de poder cuantificar estos pesos relativos nos permitirá realizar afirmaciones con la suficiente validez y rigor en el campo socioeducativo.

8. La estadística en el proceso de investigación. Definición de variables

1.4.4. En la definición de las variables

En el caso de la investigación empírico-experimental, en el que la hipótesis establece una relación de dependencia o de causalidad entre las variables, de tal forma que el investigador cuando formula sus hipótesis u objetivos está indicando la relación entre aquellas variables sobre las que interviene y modifica y aquella o aquellas que recogerán los efectos de esa intervención. Así podemos establecer que la motivación que reciben los estudiantes influirá de forma positiva sobre el rendimiento alcanzado en una determinada asignatura. En este caso, la variable independiente (Vi) será la motivación y la variable dependiente (VD) el rendimiento obtenido después de la intervención.

En la investigación socioeducativa la Estadística se manifiesta en la medición de las variables que intervienen en el proceso. Esta tarea requiere la utilización de instrumentos que sean fiables y válidos, de tal forma que nos proporcionen datos de calidad para tomar decisiones. Para determinar la fiabilidad y la validez de los instrumentos de recogida de datos y de medida es preciso recurrir a la Estadística, que mediante las oportunas fórmulas nos garantiza la medida precisa de esas variables. Además, nos ayudará también a definir de forma operativa las variables independientes, lo que quiere decir que nos indicará la forma en que se debe proceder a su medida y valoración.

Recordemos que, en ocasiones, las variables son constructos que no admiten una medida directa, por lo que es preciso definir conductas operativas y medibles que nos indiquen el valor real de las mismas en el contexto de la investigación socioeducativa. Esta operación se conoce como la definición operativa de las variables y en ella la Estadística desempeña un cometido fundamental.

Si queremos medir el cociente intelectual (CI) de un grupo de sujetos de Educación Primaria, que puede actuar como variable independiente en una investigación, o bien la motivación, dos variables que no tienen una medida directa, tendremos que recurrir a una serie de pruebas que nos manifiestan los dominios o las opiniones de los sujetos de la muestra, a partir de estos datos se podrán inferir puntuaciones de la inteligencia o de la motivación. En esta tarea es fundamental contar con el apoyo instrumental de la Estadística.

En el proceso de medida de las variables debemos ser conscientes que cuando estamos obteniendo un determinado valor para la variable en cuestión, éste puede venir influenciado por los errores asociados al acto de medir la variable y a la estimación indirecta que se realiza en no pocas ocasiones. Dado que no podemos hablar en términos de certeza, es conveniente reflejar en los trabajos empíricos que los datos obtenidos de la medida de la variable pueden estar sujetos a variaciones provocadas por múltiples factores y que, en determinados supuestos, puede tener unas repercusiones sobre la investigación y las aplicaciones de los resultados a la práctica real.

9. La estadística en el proceso de investigación. Contraste de hipótesis

1.4.5. En el contraste de hipótesis o comprobación de objetivos

El contraste de hipótesis o la comprobación de objetivos es uno de los momentos importantes del proceso de investigación, donde se pone de manifiesto de forma explícita la utilización de la Estadística. Si en las fases anteriores actuaba de forma latente o de apoyo al proceso, en esta fase nos ayuda a resolver y analizar los datos recogidos en la investigación.

Para realizar la operación de comprobación empírica de las hipótesis, debemos recurrir a las pertinentes pruebas estadísticas cuya utilización dependerá de la calidad de los datos recogidos en la investigación. Aunque en capítulos posteriores se analizará con más detalle estas cuestiones, sí conviene tener en cuenta que los modelos estadísticos generados por los datos serán sometidos a prueba mediante los estadísticos más adecuados a esas características.

Alcanzar una solución para el problema formulado inicialmente es una tarea inalcanzable en la investigación socioeducativa sin recurrir a la Estadística, entendida como ciencia instrumental de apoyo a los procesos de comprobación empírica, que nos ofrece una amplia gama de pruebas estadísticas, cuya selección queda en manos del investigador, si bien, debe basarse en el rigor y la veracidad de los datos, pues no todas las pruebas sirven para todos los cometidos, aunque el uso de la informática y los paquetes estadísticos de análisis de datos nos puedan ofrecer salidas, que no contemplan estas peculiaridades y cualidades de los datos.

Si queremos medir los rendimientos académicos de los estudiantes mediante una prueba objetiva, podemos garantizar que los datos recogidos nos ofrecen una medida bastante precisa de la variable. Ahora bien, si esta medida se basa en las apreciaciones del profesor, tomando como referencia la observación en el aula, los datos no tienen la misma precisión. En el primer supuesto se podrá recurrir a las denominadas pruebas paramétricas, en el segundo caso probablemente debemos emplear pruebas no paramétricas. En otras situaciones se recogen datos en categorías o grupos mediante la frecuencia, o sea, el número de sujetos que se adscriben a una determinada categoría o a un grupo, en este supuesto tendremos que recurrir también a pruebas no paramétricas.

En síntesis, nos encontramos ante una fase eminentemente estadística, pero ello no quiere decir que las aportaciones en otras partes no sean tenidas en cuenta o sean de poca entidad, pues los procesos lógicos y los estadísticos están íntimamente ligados y mutuamente condicionados en el campo de la investigación socioeducativa. La utilización de una prueba u otra debe quedar en manos del investigador, no del informático, éste nos puede ayudar y facilitar las herramientas de cálculo, mientras que los investigadores debemos conocer estas pruebas y, sobre todo, ser capaces de interpretar las salidas del ordenador.

1.4.6. En la decisión estadística

La fase anterior concluye con la obtención del denominado valor empírico del estadístico correspondiente, que será diferente según la prueba estadística que se haya seleccionado («t», «z», «F», «U», «T», «H», etc.), y que ha de ser interpretado, de tal forma que el investigador pueda concluir que las diferencias entre los grupos son diferencias reales y no meramente aleatorias, o sea que esas diferencias no son explicables por efecto del azar.

Esa interpretación consiste en decidir si la hipótesis de nulidad (H_0) se rechaza y, por consiguiente, se acepta la hipótesis alternativa o hipótesis del investigador (H_1); esta decisión se hace fijando unos niveles de confianza o unos márgenes de error. Recordemos que la investigación no puede realizar afirmaciones en términos de certeza. En esta última fase del proceso de investigación la Estadística, se pone de manifiesto al ofrecernos los valores teóricos o críticos de distribución de cada uno de los estadísticos y que vienen recogidos en tablas o bien se determinan mediante el empleo de una serie de ecuaciones. La cuantía de los estadísticos depende de muchos factores, como los tamaños de las muestras que se han utilizado, el número de grupos, los grados de libertad, el nivel de confianza, el margen de error, etc. Estas condiciones particulares asociadas a las pruebas estadísticas se recogen de forma más explícita en los capítulos finales de esta obra.

La decisión compara los dos valores del estadístico correspondiente, por un lado tenemos el denominado valor empírico, que se obtiene desde la respuesta de los sujetos que participan en la investigación. En ese cálculo la herramienta principal de cálculo y, sobre todo, de selección de la más adecuada es la Estadística; mientras que en los procesos de cálculo numérico se recurre al uso del ordenador y los correspondientes paquetes estadísticos de análisis de datos. Por su parte la búsqueda del valor crítico o teórico se realiza en las tablas correspondientes. En los análisis mediante ordenador, junto a los valores empíricos, se recoge la probabilidad de error asociada a ese estadístico, lo que nos permite su interpretación en la investigación.

La regla general asociada al contraste de hipótesis y la decisión estadística nos dice que: *cuando el valor empírico del estadístico es mayor que el valor teórico o crítico se rechaza H_0* , ello supone aceptar que las diferencias encontradas son estadísticamente significativas. Naturalmente, esta regla general tiene sus excepciones que se irán analizando en sucesivos capítulos. En este momento debemos entender que esta fase es fundamental en el desarrollo de la investigación pedagógica empírica en el campo socioeducativo, pues nos permite garantizar que los cambios producidos por la intervención del investigador sobre la variable independiente (V_I) ha producido sus frutos y podrá ser recomendada en el futuro en contextos o situaciones similares a la estudiada.

10. La estadística y las ciencias sociales. Posibilidades y límites

1.5. La Estadística y su relación con las Ciencias Sociales

El desarrollo de las Ciencias Sociales en las últimas décadas no se puede entender sin el importante papel que ha desempeñado la Estadística en esta tarea. Así su origen se encuentra unido al interés de los investigadores y científicos por cuantificar los diferentes aspectos sociales de los grupos o comunidades. En este punto queremos poner de relieve los diferentes ámbitos de las Ciencias Sociales en que actúa y aporta sus conocimientos la Estadística.

- a) Educación:** en páginas anteriores hemos puesto de manifiesto las vinculaciones estrechas que existen entre la Estadística y la investigación pedagógica empírica. Ello contribuye al estudio de esta temática en materias como los métodos de investigación en educación, los diseños de investigación, los problemas de la medida, la evaluación, el diagnóstico y la orientación, etc.
- b) Psicología:** los estudios de naturaleza psicológica han contribuido al desarrollo de algunas técnicas estadísticas como el análisis factorial. El estudio del comportamiento de los sujetos, las aptitudes, los rasgos de personalidad, los factores de inteligencia, etc., se basan en la utilización de la Estadística. El estudio se traslada a campos como la Psicología Experimental, la Psicometría y la Psicología Diferencial.
- c) Sociología:** el estudio de los fenómenos y las relaciones sociales forman el cometido principal de la Sociología. Para comprender y valorar el desarrollo de los comportamientos colectivos, describir instituciones sociales, su organización e interrelaciones, el análisis y la comparación de las estructuras sociales subyacentes a los grupos, etc., es preciso recurrir a la Estadística.
- d) Economía:** su cometido consiste en el manejo de datos numéricos. Para su interpretación y valoración es preciso emplear los métodos estadísticos. Entre otros se pueden citar: el índice de precios al consumo, el análisis de mercados, la estimación de la demanda y las series temporales. Además muchas de las teorías económicas recurren a modelos estadísticos para describir los fenómenos económicos. Un campo especial de estudio lo constituye la Econometría y los Modelos Econométricos.
- e) Demografía:** se ocupa del estudio de la población, a través de diversos censos, la distribución por edades o sexo, localización geográfica, profesiones, religión, nacionalidades, tasas de nacimiento o defunción (crecimiento vegetativo) y movimientos sociales migratorios. La simple enumeración pone de relieve el importante papel de la Estadística para perfilar y desarrollar estas tareas.
- f) Administración Pública:** los estudios de la Administración sobre los censos de habitantes, su distribución, las fuentes de riqueza, los temas laborales y sectoriales. Todos estos conocimientos son precisos para abordar una planificación de las actuaciones que son más necesarias en cada zona, de forma que puedan contribuir al bienestar social. Estas tareas para que se puedan desarrollar de forma eficaz necesitan del apoyo de la Estadística.
- g) Humanidades:** las nuevas metodologías de investigación en la historia, la

geografía, la antropología o la literatura, requieren el concurso de la Estadística, que aporta métodos más rigurosos y contrastados en los campos de estudio propios de las humanidades.

h) Ciencias Jurídicas: quizá uno de los campos del derecho en que nos encontramos alguna de las aplicaciones de la Estadística es el de la Criminología, en los estudios de prevención de delitos. En muchos juicios civiles se precisa el concurso de peritos estadísticos para testificar y valorar algunos datos de interés para la justicia.

1.6. Posibilidades y limitaciones de la Estadística

En las páginas anteriores hemos puesto de relieve el importante papel que ha de desempeñar la Estadística, como ciencia instrumental de apoyo a la investigación, en el campo social y educativo. Ello no lleva consigo que servirá para resolver todos los problemas que aquejan al campo de estudio que nos ocupa. Recordemos que nos permite el manejo de datos y nos facilita su análisis e interpretación, pero si la calidad de éstos no es la adecuada se puede cuestionar la validez de las conclusiones. También debemos recordar que la tarea del diseño y elección de pruebas estadísticas es tarea del educador, pues no todos los análisis son pertinentes en todas las ocasiones, como pondremos de manifiesto en páginas posteriores.

Así, la Estadística nos ayudará en la decisión de rechazar las hipótesis de nulidad (H_0), pero las garantías que han conducido a este punto deben ser revisadas y controladas, es decir, no podemos olvidar que sin un buen control sobre las variables extrañas, podemos estar considerando que los cambios generados son achacables a una variable independiente (V), cuando en realidad han sido otros factores o variables no controladas las que han generado los cambios.

Debemos ser conscientes que la elección de las pruebas de contraste va asociada al cumplimiento de una serie de requisitos por parte del modelo estadístico generado por los datos de la investigación. Aunque la Estadística no permitiría usar diferentes pruebas estadísticas, la decisión sobre la más adecuada y la que cumple mejor los cometidos del trabajo y respeta los requisitos del modelo, es del responsable del trabajo. En suma que la Estadística a través de los paquetes de análisis de datos es ciega y puede emplear diferentes procedimientos de cálculo, pero el investigador es el que debe tomar la decisión en elegir unas y descartar otras, para que los resultados finales de los trabajos empíricos sean válidos en la toma de decisiones.

Nota: Este curso es un fragmento del libro: "**Estadística aplicada a la educación**", de los autores Ramón Pérez, José Luis García, Juan Antonio Gil, Arturo Galán, publicado por **Pearson Educación** (ISBN: 9788483226360). Puedes descubrir y adquirir libros de **Pearson** en: www.jetlibros.com.