CE 02 1330 02

Buenas noches, en Revista de Educación nos vamos a ocupar de materias del curso de Adaptación y de Segundo Curso. La primera propuesta gira en torno a la investigación en educación con los profesores de Pedagogía Experimental Adaptación. Después, la profesora María García Milburú nos hablará de cómo enfocar un tema del programa de Filosofía de la Educación y, como ejemplo, hablaremos de la educación moral.

En el último espacio de la revista volveremos a hablar de Pedagogía Experimental, en concreto, de la elección no paramétrica en las pruebas para más de dos grupos, una propuesta del profesor de Segundo Curso, Juan Antonio Gil Pascual. Nuestra primera propuesta, tal como avanzábamos hace un momento, se centra en la investigación en educación de la mano de los profesores de Pedagogía Experimental Adaptación, José Luis García Llamas y Catalina Martínez Mediano. Buenas noches.

Buenas noches. Si les parece, comenzamos con una presentación de este espacio dedicado al curso de Adaptación presumidamente de interés para los estudiantes de esta carrera universitaria. Profesor García Llamas.

Bien, este programa es una continuación del que se emitió en el primer cuatrimestre donde allí presentábamos aspectos generales relacionados con lo que era la investigación en educación, centrándonos sobre todo en el proceso de investigación. Ahora vamos a entrar en una parte un poquito más técnica que se refiere al estudio de los diseños de investigación, es decir, aquella parte que nos va a permitir hacer un control sobre los procesos. Entonces abordaremos en este programa los aspectos generales del diseño y haremos una presentación breve sobre los principales modelos dentro del programa de la asignatura que se corresponde con los diseños de un grupo, los diseños de dos grupos y aquellos de más de dos grupos.

Y comentaremos brevemente las pruebas principales que van a servir para resolver estadísticamente estos diseños. Bien, pues vamos a comenzar por el principio, como se suele decir. Profesora Martínez Mediano, ¿cuáles son los objetivos del diseño de investigación? Bien, vamos a hablar del diseño experimental y claramente de sus objetivos.

La experimentación consiste en modificar deliberadamente y de manera controlada las condiciones que determinan un hecho y en observar e interpretar los cambios que ocurren en este hecho después de actuar deliberadamente sobre él. Kerlinger, un autor clásico en estas materias, define el diseño experimental como el plan, la estructura y la estrategia de investigación concebidos para obtener respuestas a preguntas de investigación y controlar la varianza del diseño. El plan es el esbozo general del proyecto de investigación e incluye desde la formulación de la hipótesis como tentativa para solucionar el problema hasta el análisis de datos.

La estructura del diseño se refiere al esquema del papel que van a jugar las variables de

estudio y la estrategia hace referencia a los métodos de recogida y análisis de datos. Todo diseño de investigación, claro, ésta tiene como finalidad obtener respuestas a las preguntas de investigación o lo que es lo mismo, solucionar los problemas planteados. Arnaud, otra autora clásica, le da un sentido más general al diseño experimental incluyendo en él todos los procedimientos requeridos en una investigación experimental que van partiendo de la formulación del problema a la formulación de hipótesis, la selección de variables independientes, las variables dependientes, el control de las variables experimentales, etcétera, etcétera, hasta el análisis de varianza y la inferencia de las conclusiones del estudio.

En definitiva, el diseño de investigación nos proporciona un modo de poner a prueba una hipótesis. Una vez probada la hipótesis podremos establecer inferencias válidas para poder generalizar desde la muestra en la que hemos trabajado a la población que representa dicha muestra. Profesora Martín Empediano, por lo que nos comenta parece que es importante la validez interna de los diseños.

¿Qué factores afectan a la validez de los diseños, profesora? Claro, la validez de los diseños fundamentalmente son dos, validez interna, es decir, garantizar cómo se ha realizado el experimento y la validez externa o el poder de generalización de esos hallazgos a poblaciones parecidas y situaciones y sujetos parecidos. Pero claro está, lo fundamental es la validez interna, cómo se ha realizado el experimento. Y entre los factores que afectan a esta validez están, por ejemplo, los hechos externos, también llamada la historia, diferentes al tratamiento que pueden interferir con esa variable independiente o tratamiento.

Otro factor puede ser la maduración de los sujetos, si el tratamiento, la investigación es larga, los sujetos pueden madurar biológica y psicológicamente e introducir cambios en interactuar con esa variable independiente o tratamiento. También pueden incidir en la validez interna las pruebas para medir los efectos del tratamiento, porque a veces no discriminan bien a los alumnos, a los sujetos. Luego también puede influir la habilidad del experimentador, que debe cuidar cómo trata el grupo experimental que recibe el tratamiento y el grupo de control del mismo modo.

La regresión estadística, que es la tendencia de valores extremos hacia la media. Y otro factor muy importante que puede influir en la validez interna es la selección de los sujetos, tienen que ser representativos y esta selección no siempre es posible cuando se trabaja con grupos naturales en los colegios. Otro factor es el abandono de los sujetos de la investigación, que puede repercutir en el grupo y en el equilibrio entre los grupos experimentales y de control, además de que se comuniquen los sujetos del grupo de control y experimental.

Estos son los principales factores que pueden afectar a la validez interna de las investigaciones experimentales. Una vez asegurada esta, la validez externa es posible, la generalización de los hallazgos es posible. Bueno, yo creo que entonces ya podemos ir hablando de cuestiones prácticas y ver cómo proceder a la resolución de un diseño de investigación, profesor García Llamas.

Bueno, una vez identificado el diseño y conocido las variables que intervienen, la siguiente fase sería la fase que podríamos llamar más operativa, es decir, nosotros tendríamos que entrar ya en el proceso de recogida de datos mediante los instrumentos elaborados o seleccionados previamente. Esos instrumentos nos van a facilitar unos datos que generan lo que llamamos el modelo estadístico. En función de ese modelo estadístico, la vía de resolución de estos diseños puede ser por la vía paramétrica o por la vía no paramétrica.

Desde el punto de vista de la exigencia y calidad de datos, lógicamente la vía paramétrica exige que se cumplan unos requisitos más rigurosos como, por ejemplo, la normalización de los datos, la distribución normal, que se dé una homogeneidad de varianzas, una medida más rigurosa de intervalo en cuanto que se trabaja con puntuaciones y, en el caso de la prueba de la análisis de varianza, la linealidad, es decir, que los efectos sean aditivos. Ese es un modelo más exigente frente al modelo no paramétrico que, como requisitos de referencia, solamente exige la independencia en las observaciones y los niveles de medida, digamos, más pobres de tipo cualitativo, como son los ordinales, o mediante frecuencias, como son los datos nominales. Si seguimos avanzando un poco más, yo creo que ya podríamos empezar a comentar, aunque sea de forma breve, las ideas básicas relacionadas con los diseños de un solo grupo, profesor García Llamas.

Bien, como dijimos en la introducción, vamos a comenzar con los diseños pre-experimentales o precientíficos, que son los diseños de un solo grupo, son propios de lo que podríamos llamar como investigación aplicada, donde el investigador lo que le importa es resolver un problema y en un contexto muy determinado. Desde ese punto de vista, podemos trabajar con tres modelos básicos, uno que llamaríamos el diseño post-test, es decir, hay una intervención del investigador sobre la variable independiente o experimental y a continuación se miden los efectos, con lo cual es un diseño muy pobre. Después hay otro modelo de diseño que se utiliza bastante, que sería el que llamamos pre-test-post-test, que permitiría medir antes, introducir la variable independiente y medir después, de tal forma que nos permite la comparación entre el antes y el después y las previsibles ganancias que se van a producir al introducir esa variable experimental.

También hay un tercer modelo, que sería el de comparación con un grupo estático o grupo testigo. En cuanto a la resolución, como hemos apuntado anteriormente, lógicamente estos diseños, estos datos del post-test en suma, pues van a tener una determinada calidad y se podrán resolver mediante la vía paramétrica, que tenemos fundamentalmente dos pruebas, la prueba T y la prueba Z, cuya elección dependerá del tamaño de la muestra. Si trabajamos con muestras pequeñas, entendiendo N menor o igual a 30, elegiríamos el estadístico T. Si trabajamos con muestras superiores, pero próximas a 30, es más recomendable T y ya cuando son muestras grandes se utiliza el estadístico Z como contraste.

En cuanto a la vía no paramétrica, fundamentalmente se puede hablar de la prueba de eje cuadrado, que nos sirve precisamente para cumplir o para analizar el requisito de normalidad en la distribución de datos. Es una de las utilidades principales que nos ofrece la prueba de eje

cuadrado. También está la prueba de Kolmogorov, la prueba de McNemar, que es una prueba pretest-post-test de significación de cambios.

Y, por supuesto, también es interesante que el alumno sea consciente de que aquellos coeficientes que estudió en el primer cuatrimestre ahora, desde la estadística inferencial, tendrá que analizar si la correlación, al margen de que sea alta, baja, media, es estadísticamente significativa. Es decir, si la relación entre variables se puede explicar por efecto del azar o porque realmente existe una relación de dependencia entre las variables. Y eso lo va a estudiar también en estos temas.

Otra cuestión básica es ver cómo se resuelven los diseños de dos grupos, profesora Martínez Mediano. Pues muy bien, antes estaba hablando de los factores que afectan a la variedad interna de los diseños y, efectivamente, un diseño de dos grupos puede facilitar ese control de variables para tratar de asegurar que los efectos del tratamiento se deben a ese tratamiento, de modo que los diseños de dos grupos, uno actúa como grupo experimental que recibe el tratamiento y otro actúa como grupo de control que no recibe el tratamiento. Por supuesto, el punto de partida de esos dos grupos debe ser igual en variables que correlacionan con esa variable que queremos probar su eficacia.

Por lo tanto, tenemos que contar que las medias iniciales de esos grupos en esas variables relacionadas con el tratamiento son iguales, no difieren significativamente y nuestra finalidad es probar después del tratamiento, cuando midamos los resultados de ese tratamiento a través de la variable dependiente, tendremos que probar si esas medias de grupo de control y grupo experimental son diferentes, qué tan diferentes son, para probar que esas diferencias se deban al tratamiento y no a efectos del azar y esa es la significación de hipótesis. Por supuesto, es importantísimo contar con ese punto de partida la igualación de los sujetos en variables relacionadas con el tratamiento y controlar después todo el proceso. Seguimos hablando en cierto modo también de cómo asegurar que esos efectos se deban al tratamiento y cómo controlar que otras variables experimentales no interactúen con la variable independiente experimental, que es el tratamiento y que es el objeto de nuestra investigación.

Entonces, para asegurar eso, lo primero que debe hacer el investigador es conocer muy bien el tratamiento, la variable experimental que va a tratar de probar. Naturalmente, los diseños de investigación hemos dicho que parten de la idea de solucionar un problema. Entonces, la hipótesis es ya la relación entre nuestro tratamiento variable independiente con la variable dependiente o resultados que pretendemos obtener.

Entonces, la mejor garantía de la validez interna es la pericia, la habilidad del investigador. El investigador debe conocer muy bien esa variable independiente, el tratamiento, conocerlo muy bien para permitirle diferenciar esa variable y otras variables extrañas que puedan interferir y detectarlas, detectarlas para controlarlas. Así, por ejemplo, si pretendemos resolver el problema de la baja comprensión lingüística de los alumnos de determinado curso de educación secundaria obligatoria, el investigador debe elaborar un programa específico para

desarrollar esa comprensión lingüística en los estudiantes.

Esa es la variable independiente, ese es el tratamiento. Conocedor de la importancia de la comprensión lingüística sabrá que otros factores pueden incidir e interactuar con su tratamiento, como puede ser que otros profesores de idiomas extranjeros estén también trabajando en esa línea. Entonces, deberá evitar en ese momento, o al menos tratar de controlarlo, que los efectos de su programa, de su tratamiento específico, se deban a su programa o estén interactuando con otros.

Entonces, también hemos dicho que la instrumentación es muy importante. Cómo aplica ese programa al grupo experimental y cómo el grupo de control está recibiendo un no tratamiento o algún nivel más bajo de ese tratamiento. Y luego hemos dicho que afectaba a la validez interna del sujeto las pruebas.

Entonces, la medida de la variable dependiente debe ser muy cuidada y ahí también está influyendo la habilidad del investigador en el conocimiento de ese tratamiento y de ese objetivo de su investigación. Entonces, tendrá que elaborar unas pruebas lo más precisas posibles y lo más representativas de los contenidos en los que ha estado incidiendo, totalmente relacionados con lo que quiere mejorar. Pruebas, seleccionará unos textos, hará unas preguntas claves para comprobar la comprensión de los alumnos.

Utilizará, pues, algunas pruebas de debate o algún ensayo para que el alumno argumente las ideas claves que están contenidas en ese texto, de modo que se pueda evaluar con plena garantía que sí ha mejorado o no esa comprensión lectora, esa comprensión lingüística que era el problema a resolver. Bien, una vez que se recogen los datos a través de esa variable dependiente, a través de la medición con instrumentos de la variable dependiente, después de un tiempo en el que se ha estado trabajando la variable independiente de tratamiento, es el momento en el que entran las técnicas estadísticas para comprobar, para ver si la hipótesis cabía de que con este tratamiento se mejorará la comprensión lingüística, se confirma. Entonces, por ejemplo, un diseño de dos grupos que ha sido medida la variable dependiente con una prueba donde se puede alcanzar el nivel de medida de intervalo, pues, utilizaría la prueba estadística T para ver la diferencia de medias.

La prueba consistiría en comparar las diferencias empíricas de las medias entre el grupo experimental y el grupo de control y dividirlas entre el error típico de esos grupos. El valor empírico que se obtiene en esa prueba sería el valor empírico de T, el valor real que se ha producido, y sería comparable, tendríamos que compararlo con el valor teórico de T que se recogen en las tablas. Entonces, el investigador funciona con unos niveles de significación previamente fijados y dice si mi hipótesis se confirma con un nivel de significación del 0,05 o el 0,01, depende de la exigencia del investigador, si nuestro valor empírico T es mayor que el valor teórico de la tabla, nuestra diferencia significativa al nivel de significación previamente fijado.

Luego, podremos afirmar que las diferencias entre los dos grupos se deben a los efectos del

tratamiento y no a los efectos del azar, y si hemos controlado, pues, la selección de los sujetos, la asignación de los tratamientos, sobre todo a través de técnicas del azar, pues, podremos realizar inferencias de los hallazgos de nuestra investigación. Bien, pues yo creo que para completar un poco este repaso, evidentemente los alumnos tienen que estudiar, tienen su material didáctico para profundizar en cada uno de estos aspectos, pero nos quedaría ver cuáles serían las principales pruebas estadísticas para la resolución de los diseños de más de dos grupos, profesor García Llamas. Enlazando con lo que comentaba la compañera, la profesora Martínez Mediano, yo creo que lo fundamental es que el alumno sea consciente de ese proceso, ¿no? Entonces, ya a la elección de la prueba tendrá que conocer las características del modelo.

Ella ha hablado del modelo paramétrico mediante la prueba T, bueno, y si somos tarragantes, la prueba Z, pero, naturalmente, también tenemos la opción no paramétrica, donde ahí dentro de los diseños de dos grupos habría que destacar fundamentalmente la prueba U de Mann-Whitney, que trabaja con datos ordinales, y también la prueba de Kolmogorov-Smirnov. En grupos independientes y para grupos relacionados, la prueba de Wilkerson. Es decir, yo creo que lo fundamental es que el alumno entienda el proceso y que sepa que la elección de esa prueba va a depender de las características del diseño que se está resolviendo y, bueno, del modelo estadístico que han generado los datos.

En cuanto al diseño de más de dos grupos, pues yo creo que los comentarios que ha hecho la profesora son válidos también, naturalmente, para estos diseños de más de dos grupos, lo que ocurre que, claro, las pruebas estadísticas que nos van a permitir su resolución van a ser distintas. Por ejemplo, si estamos ante un diseño de más de dos grupos independientes, la prueba que tendríamos que elegir sería la prueba de Lanava, análisis de varianza, que se basa en el estadístico F. Este estadístico nos va a establecer diferencias entre los tratamientos de los tres o más de tres grupos. Una vez determinado ese valor de diferenciación, y si F es significativa al compararlo con el valor de las tablas, entonces el investigador lo que tendrá que probar a posteriori es entre qué grupos de esos tratamientos las diferencias son estadísticamente significativas o no lo son, y para ello el alumno debe saber que dispone de dos pruebas principales, una que es la prueba de Tukey, que utilizaría en aquellos casos en que trabajemos con grupos igualados, es decir, el número de sujetos de los grupos sea similar, sea idéntico, y otra sería la prueba o el contraste de ese F que se utilizaría cuando el número de sujetos de los grupos es distinto.

Por supuesto, si los grupos son relacionados, ahí tendremos una prueba, una variante de Lanava, que es el Lanava de bloques, que es una de las ideas también que expuso la profesora, es decir, lo que se trata en ese caso es de controlar la influencia de alguna variable extraña que el investigador considera que es relevante para la investigación, y entonces lo que hace es formar grupos de sujetos o conjuntos de sujetos que puntúan de forma similar en esa variable, de tal forma que, asignados a tratamientos distintos cada uno de los componentes de esos conjuntos, si al final del proceso, en el posttest, se producen diferencias significativas, probablemente sean achacables a la intervención del investigador sobre la variable

experimental. Entonces, en ese caso está el Lanava de bloques, análisis de varianza de bloques con grupos igualados, donde en función, naturalmente, de esa variable de bloqueo interviene una fuente de variación que es debida al bloqueo. En cuanto a la opción no paramétrica, pues, por ejemplo, tenemos unas variantes del análisis de varianza, por ejemplo, en el Lanava simple, la alternativa no paramétrica, cuando los datos no reúnen esos requisitos, sería el análisis de varianza por rangos de Kruskal-Wallis, o también llamada prueba de Kruskal-Wallis, donde, lógicamente, como los datos tienen menor nivel de precisión, se transforman las puntuaciones en rangos y se opera con los rangos.

En el caso de las pruebas no paramétricas para grupos relacionados o de bloqueo, ahí tenemos la prueba de Konrad y, fundamentalmente, el análisis de varianza por rangos, o prueba de Friedman, también utiliza los rangos, si bien la diferencia en que en grupos independientes la asignación se hace de forma global, en los grupos relacionados lo que importa es ver qué ocurre con esos sujetos que, previamente, estaban igualados, con lo cual, la asignación de rangos se va a hacer por filas. Naturalmente, después, habría que resolverlo e interpretar estos valores. Yo creo que podemos dedicar los últimos minutos a algunas recomendaciones para los alumnos.

Quizá tengamos, después de las vacaciones de Semana Santa, otro pequeño espacio para ese momento del examen, para esas recomendaciones finales, pero en este momento de estudio, profesora Martínez Mediano, ¿qué recomendación o qué sugerencia nos puede hacer? Las recomendaciones van a ir dirigidas a la resolución práctica de los diseños en esa segunda prueba de pedagogía experimental. Por supuesto, el alumno tiene modelos de examen, el alumno que asista a tutorías recibe información del tutor sobre cómo enfocamos los exámenes, pero tiene que tener una clara fundamentación teórica el alumno para resolver los diseños de investigación práctica, elaborar el esquema donde diferencie los distintos grupos que forman el diseño, cómo se han seleccionado los sujetos, cómo se ha asignado el tratamiento, si existe prueba pretés o no, cuál es el tratamiento o variable independiente y la variable dependiente. Tiene que decidir en función de ese esquema del diseño, la prueba estadística adecuada, aplicarla, resolverla, contrastarla con las tablas, el valor empírico que obtenga y con el valor teórico, interpretar los resultados y realizar inferencias en las conclusiones.

Tiene que ser cuidadoso en buscar la coherencia entre las distintas partes del diseño y, sobre todo, también tiene que repasar las operaciones para buscar evitar errores y comprender la coherencia de los resultados con lo esperado de acuerdo con el planteamiento del diseño. Muchas gracias a los dos por habernos acercado a esta segunda parte de la asignatura pero también sobre todo a esas pautas para realizar diseños de investigación, algo básico en este caso para los alumnos del curso de adaptación y para cualquier persona que esté haciendo esta carrera en la Facultad de Educación. Gracias y buenas noches.

Recuerden que los programas ya emitidos se pueden volver a escuchar a través de las páginas web de nuestra universidad. La dirección www.uned.es. A los centros asociados también se envían copias en audio casete de los programas ya emitidos pero tardan un poco más en llegar.