Estadística de las ciencias humanas. La estimación. Guión número 4. Introducción. Guión número 5. Según nos dice el profesor Ríos, al hablar de dependencia aleatoria, introducimos la idea de modelo para expresar matemáticamente una relación de dependencia amplia al objeto de lograr científicas.

la previsión. El modelo que utilizamos en la estimación es paralelo al de la regresión en cuanto que es una forma del modelo lineal general. Las estimaciones de parámetros las realizamos a partir de las distribuciones que nos permitirá, según el caso, conocer la dependencia entre dos o más variables. Profesor, ¿por qué hablamos de estimación del parámetro? Sabemos de la dificultad de obtener un parámetro o valor de la población. Para paliar esto hemos recurrido a la muestra. Ahora bien, una vez obtenido el valor de la muestra, el estadístico, inferimos una estimación del valor que correspondería a ese parámetro si hubiésemos logrado obtenerlo, algo que, como sabemos, resulta en la práctica imposible.

recuerdo mal, el pasado curso hablamos de diferentes modos de estimación, ¿verdad? Efectivamente. La podemos realizar por puntos y por intervalos. Veamos en qué consiste cada una de ellas. Lo que los psicólogos perseguimos para nuestros estudios es lograr resultados que tengan la suficiente validez para estimar por ellos a una población. Inferir unos datos que ofrezcan la garantía científica suficiente en nuestras generalizaciones sobre un colectivo. Claro que a nadie escapa la observación de que es imposible poner en funcionamiento al colectivo total. Así aparecen los métodos de muestreo con los que hemos de obtener la muestra, que es un grupo representativo de la población a la que esa muestra pertenece. Métodos de muestreo, ha dicho, profesor. Sí, sí. Métodos utilizados en la elección de la muestra y que sean los más usados. A. Aleatorio simple. Consiste en elegir al azar de toda la población objetos, objetos de nuestro estudio, los individuos.

que compondrán nuestra muestra sin efectuar reemplazamiento. Entonces, los que estamos aquí podemos ser un grupo aleatorio, ¿no es así? Bueno, si consideramos todos los posibles estudiantes de vuestro curso de psicología en toda España, sí. Aunque relativamente, porque otros también querrían estar aquí, y si no lo están, no es precisamente por efectos del azar. Pero sigamos con los métodos y esperemos que vosotros casi solo sacaréis las conclusiones oportunas, ¿vale? Sí, de acuerdo. Bien. B. Otro método es el estratificado. Se divide la población en estratos y de cada uno de ellos se toma al azar. Supongamos que vamos a aplicar un test a un grupo de 100 niños de octavo de eje B. Para ello, podemos seguir, por ejemplo, un criterio geográfico. Entonces haríamos el reclutamiento eligiendo cuatro de Canarias de ese curso, al azar, uno de Melilla, otro de Ceuta, y así sucesivamente hasta obtener la muestra de la misma clase de estudiantes. Y, por último, el reclutamiento de la clase de estudiantes. Y, por último, el reclutamiento de la clase de estudiantes. Representativa de toda la población.

española. La estratificación podría ser uniforme, proporcional, óptima, valoral. Por último, cuando quisiéramos afinar mucho en nuestro estudio, usaríamos el controlado. Este método tiene como base al anterior, por cuanto divide la población también en estratos. Ahora bien, después escoge de cada estrato la parte proporcional al número de componentes de la población que hay en él. Supongamos la muestra anterior. Ahora, para escogerla, deberíamos ver cuántos habitantes hay en España. Supongamos 40 millones. Pues si vemos las regiones como antes, tenemos Cataluña, 4 millones de habitantes. Luego

escogeríamos 10 individuos de su población de octavo de básica. Galicia, 2 millones de habitantes. Escogeríamos al azar 5 individuos. Canarias, Ceuta y Melilla, 1 millón 400 mil. Pues 3.

Y así hasta completar los 100 respetando la proporción correspondiente en la población. Hay otros métodos, como hemos dicho, entre los que están muestreo con reemplazamiento, muestreo por áreas, muestreo polifásico. Objetivos de la estimación. Antes de señalarlos, conviene que precisemos algunos conceptos muy utilizados en nuestro argot. Hablamos de precisión de un estadístico. Estadístico es cualquier valor obtenido en una muestra. Así, estadístico puede ser la media de esa muestra en un test. Estadístico puede ser una puntuación directa de un sujeto de la misma. Precisión es la exactitud en la medida realizada. Podríamos decir que en el grado en que se correspondan el valor real y el obtenido. Un test será preciso cuanto más exacto sea el resultado.

sea en la apreciación de la cualidad que intentamos calibrar. Ahora bien, el hecho de que usemos indistintamente precisión o fiabilidad se debe a que, si hablamos de exactitud en la medida, nos referimos de algún modo a una sola situación, pero no podemos elaborar un test para cada ocasión. Por otro lado, nos interesa que los resultados que obtengamos cada vez que se repitan las mismas condiciones sean iguales o aproximadamente parecidos, que haya constancia. Si se da esta constancia, que es tanto como la repetición de los mismos resultados en iguales circunstancias, podemos hablar de fiabilidad. Fiabilidad es, pues, la constancia en la precisión. En la medida en que los resultados que obtengamos a distintas aplicaciones podamos ver que se ajustan más a las verdaderas, podremos hablar con más fiabilidad, o sea, con mayor seguridad de que nuestras estimaciones son ciertas. ¿Hasta qué punto es necesario?

que se aproximen nuestras estimaciones a la verdad? En primer lugar, siempre los valores estimados del parámetro son función de la variabilidad de la muestra con la que estamos trabajando. Por lo general, partimos del hecho de que la distribución del estadístico considerado en infinitas muestras del mismo tipo y tamaño que la considerada es normal. Por ello, al estimar un intervalo dentro del cual esperamos se encuentre nuestro parámetro, lo hacemos con una seguridad en términos de probabilidad. Así, aparecen el concepto de nivel de confianza y coeficiente de riesgo. Si nuestra estimación tiene la posibilidad de ser cierta, el 95% de las veces corremos el riesgo de equivocarnos el 5% restante. Profesor, ¿por qué no vemos un caso en la práctica? Veamos. Supongamos que estamos trabajando con una muestra de 170 alumnos a los que hemos aplicado un test, cuya media vale 30. Su derivación estándar... 10.

Y queremos ver con un nivel de confianza del 5%, o sea, con una seguridad del 95%, cuál será el intervalo en el que se halla el parámetro al que esta muestra representa. El proceso estadístico sería. Primero, obtenemos el valor en zetas, o sea, en desviaciones de la curva normal, una distancia que deje a ambos lados de la media el 95%, con lo cual veremos el valor de la ordenada que deje entre sí y la ordenada correspondiente a la media la frecuencia relativa que supone el 47,5% de los casos a cada lado. En las tablas de la curva normal vemos que a esa ordenada corresponde 1,96 signos o desviaciones. Ya tenemos el valor 1,96 en zeta o desviaciones de la curva normal. Segundo. Aparece ahora el concepto de error típico de la media.

o error de la media, aplicable a cualquier estadística que es. Su fórmula es distinta para cada estadístico. Aquí, y para la media, es siempre igual a sigma submedia igual raya de fracción. En el numerador, la desviación estándar, que en nuestro caso tiene como valor numérico 10. En el denominador, raíz cuadrada de n-1, o sea, raíz cuadrada de 169. Así obtenemos 10 partido por 13, que es 0,76. Este 0,76 es el error de la media en nuestra muestra. Tercero. Ahora bien, con esta muestra, inferimos a toda la población a la que pertenece esa muestra, o sea, queremos inferir el valor paramétrico. El intervalo dentro del cual, en el 95% de los casos, se hallará comprendido el parámetro. Para ello, hemos declarado que el valor paramétrico es el valor de la media. Gracias.

calcular el error máximo que estamos dispuestos a admitir en nuestras estimaciones. Su fórmula es el producto entre el error de la media y el valor en zetas correspondiente al intervalo del 95% en la curva normal. Obtenemos pues 0,76 por 1,96. Resulta ser 1,48. Cuarto. En cuarto lugar tomamos como muestra media 30 y le sumamos y restamos 1,48 con lo que tenemos 31,48 como valor superior a ella y 28,52 como valor inferior. Concluimos pues. Al nivel de confianza del 5%, o sea con una seguridad de que ocurrirá en el 95% de nuestras observaciones, la media paramétrica que buscamos se hallará comprendida entre 31,48 y 28,58. 52.

El intervalo comprendido entre ambas puntuaciones, 31,48 y 28,52, lo llamamos intervalo confidencial del 5%, puesto que a ese nivel de confianza hemos hecho nuestra afirmación. Cada uno de esos valores, 31,48 y 28,52, que limita dicho intervalo reciben el nombre de límites fiduciales. Esto es lo que conocemos como estimación por intervalo. Podríamos poner otros casos, pero sería demasiado largo. En la estimación por punto se realiza la inferencia obteniendo para el parámetro un solo valor, que aceptamos como representativo de la población. Gracias por ver el video.

¿Qué aplicaciones más usuales y qué problemas podemos resolver con la estimación? En principio son varias las aportaciones de la estimación al campo de la estadística, así como sus desviaciones, pero por razones de brevedad vamos a ceñirnos a, primero, la estimación de parámetros. Segundo, la contrastación de hipótesis. La estimación en contrastación a verificación de hipótesis que supone la investigación de un conjunto de reglas y principios a los que deberíamos someter toda nueva teoría para que nos pueda ofrecer suficientes garantías científicas. Por todo ello, mientras no realizamos este tipo de comprobación, no podemos sostener en forma definitiva ninguna hipótesis. El profesor Popper señala que la estimación de parámetros es un tipo de hipótesis que no se puede sostener en forma definitiva. Se necesitan cuatro procedimientos para llevar a cabo la contrastación de hipótesis. El que se refiere a la comparación...

con otras teorías, tratando con ello de averiguar si la teoría examinada constituye una novedad o por el contrario es algo ya superado o se encuentra incluido en otras anteriores. B. El referente a la forma lógica de la teoría, con objeto de determinar su carácter, es decir, si tenemos una teoría o es empírica o es lógica. C. El basado en la consistencia interna del sistema o coherencia, de modo que así realizamos una comparación lógica entre las distintas conclusiones. D. Por el carácter eminentemente práctico de nuestra materia, en último caso nos referimos a la aplicación empírica de una serie de conclusiones que en definitiva han sido deducidas de la teoría.

En este punto tenemos obviamente un contraste de tipo inductivo, basado en formulación de juicios sobre la realidad concreta que significa toda una teoría cuya exigencia es que el método sea cuantificable. Claro, hay que señalar que toda teoría, aunque sea cuantificable, no resulta fácilmente verificable. En resumen, las aplicaciones de la estimación son múltiples. Hemos visto en la media respecto a su nivel de confianza. También es aplicable al nivel de significación, donde podemos ver la diferencia significativa o no entre estadísticas de distintas muestras. Deberíamos insistir en el carácter aleatorio de nuestras estimaciones. Así vemos cómo nuestros intervalos son aleatorios. Esos mismos intervalos pueden ser multiplicados.

Pueden ser simétricos o asimétricos. Si el alumno desea ampliar estos contenidos, la Sexta Unidad Didáctica de Estadística de las Ciencias Humanas le ofrece más información. Asimismo, pueden consultar el libro de Murray Spiegel y el de Mariano Gela, que siguen los alumnos de la Universidad Complutense. Si desean más información, pueden dirigirse al profesor Alonso Lozano, Estadística de las Ciencias Humanas de la Universidad Nacional de Educación a Distancia. Gracias.