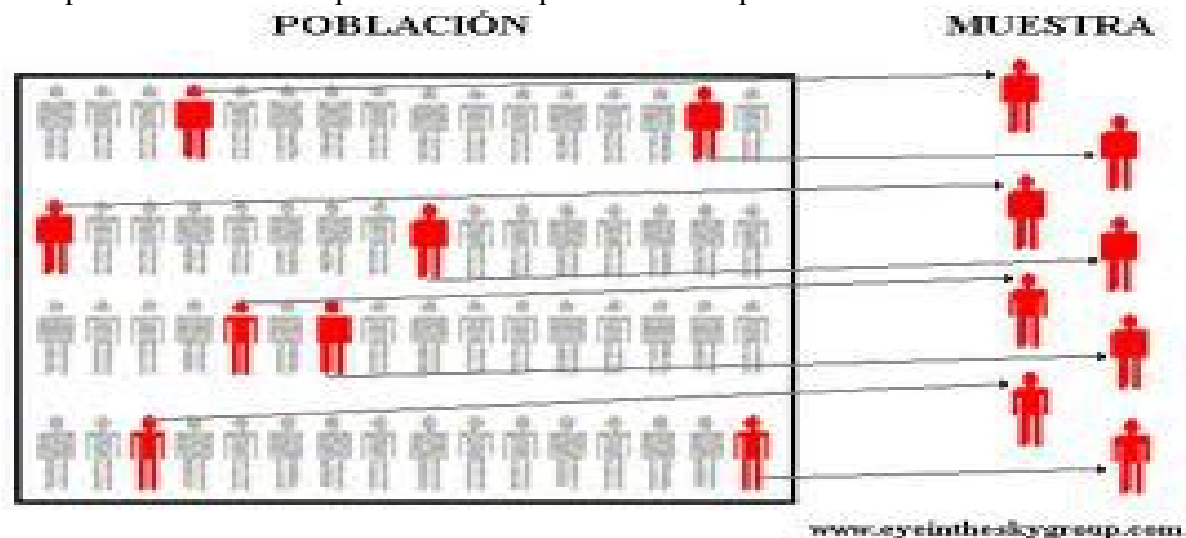


CONCEPTO Y FUNCIONES DE LA ESTADÍSTICA

La estadística es una forma de trabajar muchos valores por medio de un método empírico, y es parte de las matemáticas y como arte de esta es una ciencia. Pero una ciencia un poco diferente porque a diferencia de las matemáticas no es exacta, esto se debe a que la estadística trabaja con los fenómenos del mundo en especial los humanos y por tanto no son perfectos. Así por medio de una metodología se buscará ver la diferencia que existe entre la realidad y la teoría.

Para la educación social y la pedagogía la función de la estadística es permitir entender los artículos científicos, mejorar la evaluación, facilitar en la toma de decisiones y mejorar la profesión socioeducativa.

Por medio de fórmulas y gráficas nos permite entender las características de un gran grupo de sujetos, a la vez que nos permite comprobar si los instrumentos de recogida de información son buenos por medio de la **fiabilidad y la validez**. Permite comprobar si hay diferencias entre dos grupos, o si una característica provoca diferencias en otra característica. También nos permite intentar con un grupo pequeño conocer un grupo de mayor tamaño. Así a los valores de los grupos pequeños con los que trabajamos llamamos muestra, mientras el grupo general es la población. Los resultados que conseguimos de la muestra recibirán el nombre de estadísticos y la información que extrapolamos de esta muestra para conocer la población es el parámetro.



1.3 TIPOS DE ESTADÍSTICA

En general hay dos tipos de estadística que se pueden diferenciar. Estas dos partes se diferencian según las tareas a las que deben enfrentarse y sus objetivos. Estos tipos son la **estadística descriptiva** y la **inferencial**.

1.3.1. ESTADÍSTICA DESCRIPTIVA

Esta etapa de la estadística se basa en el análisis de los resultados recogidos en la muestra, se basa en la creación de estadísticos, que son valores con los que intentamos conocer de manera general el grupo con el que hemos trabajado directamente. Estos estadísticos son indicadores que nos explican las características de los grupos con los que trabajamos.

El problema de esta etapa de la estadística es su limitación ya que solo nos permite conocer el grupo en concreto con el que se trabaja, y a lo sumo nos permite comparar dos muestras pero no nos permite conocer al grupo general ni tomar decisiones con respecto a ellos.

Por ejemplo, si en un grupo A se obtiene una media de 5,25 y en otro grupo B de características similares 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.

Pro 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. Es un resumen de las características de un grupo en concreto, con el que trabajamos directamente.

Dentro de esta estadística hay dos tipos de estadística la ***estadística descriptiva univariada*** que es muy simple ya que solo intenta conocer las características de la muestra en una única variable y la ***bivariada*** donde analizamos los datos de más de una variable, por medio de cálculos como las correlaciones. Cuando hay más de dos variables se llamará multivariada, aunque en este curso no se trabajará.

1.3.2 ESTADISTICA INFERENCIAL

Esta parte de la estadística va un paso más allá de la estadística descriptiva. Ya que lo que busca esta etapa es conocer a la población ya que en la estadística descriptiva solo se conoce al grupo con el que se trabaja. Como investigadores utilizamos la estadística porque nos interesa conocer las características generales de las personas de una ciudad, un país, un centro educativo... pero en la estadística descriptiva no llegamos a conocer a toda la población sino al reducido grupo del que hemos recogido toda la información. De ahí la necesidad de la estadística inferencial que es la que se encarga de a partir de los estadísticos de la muestra, conocer al grupo general la población. Aportando lo que llamaremos parámetros que son cálculos que nos hablan de la población. El problema es que no trabajamos directamente con la población, por lo tanto no podemos hablar de que poseamos un cien por cien de seguridad en estos cálculos, porque al no poseer todos los datos (solo tenemos los de la muestra) podemos cometer cierto margen de error, que se tiene en cuenta por medio de intervalos de confianza. A partir de estos parámetros se darán conclusiones sobre la población.

Si en el ejemplo anterior veíamos que un grupo tenía una muestra de 5,75 y otro de 5,25. La estadística inferencial nos mostrará si la diferencia que hay entre ambas medias es significativa, o sea si esa diferencia es real en la población o es pura casualidad, azar. También nos puede interesar si el estadístico que obtenemos de la media es similar al de la población, o si la relación entre dos variables que hemos sacado de una muestra solo existe en la muestra o por el contrario la relación entre variables existe también en la población.

Gracias a esta estadística inferencial podemos extraer conclusiones.

En este tema tenemos que hablar de que es el ***nivel de confianza N.C.***, cuando hablamos del nivel de confianza tenemos que tener en cuenta que nosotros vamos a tener que realizar muchos cálculos diferentes, y tenemos que extrapolar los datos de la muestra a la población, prácticamente tendremos que hacer algo de brujita adivinando valores de la población. Claro está que estamos especulando y por tanto podemos confundirnos. Aquí entra en juego el nivel de confianza, ya que cuanto más confianza tengamos, menos nos confundiremos y mejor nos saldrá, también seremos menos estrictos porque confiamos en nosotros mismos. Lo normal en pedagogía y educación social trabajaremos con un nivel de confianza del 95% vamos que pensamos que vamos a acertar por lo menos 95 de cada 100 veces, o un 99% que confía que acertaremos 99 de cada 100 veces. Aunque esto trae implícito un problema, que si yo acierto 95 de cada 100 veces, significa que yo fallaré 5 de cada 100 veces. O si trabajamos con un nivel de confianza del 99% eso significa que fallaré 1 de cada 100 veces. Ahora tenemos que tener en cuenta un pequeño detalle que en estadística nunca trabajamos con valores en %, este tipo de valores solo se utilizan para dar a conocer datos a un público en general que no tiene amplios conocimientos de estadística, pero nosotros como investigadores tenemos que conocer los valores en proporción, vamos valores que van de 0 a 1, por lo tanto yo no me confundiré un 5% sino un 0,05, que es el margen de error que nos fijamos como investigadores por nuestra culpa, por nuestro fallo lo que a partir de ahora llamaremos ***nivel de significación*** que es lo contrario de el nivel de confianza que es el porcentaje de aciertos que comento.

En resumen en la estadística inferencial hay dos grandes campos. Por un lado la estimación de parámetros, cálculo de las características de una población, y por otro lado está el contraste de

hipótesis , que busca comprobar que la hipótesis que guía nuestro estudio es confirmada o rechazada por los datos que nos aporta la muestra.

1.4. EL PAPEL DE LA ESTADÍSTICA EN EL PROCESO DE INVESTIGACIÓN

1.4.1. EN EL PROBLEMA DE INVESTIGACIÓN

Lo primero que se tiene que hacer en una investigación científica es identificar el problema, Tenemos que saber si este problema es resoluble ya que no podemos trabajar con un problema que sabemos desde un principio que no es resoluble, porque no es medible. Yo como científico no puedo trabajar en un estudio sobre el alma humana , dado que se que el alma no es medible , entonces no podré utilizar ninguna técnica o modelo estadístico para la comprobación de la teoría. De esta idea clave, que todo ***problema tiene que ser resoluble***, podemos extraer otras conclusiones :

En primer lugar sabremos que un punto importante para el estudio estadístico es conocer las variables que vamos a estudiar y cuales serán sus funciones, o sea que papel hace en la investigación, como serán recogidos, para no llegar a un camino sin salida en la investigación.

1.4.2. EN LA FORMULACIÓN DE HIPÓTESIS Y OBJETIVOS

La creación de los objetivos y las hipótesis es el segundo paso que debe realizar un investigador. Pero es básico no confundirse estos dos términos. Cuando nos referimos a los objetivos estamos refiriendonos a lo que busca el investigador, cual es la meta que desea alcanzar y el porqué realiza todo este proceso, Son descripciones de los intereses del investigador , una mera descripción de lo que busca encontrar y que por tanto carece de imparcialidad. Es lo que quiere ***comprobar*** el investigador. Mientras que las hipótesis tienen un cariz científico es la guía de la investigación , es lo que se busca contrastar por lo tanto es ***contrastable*** . La hipótesis tiene que ser científica ya que es el inicio de la investigación y por lo tanto a de ser imparcial. No ha de presuponer ninguna relación entre variables. Como veremos más adelante.

Otro punto relacionado con las hipótesis es que como estas son la guía de la investigación nos indican cuales son los cálculos, los estadísticos, que debemos buscar para resolver el problema. Las hipótesis que estudiaremos pueden ser o ***unilaterales o bilaterales***, es decir, Hay ocasiones en las que un investigador solo esta interesado en comprobar una situación en concreto , en ese caso las hipótesis serán unilaterales, pero si el investigador deja una investigación abierta que quiere comprobar si hay diferencias entre las variables sin importarle si la diferencia es a favor de uno u otro grupo en estos casos es bilateral. Este concepto se entenderá mejor con algunos ejemplos. Para ello usare el trabajo del 2015 y el examen de septiembre de 2014. Comencemos con el trabajo del 2015.

Un grupo de profesores de bachillerato preocupados por la calidad de la enseñanza piensa que los estudiantes que cursan el bachillerato internacional aprenden más que los que siguen el bachillerato tradicional, pero no tienen pruebas de su intuición. Para contrastar su hipótesis, desean probar que los estudiantes que cursan el Bachillerato Internacional (BI) obtienen mejores notas en Selectividad que los que cursan el Bachillerato Ordinario o tradicional (BO).

En este trabajo estamos viendo un contraste de hipótesis unilateral , y ahora os explicaré porque, en un principio los profesores buscan demostrar que en el bachillerato internacional los alumnos aprenden más , por lo tanto ya solo miran en una dirección solo se plantean el hecho de que el bachillerato internacional es mejor al ordinario, en ningún momento se plantean el hecho de que el bachillerato ordinario sea mejor. Al solo mirar a un lado es unilateral.

Por el contrario en el examen de 2014, el tema era practicamente el mismo pero variaba en su visión de la hipótesis:

Un investigador quiere comprobar si los estudiantes del bachillerato internacional (BI) obtienen distinto rendimiento global en selectividad respecto a los estudiantes que cursan el bachillerato estándar (BE). Para ello, ha seleccionado dos grupos de estudiantes, según el bachillerato que estaban cursando, recogiendo finalmente la nota media obtenida en la prueba de selectividad (escala 0-12). Los resultados del contraste estadístico son los siguientes ($\alpha=0.05$):

En este caso aunque el problema es practicamente el mismo aquí el investigador lo que busca es saber si hay diferencias entre los bachilleratos, dejando abiertas las posibilidades, puede por tanto ser mejor el

ordinario o el internacional indistintamente. Al abrir el punto de mira de nuestro estudio pasa a ser bilateral, ya que ambos grupos pueden poseer valores más altos.

Realmente la situación puede complicarse un poco más ya que cuando nuestro estudio es unilateral, hay un grupo que es mayor que el otro, podemos especificarlo un poco más y decir si es unilateral izquierdo o si es unilateral derecho. Cuando especificamos esto, no es tan complicado como parece. En primer lugar tenemos que ver que en el tipo de experimentos con los que trabajareis vosotros en esta asignatura, solo vais a trabajar con dos grupos, por lo general codificados como 0/1. Lo que teneis que hacer para saber si es unilateral derecho o izquierdo es simplemente colocar las variables en orden por ejemplo en este caso los grupos eran bachillerato ordinario y el internacional. Codificados 0 el ordinario y 1 el internacional. En el caso del trabajo que era unilateral, nos preguntamos hacia donde tenemos que mirar y pensando que nosotros estamos en medio, miraremos hacia el lado derecho porque creemos que el internacional es mayor que el ordinario. No nos interesa mirar a la izquierda. Si por el contrario hubiesen dicho que el bachillerato ordinario sacan mejor nota que en el bachillerato internacional en este caso sería unilateral izquierdo porque solo miraría a la izquierda al bachillerato ordinario que supongo que sería mejor.

Dentro de los problemas de investigación un punto relevante es el saber que papel juega cada variable dentro del estudio estadístico según la intención del investigador. Las investigaciones con las que trabajareis solo tendrán dos variables, aunque pueden haber muchas más, estas variables tendrán funciones diferentes según lo que quiera el investigador. Así podremos encontrar variables dependientes y variables independientes. Antes de explicarla veremos cuales son los objetivos del investigador que afectarán en esta clasificación de las variables. El investigador buscará demostrar que una variable produce cambios significativos en la otra variable, un ejemplo demostrar que la variable horas de estudio afectan a la variable nota. En este caso las horas de estudio serán la variable independiente, porque el investigador si que puede controlar las horas de estudio que dediquen sus estudiantes, mientras que la variable dependiente que es la variable que el investigador no puede controlar serán las notas porque el investigador no decide las notas, las notas son las que se merecen los alumnos de lo contrario estaría haciendo trampas. Por eso podemos ver que hay una variable dependiente que es la variable efecto que busco observar si hay cambios, y la variable independiente es la variable que controla el investigador, es la que el investigador decide coger mas gente de un grupo u otro o manipular la cantidad de esa variable.

1.4.3. EN EL CONTROL DE VARIABLES EXTRAÑAS

Cuando trabajamos en el ámbito social nos vamos a encontrar muchos más problemas técnicos que en los estudios físicos ya que al trabajar con seres humanos nos encontramos con problemas de carácter moral, problemas de cuantificación de variables y problemas de imparcialidad. Hay muchas y diversas formas de intentar solventar esta problemática, entre estas formas encontraremos el control de las variables extrañas. Porque un problema que se plantea en el ámbito social es el hecho de que existan multitud de variables que pueden intervenir y transformar nuestra variable dependiente. Al controlar al máximo posible el contexto, nos permite minimizar los efectos de otras circunstancias.

1.4.4. EN LA DEFINICIÓN DE LAS VARIABLES

Los estudios que se llevarán a cabo son estudios experimentales o empíricos, donde se busca comprobar si existen relaciones de causa efecto o dependencia entre las variables, como hemos visto al hablar de variable independiente y variable dependiente. Para ello se han de recoger la información pertinente de la muestra a partir de unos instrumentos adecuados. Instrumentos que veremos que son adecuados a partir de los cálculos del tema 8 que son la fiabilidad y la validez. Una vez que poseemos los datos buscaremos tomar decisiones sobre si existe o no relación entre variables para ello se realizarán toda una serie de fórmulas que debremos elegir dependiendo del tipo de variable con la que trabajamos.

Este proceso es especialmente importante cuando trabajamos con variables que no son medibles, sino que se calculan a partir de otras muchas características. En estos casos debemos definir la nueva variable y la conducta operativas y medibles de esta variable. A este proceso se le llama **definición operativa de las variables**. Por ejemplo el cociente intelectual no es algo que exista una regla o una balanza que nos muestre su valor. Sino que para medirlo tenemos que tener en cuenta toda una serie de variables. Un ejemplo claro sería el C.I., la inteligencia de una persona se mide a partir de diferentes características que juntas nos permiten ver la inteligencia de una persona. Cuanto más tengamos que

trabajar con una variable para medirla mas posibilidades de cometer errores que tenemos que tener en cuenta.

1.4.5. EN EL CONTRATE DE HIPÓTESIS O COMPROBACIÓN DE OBJETIVOS

Una vez que se tiene clara la importancia de clasificar a las variables, tenemos que pasar a entender la importancia clave que para un estudio estadístico tiene la **comprobación de los objetivos o el contraste de la hipótesis**. Hay toda una serie de pruebas estadísticas para el contraste de hipótesis, como se verá en el tema 12. Y escogeremos una prueba u otra dependiendo del tipo de variable con la que trabajemos y las características de la muestra con respecto a esa variable en concreto. Por lo tanto el investigador debe tomar la decisión de que cálculo utilizar en cada momento y la interpretación de todos sus resultados.

1.4.6. EN LA DECISIÓN ESTADÍSTICA

Una vez que he escogido una u otra de las pruebas estadísticas posibles tengo que interpretar los resultados que me da para poder concluir si se cumple o se rechaza la hipótesis central y así saber si hay diferencias entre los grupos reales o solo son aparentes.

Aquí ya no queda más remedio que hablar de los tipos de hipótesis que hay. En estadística podemos clasificar a las hipótesis de dos tipos diferentes según como se usen en la investigación. Así hablaremos de **hipótesis de nulidad o hipótesis nula y la hipótesis alternativa o hipótesis del investigador**. La hipótesis nula será el pilar central de mi estudio estadístico porque como dijimos antes lo más importante de la estadística es mantener su estatus de ciencia y por tanto centrarse en una visión imparcial de la existencia o no de relaciones entre variables. Para ser imparcial tengo que tener en cuenta que no puedo pensar que existe una relación si todavía no la he probado por lo tanto la hipótesis nula dice que en principio no hay diferencias estadísticamente significativas entre las variables, y que si vemos alguna diferencia es explicable por el azar. De ahí viene su nombre nula, porque habla de la nula relación entre las variables. Esta hipótesis es el pilar central de la investigación. Y todo gira entorno a si se acepta o no. Cuando no se acepta tenemos que pensar que si que existe relación entre las variables y por lo tanto se aceptará la hipótesis alternativa (la alternativa a la hipótesis nula), que en el fondo coincide con la hipótesis de investigación por eso se le llama hipótesis del investigador.

En esta última etapa de la estadística debemos saber si se cumple o se rechaza la hipótesis nula, aunque claro esta teniendo en cuenta que todos los cálculos que estamos realizando tienen un margen de error porque no trabajamos con todos los valores de la población sino con un grupo reducido de individuos. Por lo tanto podemos confundirnos en las conclusiones porque los individuos de nuestra muestra no se parezcan a los de nuestra población. También podemos cometer otro tipo de error que es el error personal, nosotros podemos confundirnos porque somos personas y por tanto nos podemos confundir.

Un punto que comenta este primer capítulo es el método más habitual de comprobación de si se cumple o no la hipótesis nula. La hipótesis nula se comprueba de varias formas como se verá en el tema 11 y 12 pero la más importante y fiable de estas formas es por medio del estadístico de contraste. Unas fórmulas que nos permiten ver si la diferencia de medias en la población es significativa o no lo es (si existe o no) para saberlo se pone en relación la ecuación que habla de esta diferencia, con el valor ideal que debería de existir para esa situación en concreto, y como podemos saber cual es ese valor ideal. Ese valor ideal nos lo proporcionan las tablas que se encuentran en nuestro formulario. En resumen tenemos que poner en relación el valor que nos proporciona el estadístico de contraste, con el valor ideal de la tabla, por supuesto para que se cumpla la hipótesis nula el valor ideal (llamado teórico o crítico) a de ser más elevado que el valor empírico (el de la fórmula). Cuando ocurre esto podemos decir que se cumple la hipótesis nula.

1.5. LA ESTADÍSTICA Y SU RELACIÓN CON LAS CIENCIAS SOCIALES.

La estadística es un punto técnico básico para las ciencias sociales, ya que es el metodo más eficaz para dar valor científico a estos ambitos del conocimiento sociales. Ya que por si solo no lograría mostrarse de manera científica. Ya que al trabajar con variables humanas es difícil de cuantificarlas.

1.6. POSIBLES LIMITACIONES DE LA ESTADÍSTICA

La estadística es un instrumento de apoyo a la investigación en el ámbito social. Facilitando el trabajo con los datos y su análisis e interpretación de resultados. Aunque claro está hay un problema de fondo que es el control de la variable extraña que son estas variables extrañas, las variables extrañas son variables que si bien no nos interesaba estudiar si que tiene importancia en los resultados obtenidos por la variable dependiente. Son variables que influyen directamente en ella aunque en nuestro estudio no

nos interesa mirar. Y por tanto deberíamos controlarlo para saber si puede transformar todos nuestros resultados.

Otro problema que podemos encontrar otros dos problemas que son el error que se puede llegar a cometer en un estudio porque exista sesgo. Osea que la muestra con la que trabajamos no se parezca a la población y por tanto el contraste que establezca tenga cierto margen de error , al no ser similares la muestra y la poblacion. Otro margen de error que tenemos que tener en cuenta es el error que podemos cometer como humanos, ya que podemos llegar a confundirnos. Este será por tanto otro tanto por ciento de error que puede cometer un estudio estadístico.