

AC 10 105

Hola, buenas noches, bienvenidos al tiempo del curso de acceso directo a la universidad para mayores de 25 años, el CAT. Esta noche en nuestro programa Matemáticas Básicas, vamos a hablar del cálculo de probabilidades y estadística. Con la compañía del profesor Eduardo Ramos veremos las lecciones 10 y 11, cálculo de probabilidades y estadística.

Esta noche en nuestro programa de Matemáticas Básicas, vamos a hablar del cálculo de probabilidades y estadística. Nos acompaña como es habitual el profesor Eduardo Ramos. Muy buenas noches.

Muy buenas noches y un saludo muy cordial a todos nuestros alumnos al tiempo que quiero desearles una palabra de ánimo para esta renta final del curso. Ya se acercan las pruebas presenciales y hay que hacer un esfuerzo final para afrontarlas con éxito. Bien, la ciencia siempre trata de encontrar las leyes que rigen los fenómenos que observamos.

Hay una clase de fenómenos que permiten prever cuál va a ser el efecto producido en otros fenómenos. Sin embargo, se dice que el resultado es consecuencia del azar, pero el azar está sometido a leyes. Esta noche vamos a ver las dos especialidades de la matemática que abordan este tipo de leyes no necesarias.

El cálculo de probabilidades y la estadística, ¿es así verdad? Sí, los comentarios de esta noche se refieren precisamente a estas dos lecciones que forman un bloque que probablemente constituya una de las partes más aplicadas del curso. Todos los alumnos que vayan a estudiar psicología, pedagogía, sociología y disciplinas parecidas encontrarán en estas lecciones de probabilidad y estadística los primeros pasos de algo que luego tendrán que estudiar con más fuerza, con más dedicación. Como muy bien has comentado, la vocación del hombre ha sido siempre descubrir las leyes que gobiernan el universo.

Esas leyes necesarias, como la que observamos desde muy jóvenes, que una piedra se cae cuando es un niño que se divierte tirando objetos al suelo. Pues un objeto se cae si no hay nada que lo sostiene. Esas leyes el hombre encontró, se interesó por ver cómo predecir el comportamiento de esos fenómenos y encontró leyes que de alguna manera formulaban el comportamiento del fenómeno observado y eran capaces de predecir lo que iba a pasar en una repetición del fenómeno futuro.

Sin embargo, hay otro tipo de fenómenos, por ejemplo, el más característico es el de lanzar una moneda al aire, de los cuales no podemos predecir con certeza qué ocurrirá en cada observación del fenómeno. Sabemos que puede salir cara, que puede salir cruz, probablemente incluso que pueda caer de canto la moneda, pero nadie se jugaría la vida a predecir qué es lo que va a ocurrir en una observación del experimento. Esto, el hombre, a mediados del siglo XV, como resultado de observar los juegos de azar precisamente, un jugador fue el que predominantemente se planteó este tipo de problemas, un jugador de

dados, y también intentó averiguar leyes que expliquen cómo se van a comportar estos fenómenos que son gobernados por el azar.

El hecho de que los fenómenos no siempre podamos predecir lo que va a ocurrir, no quiere decir que el hombre no renunció a investigar si era posible encontrar leyes que le pudieran permitir predecir el comportamiento de estos fenómenos. Y esto es lo que hoy se engloba bajo el nombre de cálculo de probabilidades. ¿Pero cuál es el concepto de probabilidad? Sí, el concepto de probabilidad surge de una observación empírica también.

Si lanzamos, ponemos otra vez el mismo ejemplo, si lanzamos una moneda al aire, no sabemos, no podemos predecir con certeza si en el lanzamiento va a salir cara o va a salir cruz. Pero, sin embargo, sí sabemos esta otra cosa. Si lanzamos indefinidamente la moneda, una serie larga de veces la moneda, por ejemplo mil veces, lo que sí podemos afirmar con bastante certeza, podemos tener cierta confianza, es que el número de caras, por decir algo, va a estar comprendido entre 400 y 600.

Es decir, aproximadamente a la mitad de las veces saldrán caras, aproximadamente a la mitad de las veces saldrán cruces, y podríamos tranquilamente jugarnos, hombre, no la vida, pero sí mil pesetillas, a que el número de caras estarán entre 400 y 600. Casi todos estaríamos dispuestos a ese juego, y con la tranquilidad de que para una vez que íbamos a fallar íbamos a acertar un buen número de veces. Entonces, esta estabilidad del número de veces que se presenta cada uno de los resultados que puede andarse a observar en un experimento aleatorio, esta estabilidad en largas series, en largas repeticiones del experimento, es una cierta ley, en definitiva, es una ley de estabilidad de las frecuencias que nos induce a considerar ese número, es decir, el número aproximadamente 0,5, la mitad de las veces que tiramos una moneda saldrá aproximadamente caras, nos induce a considerarlo como una medida del grado de ocurrencia de esa posibilidad.

Y ese grado de ocurrencia son las ideas intuitivas de lo que luego se formaliza y se define con precisión como probabilidad. ¿Y cuál es el modelo matemático de los sucesos aleatorios? Después de estas ideas intuitivas o de estas reflexiones más o menos de discurso, de palabras, entran las matemáticas con toda su artillería de siempre a tratar de formalizar las cosas, a tratar de precisar la terminología y a tratar de decir lo que es cada uno de los conceptos que estamos manejando. Una vez más, las matemáticas nos sirven para ese rigor, esa disciplina intelectual que ya vamos comentando en todo el curso, que es uno de los objetivos principales de este curso.

Entonces, para construir el modelo matemático de los sucesos aleatorios, el concepto clave es el de espacio muestral, el espacio muestral o espacio de posibilidades. Esto está formado por todos los resultados posibles que pueden darse, o sea, es el conjunto formado por todos los resultados posibles que pueden darse al efectuar un experimento aleatorio. Y luego, una vez definido o considerado el espacio muestral que suele designarse, como lo ponemos en el texto con la letra griega omega, pues pueden definirse sucesos como conjuntos, su conjunto del

espacio muestral, y pueden considerarse algunas operaciones con estos sucesos, cuáles son la unión, la intersección, etc.

Algo parecido a lo que se puede hacer en la teoría de conjuntos. También aprovecho para decir que todo esto no va a costar mucho de estudiar porque ya lo sabemos, es muy similar a lo que se ha estudiado en teoría de conjuntos. Y una vez identificada la familia de sucesos correspondientes a un determinado fenómeno aleatorio, ¿cuál sería el modelo matemático de la probabilidad? La segunda parte, el segundo paso que dan las matemáticas para describir mediante un modelo el estudio de esos experimentos aleatorios es la consideración de la probabilidad.

La probabilidad, tratando de formalizar esas ideas intuitivas que decíamos antes, no es más que un número que nos mide el grado de ocurrencia de un determinado suceso. Es decir, qué valora, qué mide, con qué grado de certeza podemos tener de que va a ocurrir un suceso. Entonces, una probabilidad es un número, es una aplicación, para hablar en términos matemáticos, que asigna a cada suceso un número real entre 0 y 1. Claro, no sirve cualquier aplicación que asigne a cada suceso un número real entre 0 y 1, sino que a esa aplicación le debemos exigir algunas condiciones de coherencia, algunas condiciones de regularidad, por decirlo así.

Y son esencialmente dos las que se consideran como definición de probabilidad. Una es que si dos sucesos no tienen A y B , por decir, no tienen ningún resultado común, o sea, son incompatibles en términos matemáticos, si su intersección es vacía, entonces la probabilidad de la unión de sucesos es la suma de las probabilidades. Es decir, la probabilidad es una aplicación que traspasa las uniones y las convierte en sumas.

Y la otra condición de regularidad, para que una aplicación definida en un subconjunto del espacio muestral entre 0 y 1 sea una probabilidad, es una condición de normalidad. Es decir, la probabilidad de omega, del suceso seguro, por decirlo así, de lo que sabemos que va a pasar con toda seguridad, es igual a 1, es la cota superior. Bueno, entonces, este par formado por el espacio muestral omega y la aplicación P , la aplicación probabilidad, constituye lo que en matemáticas, en materia de probabilidades, se llama el espacio de probabilidad y es la descripción matemática completa de cualquier experimento aleatorio.

La probabilidad también tiene propiedades, ¿cuáles son estas? Sí, bueno, con la probabilidad se pueden, una vez formalizada así, con esta definición que hemos dado, de la probabilidad de una aplicación del conjunto de sucesos en el intervalo 0-1, con las dos condiciones de que la probabilidad de la unión de sucesos disjuntos sea la suma de las probabilidades y que la probabilidad del suceso seguro sea todo 1, entonces, con esa definición se pueden considerar una serie de probabilidades que no voy a repetir aquí, me voy a remitir a la sesión 15 del texto y recordar a los estudiantes que deben trabajarlas, comprenderlas y ver cómo en el fondo son todas bastante intuitivas y bastante sencillas. ¿Y los problemas de asignar probabilidades, cuáles son? Sí, esto es uno de los, siempre de las lagunas, es que se leen de alguna manera las

cosas que quedan en el aire cuando se hace el modelo matemático para trabajar o para explicar los experimentos aleatorios, es cómo asignar probabilidades. En esto, de hecho, los problemas matemáticos empiezan en el momento en que tenemos ya una función que nos da la fredad de los sucesos, ahí es cuando ya podemos trabajar y aplicar, por ejemplo, esas condiciones, esos resultados, esas consecuencias de la definición de probabilidades que acabamos de señalar.

No hay reglas fijas para asignar probabilidades, depende un poco del interés del observador, del experimento que estemos considerando, de la experiencia que se tenga, etc., del sentido común, en definitiva, y seguramente con cualquier modelo o cualquier problema o técnica de asignación de probabilidades debe estar sustantamente revisándose según se vaya conociendo y ganando experiencia y ganando conocimiento del experimento aleatorio. Puedo poner un ejemplo de lo que quiero decir. Si nos dan una moneda y no tenemos mayor información, la intuición nos dice que deberíamos asignar probabilidad igual al hecho de salir cara y al hecho de salir cruz, y cero al hecho de salir de canto.

Eso lo podríamos simbolizar con probabilidad de cara igual a un medio, probabilidad de cruz igual a un medio y probabilidad de salir de canto igual a cero. Si después de tirar, por ejemplo, mil veces la moneda, observamos que han salido doscientas caras y ochocientas cruces, probablemente lo que nos conduzca a la intuición es a revisar nuestra asignación de probabilidades y deducir que la moneda está cargada, y por lo tanto esa asignación de probabilidades que hemos hecho de un medio a un medio no nos parezca correcta para describir el comportamiento de esa moneda y tendríamos que revisarla. Claro que ahí siempre tendremos alguna decisión que tiene cierto riesgo, porque el resultado de doscientas caras y ochocientas cruces bien pudo haberse dado con una moneda equilibrada, pero en realidad eso nunca lo sabremos con certeza, y es la negociación entre el que hace el modelo y la contrastación del modelo luego con la realidad.

En cualquier caso, este mismo ejemplo me sirve para comentar una regla bastante intuitiva, que es la regla de Laplace o la regla de la equiprobabilidad, que es cuando la probabilidad de un supuesto que todos los casos que pueden darse al efectuar un experimento aleatorio son igualmente probables, la probabilidad de un suceso es el número de casos favorables dividido por el número de casos posibles. Con esta regla de equiprobabilidad o regla de Laplace podemos empezar a trabajar para iniciar muchos modelos, muchos casos y muchas aplicaciones. Bien, pues después de ver este capítulo 10, el cálculo de probabilidades, vamos a continuar con el capítulo 11, la estadística.

Desde luego, entre todas las ciencias matemáticas, la estadística es la más cercana a las preocupaciones y problemas del hombre de hoy. ¿Podríamos hacer un poco de historia? Etimológicamente, estadística significa ciencia del Estado. Ya hace siglos el Estado se interesó en averiguar, en saber cuántos súbditos tenía, cuántas tierras, cuántos hombres en edad militar.

Las razones sin duda eran muy prosaicas, eran de esta índole, eran sujetos que iban a pagar tributos. Entonces, esto para representar los Estados producía tablas de números, tablas de números con los habitantes, con la edad que tenía cada uno de ellos, con las posiciones que tenía cada uno de ellos. Y estas tablas de números que representaban o que eran las cosas que le interesaban al Estado, pues cuajaron en la palabra estadística.

Ese es, sin duda, el étimo. También el hombre, como comentábamos antes, que se preocupaba o se ha preocupado siempre en toda la historia de encontrar leyes que expliquen los fenómenos que le rodean, que expliquen la naturaleza. Durante los siglos XVII y XVIII, fundamentalmente, la ambición del hombre, el interés del hombre era dominar el mundo que le rodeaba, el hacerlo más habitable.

Era un hombre constructor, era un hombre ingeniero. Poco a poco, con el desarrollo de la sociedad y el desarrollo del pensamiento, el hombre giró un poco en sus intereses y empezó a preocuparse de lo psicológico, de lo económico, de lo social. El hombre se interesó por aspectos que contemplaban experimentos del tipo que comentábamos antes, aleatorios.

De experimentos que no estaban gobernados por las leyes necesarias, sino por las leyes de la clara. Y en este ámbito se hace imprescindible la utilización de la herramienta matemática para estudiar y contemplar los fenómenos aleatorios, que es el cálculo de probabilidades, y la herramienta para el estudio y la manipulación de esas masas de datos, de números que representan los intereses del hombre en estos ámbitos. Por eso, en los siglos XIX y, fundamentalmente, a lo largo de este último siglo, siglo XX, el desarrollo de la estadística ha sido muy importante.

Y quizás podemos decir, a lo mejor exagerando un poco, que el pensamiento del hombre actual es un pensamiento estadístico. Nada se puede afirmar en el ámbito de lo social, en el ámbito de lo económico, que no esté sustentado por una estadística. Sí, hasta tal punto que hay conceptos que están en la calle y todos hablamos de ellos, aunque quizá no los conozcamos matemáticamente.

Por ejemplo, el de población y muestra. Sí, la población es, digamos, el núcleo, el corazón de lo que quiere estudiar la estadística. Se llama población al conjunto de seres o de objetos o de individuos sobre los que se quiere obtener información.

Muestra es una parte de esa población, parte, un subconjunto, por decirlo así, de una población que es la que se examina para obtener los datos. Y lo que se mide, lo que se observa en las unidades de la población, en los individuos de la población, son variables estadísticas, es decir, características que se pueden medir o que se pueden observar. Estos son tres conceptos, población, muestra y variable estadística, alrededor de los cuales gira toda la parte de estadística descriptiva que estudiamos en esta lección.

Quizá sería conveniente saber cuáles son las clases de variable estadística. Sí, los datos en estadística proceden de observar una magnitud a cada uno de los elementos de la población.

Como hemos dicho, estas magnitudes que observamos en los elementos de la población se llaman variables estadísticas, que pueden tomar diferentes valores o diferentes modalidades, también se dice así.

Voy a poner un ejemplo muy sencillito. Si observamos la variable estadística sexo en una población, puede tomar dos valores, masculino o femenino. Si observamos la variable número de hijos de un matrimonio, puede tomar valores como 0, 1, 2, etc.

Las variables, estas variables estadísticas, pueden clasificarse atendiendo a su carácter. Podemos hablar de variables cualitativas, cuando los valores que pueden tomar son cualidades de los individuos, es decir, aspectos que no pueden ser ordenados ni operados, es decir, no se puede sumar ni multiplicar. El ejemplo más característico es el que acabo de decir.

La variable sexo es una variable que es de tipo cualitativo. Los valores masculino y femenino no tienen sentido de que sean ni uno mejor que el otro, ni mayor ni menor, ni se pueden sumar. Otro tipo de variables que añaden un poco más son las variables ordinales.

Son aquellos cuyos valores, sin que se pueda operar algebraicamente con ellos, sí son susceptibles de una cierta ordenación. Son variables del tipo cuando las posibles respuestas o valores que pueden tomar son malo, regular, bueno, óptimo, de ese estilo, es decir, sabemos que malo es peor que regular y es peor que bueno, pero no sabemos cómo se podría sumar bueno y malo. Un paso más en la clasificación, otro tipo de variables son los que ya se llaman cuantitativas, cuando los valores que pueden tomar son susceptibles de ser operados algebricamente.

Es algo que es numerable, que el resultado es un número. Por ejemplo, esa que dijimos antes, el número de hijos de una familia, 0, 1, 2 y 3, pues podemos sumar esos valores. 3 más 2 es igual a 5. Y dentro de este tipo de variables cuantitativas podemos todavía distinguir dos categorías.

Las variables que son discretas, es decir, aquellas que pueden tomar números, pueden tomar una serie de valores determinados, pero no los valores intermedios de ese tipo, 0, 1, 2, como le acabo de decir, y variables que son continuas, de las cuales, a priori, podemos pensar que pueden tomar cualquier valor real. Por ejemplo, una variable que midiese la altura de una persona, expresada, por ejemplo, en centímetros, podría tomar cualquier valor en centímetros. También si se expresa en milímetros podría tomar cualquier valor en milímetros.

Si se expresa en décimas de milímetros podría tomar cualquier valor en décimas de milímetros. Estas variables, repito, se llaman continuas. Y esta es una clasificación suficientemente válida y amplia para los objetivos de este curso.

¿Y el concepto de frecuencia? Sí, ahí empezamos ya a pensar cómo trabaja la estadística. Los datos que tiene son masas de datos brutos, es decir, el resultado de cada observación a cada individuo. Y nos podemos empezar a imaginar cómo es la fuente bruta de datos de un

problema de estadística.

Son cientos y cientos de números. Lo que pretende a continuación el estadístico es extraer información de esos números, resumir esos números. El poderles hacer algo luego con ellos que les proporcione un conocimiento de la población que no tenía.

Y el primer paso es contar. Contar cuántas veces se produce determinado valor de una variable. Y esa es la idea de frecuencia.

Que se puede contar en términos absolutos, es decir, el número de veces que aparece en toda la tabla de datos, o bien en términos relativos. El número de veces que aparece el valor de una variable, pero dividido por el número total de observaciones que tenemos. Eso sería la frecuencia relativa.

Digamos que eso es el primer paso que se nos puede ocurrir para resumir esa basa de datos bruta, que es el objeto fuente, el primer resultado de las observaciones que hace la estadística. ¿Se puede representar gráficamente las frecuencias? A continuación, después de resumir en unas tablas de frecuencias, después de haber contado cuántas veces sale cada variable, cada valor de la variable, perdón, entonces una siguiente técnica para ganar información sobre esa población es hacer una representación gráfica de esa tabla de frecuencias. Y entonces ideas sobre eso hay muchas.

Las más características, que son bastante familiares para el hombre de la calle porque aparecen todos los días en el periódico, son los llamados diagramas de sectores y diagramas de barras, que no es momento ahora en este programa de explicar detalladamente cómo se construyen, pero sí puedo remitirme al texto en donde está muy detallado cómo hay que hacer eso, cómo es la lógica que tienen estos diagramas, cómo se construyen, aunque debo decir también que hoy en día efectos plásticos, todo este tipo de gráficos, diagramas, etc., se hacen de una manera muy sencilla con los ordenadores. Pero en estos ejemplos de representación gráfica de las frecuencias no hay ningún ejemplo de variable continua. Sí, las técnicas variables continuas, en el texto también tenemos cómo se puede hacer la representación tanto para variables discretas como para variables continuas.

Las ideas son muy similares. Y lo que me importaría aquí destacar es que es el primer paso para ganar información sobre la población, recordando el viejo afroismo de que una imagen vale más, en este caso, que una tabla de números. ¿Y la media aritmética qué es? Otro tercer paso.

Estábamos hablando de resúmenes. Estoy repitiendo siempre la misma idea. Resumir la tabla bruta de datos.

Primero lo hacemos con ambas frecuencias, hacemos una técnica de representación y luego hacemos otro resumen, que son resúmenes numéricos. La media aritmética no es, ni más ni menos, que el primer intento de reducir la tabla de números que proporciona la masa bruta de

datos a un solo número que lo representa a todos ellos. Y hay otras ideas, además de la media.

La media, como probablemente, una vez que se estudie el texto, pero es algo también bastante popular, la idea de media es simplemente sumar los valores que se tienen y dividir por el número de ellos. Es una idea bastante intuitiva. Eso es una idea, ya digo que no es la única, como hay otras, la moda, la mediana, etc.

Pero es una idea para intentar concentrar en un solo número toda esa masa de datos, que es el primer objeto, el objeto fundamental, o sea, lo que nos produce principalmente la estadística. Hasta aquí solo se ha considerado una variable. Sin embargo, una de las aplicaciones más importantes de la estadística es el análisis de las relaciones que pueden existir entre una variable respecto de otras o varias entre sí.

Sí, la estadística de una variable lógicamente es lo más elemental y por eso hay que empezar a estudiar esa parte, pero desde luego en cualquier aplicación real observaremos varias variables estadísticas a un mismo individuo en la población. Por ejemplo, podemos observar la talla, el peso, el color de los ojos para un individuo determinado, el número de hijos que tiene. Y es cuando realmente cobra su interés la estadística, el poder describir una población mediante un conjunto de variables.

Y como es fácil suponer, el número de posibilidades que se abren es muy grande, es decir, ahora ahí aparecen nuevos problemas, como pueden ser las relaciones que hay o las dependencias o las asociaciones que hay entre los valores de una variable y los valores de la otra variable. En esta parte, simplemente en este curso y en esta parte de la lección nos tenemos que limitar a comentar brevemente esos problemas de la asociación que hablamos y es lo que debe estudiar si comprende bien todo lo que puede ser la dependencia funcional o la independencia entre dos variables o la dependencia estadística que es un grado intermedio entre la dependencia funcional y la independencia. Por poner un ejemplo muy sencillo, si consideramos dado una velocidad de un coche por ejemplo, una variable que mide el tiempo y otra variable que mide el espacio recorrido, la física nos enseña que para un valor del tiempo le equivale a un espacio recorrido, si vamos a 100 km por hora, en una hora andamos 100 km, en dos horas 200 es decir, dada la variable tiempo su valor, sabemos cual es la variable espacio eso es el caso de dependencia funcional también habrá un caso en el que lo que nos dice una variable no nos informa nada de lo que puede pasar a otra, como puede ser por ejemplo si consideramos en una persona el número de hijos y el número del zapato que gasta el que tenga dos hijos no nos dice nada que si lo sigue hasta el 40 o el 41, si tiene tres hijos pues puede gastar el 38 o el 39 es decir, la información que nos da la variable número de hijos no gana nada de lo que sabemos sobre la otra variable, el número del zapato, sería un caso de independencia estadística pero entre estos dos extremos de la dependencia funcional y la independencia y la independencia estadística podemos considerar otro tipo de asociación, otro tipo de relación entre variables como puede ser que es la que se denomina dependencia estadística el ejemplo más típico que se nos puede ocurrir es el de la talla y el peso, es decir, si nos dicen que una persona pesa 60 kilos, pues ya nos dan cierta información sobre su talla, lo

[illegible]