



Mi Universidad

Nombre del Alumno: Andri Guadalupe Pérez Aguilar

Nombre del tema: Estadística descriptiva y Cálculo de probabilidades

Parcial: 1°

Nombre de la Materia: BIOESTADISTICA I

Nombre del profesor: Rosario Gómez Lujano

Nombre de la Licenciatura: Enfermería

Cuatrimestre: 4°

Estadística descriptiva

1.1 La estadística en enfermería

En las ciencias de la salud, la estadística tiene una gran importancia ya que posee numerosas ventajas. La principal ventaja del pensamiento estadístico interviniendo en las ciencias de la vida es que no solo resuelve, sino que también comprende una compleja metodología para dar respuesta a las hipótesis. En Salud Pública la estadística permite analizar situaciones en las que los *componentes aleatorios* contribuyen de forma importante en la variabilidad de los datos obtenidos.

1.1.1 Introducción histórica

El primer médico que utilizó métodos matemáticos para cuantificar variables de pacientes y sus enfermedades fue el francés Pierre Charles-Alexandre Louis. Los primeros intentos de hacer coincidir las matemáticas de la teoría estadística con los conceptos emergentes de la infección bacteriana tuvieron lugar a comienzos del siglo XX. La enfermera inglesa Florence Nightingale. Durante la guerra de Crimea, Florence Nightingale observó que eran mucho más numerosas las bajas producidas en el hospital que en el frente.

1.2 La estadística como herramienta de trabajo en enfermería

El análisis y las técnicas estadísticas son un componente esencial en toda investigación biomédica, y la utilización de las técnicas estadísticas ha evolucionado considerablemente en los últimos años en las áreas de la investigación de ciencias de la salud. Esta disciplina es usada en diversos campos de la medicina y la salud pública, como la epidemiología, nutrición y salud ambiental. Asimismo, sus métodos son aplicados en estudios relacionados con la ecología y la genómica. La variable sería el color de ojos y sería una variable cualitativa. Es decir, no se expresa con número.

1.3 Descripción de una variable estadística

Cuando hablamos de variable estadística estamos hablando de una cualidad que, generalmente adopta forma numérica. Claro que no todas las variables estadísticas son iguales y, por supuesto, no todas se pueden (en principio) expresar en forma de número. Así, otra variable que podríamos encontrarnos es el color de ojos de una persona. La variable sería el color de ojos y sería una variable cualitativa. Es decir, no se expresa con número.

1.3.1. Definiciones básicas

Variable estadística: Una variable estadística es una característica de una muestra o población de datos que puede adoptar diferentes valores.

Variable cuantitativa: Son variables que se expresan numéricamente.

□ **Variable continua:** Toman un valor infinito de valores entre un intervalo de datos.

Variable cualitativa: Son variables que se expresan, por norma general, en palabras.

□ **Variable ordinal:** Expresa diferentes niveles y orden.

1.4 Representaciones gráficas

Una gráfica o una representación gráfica o un gráfico, es un tipo de representación de datos, generalmente cuantitativos, mediante recursos visuales (líneas, vectores, superficies o símbolos), para que se manifieste visualmente la relación matemática o correlación estadística que guardan entre sí.

Tipos de representación grafica

Diagramas de barras: muestran los valores de las frecuencias absolutas sobre un sistema de ejes cartesianos, cuando la variable es discreta o cualitativa.

Histogramas: formas especiales de diagramas de barras para distribuciones cuantitativas continuas.

Polígonos de frecuencias: formados por líneas poligonales abiertas sobre un sistema de ejes cartesianos.

Pictogramas: o representaciones visuales figurativas.

Cartogramas: expresiones gráficas a modo de mapa.

1.5 Representación numérica

En la primera columna se ordenan de menor a mayor los diferentes valores que tiene la variable en el conjunto de datos.

2. En las siguientes columnas (segunda y tercera) se ponen las frecuencias absolutas y las frecuencias absolutas acumuladas.

Frecuencia Absoluta

La frecuencia absoluta (n_i) de un valor X_i es el número de veces que el valor está en el conjunto (X_1, X_2, \dots, X_N). La suma de las frecuencias absolutas de todos los elementos diferentes del conjunto debe ser el número total de sujetos N .

Frecuencia Relativa

La frecuencia relativa (f_i) de un valor X_i es la proporción de valores iguales a X_i en el conjunto de datos (X_1, X_2, \dots, X_N).

1.6 Características de posición, dispersión y forma.

Las medidas de posición son indicadores estadísticos que permiten resumir los datos en uno solo, o dividir su distribución en intervalos del mismo tamaño.

El cuartil: Es uno de los más utilizados y divide la distribución en cuatro partes iguales. Así, existen tres cuartiles.

El quintil: En este caso, divide la distribución en cinco partes.

La media aritmética, geométrica

oarmónica: Son tres medidas centrales que nos indican un promedio ponderado de los datos.

La mediana: En este caso, esta es la medida de posición central más reconocible. Divide la distribución en dos partes iguales.

La moda: Estamos ante una medida central de los valores más frecuentes. Por tanto, la moda nos informa sobre aquellos que se repiten en más ocasiones.

1.7 Descripción numérica de una variable estadística bidimensional

En numerosas ocasiones interesa estudiar simultáneamente dos (o más) caracteres de una población. Variable estadística bidimensional es el conjunto de pares de valores de dos caracteres o variables estadísticas unidimensionales X e Y sobre una misma población. El término variable marginal se usa para referirse a una variable del subconjunto de retenido y cuyos valores pueden ser conocidos.

1.8 Distribuciones marginales y condicionadas.

En teoría de probabilidades, la distribución marginal es la distribución de probabilidad de un subconjunto de variables aleatorias de un conjunto de variables aleatorias. Para una variable aleatoria se puede especificar probabilidades para dicha variable sin tener en cuenta los valores de cuales quiera otras variables aleatorias. separado se llama "distribución marginal".

1.9 Independencia e incorrelación

Dos variables estadísticas son estadísticamente independientes cuando el comportamiento estadístico de una de ellas no se ve afectado por los valores que toma la otra. Se dice que dos variables X e Y son independientes estadísticamente cuando la frecuencia relativa conjunta es igual al producto de las frecuencias relativas marginales.

Ejemplo

El suceso estatura de los alumnos de una clase y el color del pelo son independientes: el que un alumno sea más o menos alto no va a influir en el color de su cabello, ni viceversa.

Característica numérica

Los sistemas de numeración son conjuntos de dígitos usados para representar cantidades, así se tienen los sistemas de numeración decimal, binario, octal, hexadecimal, romano, etc.
 b = valor de la base del sistema
 n = número del dígito o posición del mismo

1.10 Regresión y correlación

En forma más específica el análisis de correlación y regresión comprende el análisis de los datos muestrales para saber qué es y cómo se relacionan entre si dos o más variables en una población. El análisis de correlación produce un número que resume el grado de la correlación entre dos variables; y el análisis de regresión da lugar a una ecuación matemática que describe dicha relación.

Definiciones

En estadística, el análisis de la regresión es un proceso estadístico para estimar las relaciones entre variables. Si los valores de una variable se modifican de manera sistemática con respecto a los valores de otra, se dice que ambas variables se encuentran correlacionadas.

Curva de regresión y coeficiente de determinación

La curva de regresión de Y sobre X visualiza como cambia la media de la variable Y de aquellos grupos de observaciones caracterizados por tener un mismo valor en la otra variable X. Es decir, como varía, por término medio, la variable Y en función de los valores de X.

Regresión Lineal

La regresión lineal simple comprende el intento de desarrollar una línea recta o ecuación matemática lineal que describe la reacción entre dos variables.

1.11 Otros tipos de regresión

Regresión Múltiple: Este tipo se presenta cuando dos o más variables independientes influyen sobre una variable dependiente. Ejemplo: $Y = f(x, w, z)$.

Ejemplo de aplicación

En la Facultad de Ingeniería de Sistemas se quiere entender los factores de aprendizaje de los alumnos que cursan la asignatura de PHP, para lo cual se escoge al azar una muestra de 15 alumnos y ellos registran notas promedias en las asignaturas de Algoritmos.

El Error Estándar de Regresión Múltiple

Mediante esta medida de dispersión se hace más preciso el grado de dispersión alrededor del plano de regresión, se hace más pequeño. En los resultados de Excel se llama error típico y para explicar la relación del aprendizaje de PHP que se viene desarrollando es de 0.861

Conclusión

El 69.70% del aprendizaje del Curso de PHP puede ser explicado mediante las notas obtenidas por las asignaturas de Algoritmos, Base de Datos y Programación.

1.12 Análisis de atributos

Su principal objetivo es el de evitar un error muy común consistente en tratar de encontrar la forma de mejorar un producto, servicio o proceso analizándolo como un todo. Muchas veces, la búsqueda de una idea global, salvadora, que mejore el todo, impide descubrir la característica específica que, por sí sola, podría producir el resultado deseado.

Características para las Gráficas de Control de Atributos

Están basadas en decisiones de pasa/no pasa.

- Se pueden aplicar en casi cualquier operación donde se recolectan datos.
- Se utilizan en características de calidad que no pueden ser medidas o que son costosas o difíciles de medir. A diferencia de las gráficas de control de datos variables, las gráficas de datos atributos se pueden establecer para una característica de calidad o para muchas.

Tipos de Gráficas de Atributos:

- Defectivos – np - número de unidades no-conformes
- p - proporción de unidades no-conformes
- Defectos – c - número de defectos

Unidad 2

Cálculo de Probabilidades

En la vida cotidiana aparecen muchas situaciones en las que los resultados observados son diferentes, aunque las condiciones iniciales en las que se produce la experiencia sean las mismas. El objetivo del Cálculo de Probabilidades es el estudio de métodos de análisis del comportamiento de fenómenos aleatorios.

2.1 La medida de probabilidad. Espacio Probabilístico

Para medir la incertidumbre existente en un experimento aleatorio, se parte de un espacio muestra M en el que se incluyen todos los posibles resultados individuales del experimento.

Axioma 1

Un experimento se denomina aleatorio cuando puede dar resultados distintos al realizarse en las mismas condiciones.

Axioma 2

Es una fórmula de agregación que se usa para calcular la probabilidad de la unión de subconjuntos disjuntos.

2.2 Probabilidad condicionada

Miraremos la forma en que cambia la probabilidad de un suceso cuando se sabe que otro suceso ha ocurrido.

Ejemplo

indica que la probabilidad de que un estudiante juegue al baloncesto dado que también juega al fútbol es del 43,75%.

Suceso independiente

Dos sucesos, A y B , son independientes cuando la probabilidad de que suceda no se ve afectada porque haya sucedido, o no.

Suceso Dependiente

En la parte del numerador tenemos la probabilidad condicionada, y en la parte de abajo la probabilidad total. En cualquier caso, aunque la fórmula parezca un poco abstracta, es muy sencilla.

2.3 Teoremas asociados

El teorema de Bayes es utilizado para calcular la probabilidad de un suceso, teniendo información de antemano sobre ese suceso. El teorema de Bayes entiende la probabilidad de forma inversa al teorema de la probabilidad total.

Fórmula del teorema de Bayes

Para calcular la probabilidad tal como la definió Bayes en este tipo de sucesos, necesitamos una fórmula. Donde B es el suceso sobre el que tenemos información previa y $A(n)$ son los distintos sucesos condicionados. Sabiendo que un envase es defectuoso, la probabilidad de que haya sido producido por la máquina A es del 25%, de que haya sido producido por la máquina B es del 28% y de que haya sido producido por la máquina C es del 47%.

2.4 Variable aleatoria

Se llama variable aleatoria a toda función que asocia a cada elemento del espacio muestra E un número real.

Tipos de variable aleatoria

Variable aleatoria discreta: Una variable aleatoria es discreta si los números a los que da lugar son números enteros.

Variable aleatoria continua: Una variable aleatoria es continua en caso de que los números a los que dé lugar no sean números enteros.

Ejemplo de variable aleatoria

- . **Dado:** No es la variable aleatoria. El dado es simplemente un objeto.
- Lanzamiento de un dado:** No es la variable aleatoria. El lanzamiento de un dado es el experimento aleatorio.
- Resultados del lanzamiento de un dado:** Sí es la variable aleatoria. Es la función que recoge los resultados del lanzamiento del dado.

2.5 Concepto de variable

Una variable es un símbolo que actúa en las funciones, las fórmulas, los algoritmos y las proposiciones de las matemáticas y la estadística. Según sus características, las variables se clasifican de distinto modo.

variable aleatoria

Se denomina variable aleatoria (o estocástica) a la función que adjudica eventos posibles a números reales (cifras), cuyos valores se miden en experimentos de tila probabilidad obtenida de esta manera se denomina probabilidad inducida. valores inferiores o iguales a x es igual a la probabilidad del suceso formado por el conjunto de resultados elementales sobre los que el valor de la variable es menor o igual que x . Se puede comprobar que, a partir de la condición requerida, se pueden obtener probabilidades sobre cualquier tipo de intervalo de la recta real. por aleatorio.

2.6 Función de distribución

En la teoría de la probabilidad y en estadística, la Función de Distribución Acumulada (FDA, designada también a veces simplemente como FD) o función de probabilidad acumulada asociada a una variable aleatoria real: X (mayúscula) sujeta a cierta ley de distribución de probabilidad, es una función matemática de la variable real: x (minúscula); que describe la probabilidad de que X tenga un valor menor o igual que x .

Intuitivamente, asumiendo la función f como la ley de distribución de probabilidad, la FDA sería la función con la recta real como dominio, con imagen del área hasta aquí de la función f , siendo aquí el valor x para la variable aleatoria real X .

La FDA asocia a cada valor x , la probabilidad del evento: "la variable X toma valores menores o iguales a x ". El concepto de FDA puede generalizarse para modelar variables aleatorias multivalentes.

2.7 Variables aleatorias discretas y continuas

Una variable aleatoria es una función que asigna un valor numérico, al resultado de un experimento aleatorio. Una variable aleatoria puede ser discreta o continua.

Las variables aleatorias discretas son aquellas que presentan un número contable de valores; por ejemplo, el número de personas que viven en una casa (3, 5 o 9).

□ **Las variables aleatorias continuas** son aquellas que presentan un número incontable de valores.

variable aleatoria

Una variable aleatoria es una función que asigna un valor numérico, al resultado de un experimento aleatorio.

Ejemplo

Tenemos una moneda que en sus caras tiene por un lado un gato y por el otro, un perro.

variable aleatoria discreta

Una variable aleatoria discreta es aquella que puede asumir un número contable de valores.

variable aleatoria continua

Es aquella que puede asumir un número incontable de valores.

2.8 Características de una variable

Están contenidas esencialmente en el título, el problema, el objetivo y las respectivas hipótesis de la investigación. En virtud de ello es que no se puede agregar nuevas variables de las que ya existen en los ítems mencionados.

Son aspectos que cambian o adoptan distintos valores. Esto significa que las variables al ser medidas y observadas expresan diferencias entre los rasgos, cualidades y atributos de las unidades de análisis.

Son enunciados que expresan rasgos característicos de los problemas medibles empíricamente. Estas variables en la práctica social pueden ser medidas y observadas con instrumentos convencionales, en mérito de que contienen rasgos, propiedades y cualidades.

Son susceptibles de descomposición empírica. Dicho de otro término, que las variables pueden desagregarse en indicadores, índices, subíndices e ítems.

2.9 Esperanza de una variable aleatoria

En estadística la esperanza matemática (también llamada esperanza, valor esperado, media poblacional o media) de una variable aleatoria, es el número que formaliza la idea de valor medio de un fenómeno aleatorio. La esperanza matemática de una variable aleatoria es una característica numérica que proporciona una idea de la localización de la variable aleatoria sobre la recta real.

La varianza de una variable, si existe, es el valor medio de las dispersiones cuadráticas de los valores de la variable respecto de su media. Por este motivo, tanto la varianza como su raíz cuadrada, σX , que se denomina desviación típica, se usan, como se verá posteriormente, como medidas de la dispersión de la variable.

2.10 Momentos de una variable aleatoria

Cuando la distribución de probabilidad de una variable aleatoria no es conocida, diversas características de ella pueden proporcionar una descripción general de la misma.

Momentos no centrados

□ **Momentos centrados en media**

Los momentos centrados se calculan, como los no centrados, teniendo en cuenta la definición de esperanza de una función de una variable aleatoria.

La varianza de una variable, si existe, es el valor medio de las dispersiones cuadráticas de los valores de la variable respecto de su media. Por este motivo, tanto la varianza como su raíz cuadrada, σX , que se denomina desviación típica, se usan, como se verá posteriormente, como medidas de la dispersión de la variable.

2.11 Funciones asociadas a una variable aleatoria

Una función que asocia un número real, perfectamente definido, a cada punto muestra. A veces las variables aleatorias (v.a.) están ya implícitas en los puntos muestrales.

Las funciones de densidad discreta y continua tienen, por tanto, un significado análogo, ambas son las funciones que acumuladas (en forma de sumatorio en el caso discreto o en forma de integral en el caso continuo) dan como resultado la función de distribución.

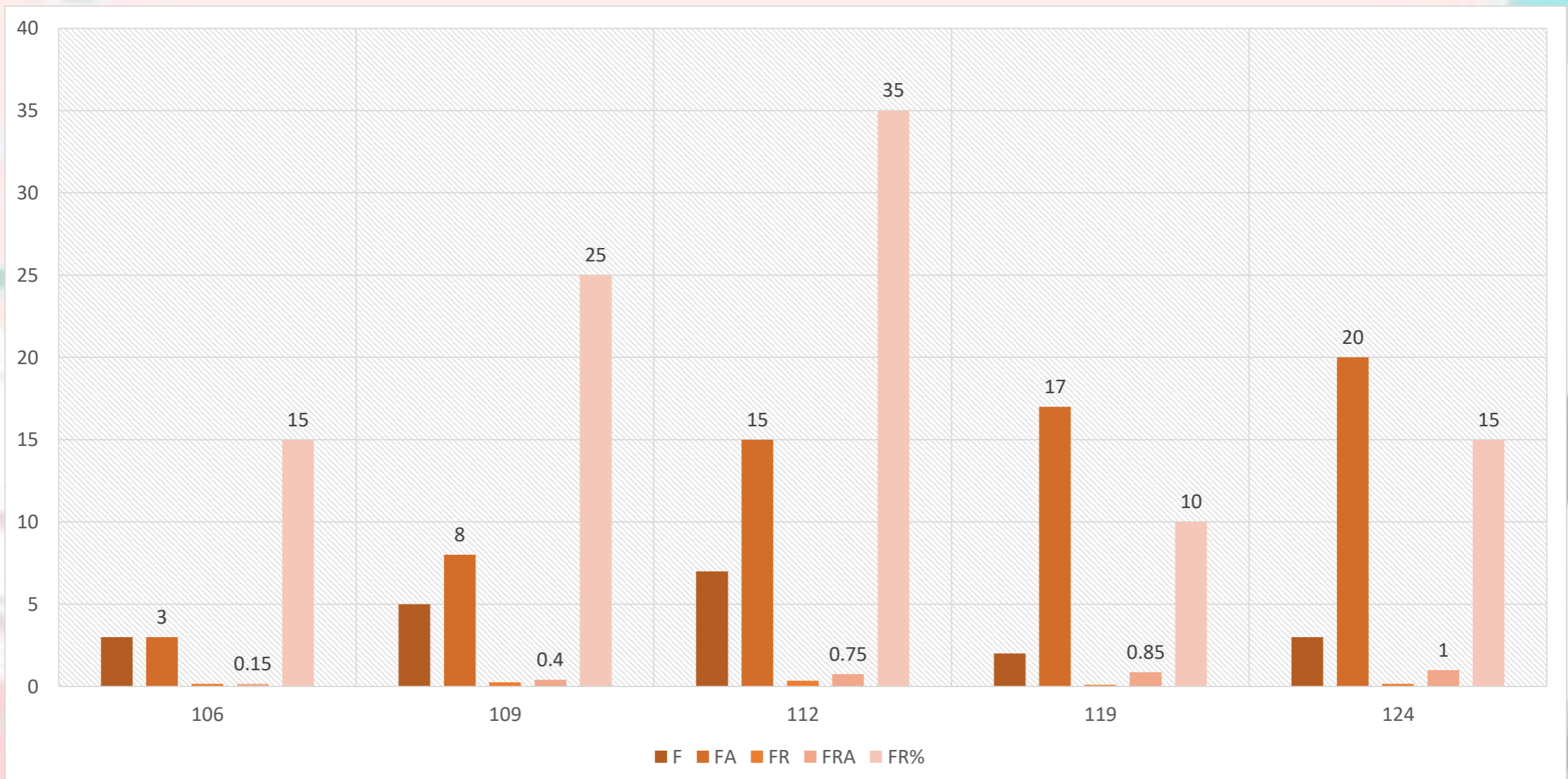
La diferencia entre ambas, sin embargo, es notable. La función de densidad discreta toma valores positivos únicamente en los puntos del recorrido y se interpreta como la probabilidad de la que la variable tome ese valor $f(x) = P(X = x)$. No está acotada por 1, puede tomar cualquier valor positivo. Es más, en una variable continua se cumple que probabilidades definidas sobre puntos concretos siempre son nulas.

Resuelve el siguiente ejercicio

Cierta universidad realizo un experimento sobre el coeficiente intelectual (C.I) de sus alumnos, para lo cual aplico un examen de C.I a un grupo de 20 alumnos escogidos al aza, obtenidos los siguientes resultados:
119,109,124,119,106,112,112,112,112,109,112,124,109,109,106,124,112,112,106.

DATOS	F	FA	FR	FRA	FR%
106	3	3	0.15	0.15	15
109	5	8	0.25	0.4	25
112	7	15	0.35	0.75	35
119	2	17	0.1	0.85	10
124	3	20	0.15	1	15

TOTAL	20	0	1	0	100
--------------	-----------	----------	----------	----------	------------



Bibliografía

Artículo (SD). Distribución Hipergeométrica. 22/05/2021, de Proyecto Descartes
Sitio web:

https://proyectodescartes.org/iCartesiLibri/materiales_didacticos/EstadisticaProbabilidadInferencia/VAdiscreta/4_1DistribucionHipergeometrica/index.html

Aula Fácil. (2019). Independencia de sucesos. 13/08/2021, de Aula Fácil Sitio web:
<https://www.aulafacil.com/cursos/estadisticas/gratis/independencia-de-sucesos-l11238>

Arrondo, V. (2020). Regresión y correlación. 13/08/2021, de Sites Sitio web:
<https://www.ugr.es/~jsalinas/apuntes/C5.pdf>