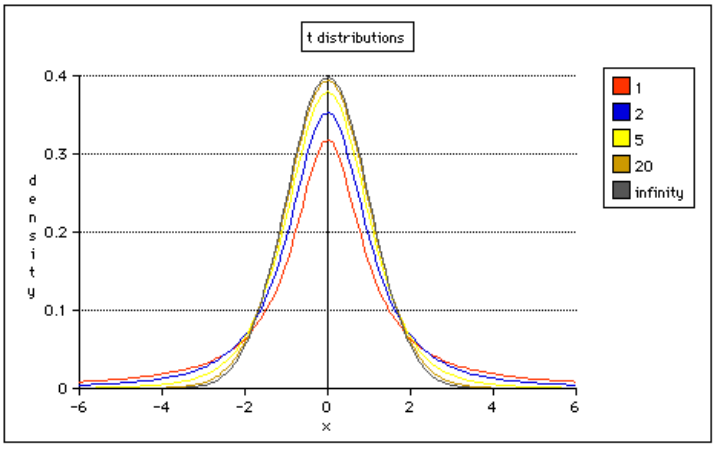
UNIVERSIDAD TECNICA DE ORURO

FACULTAD DE CIENCIAS AGRARIAS Y NATURALES

MANUAL ESTADISTICA BASICA

$$\chi^2 = \sum_{i=1}^{n} \left(\frac{X_i - \mu_i}{\sigma_i} \right)^2 = \sum_{i=1}^{n} Z_i^2$$



Por: Freddy Fernández Camacho

2019

PROLOGO

La materia prima de la Estadística Básica consiste en conjuntos de números obtenidos al contar o medir elementos. Al recopilar datos estadísticos se ha de tener especial cuidado para garantizar que la información sea completa y correcta.

El primer problema para los estadísticos reside en determinar qué información y en que cantidad se ha de reunir. En realidad, la dificultad al compilar un censo está en obtener el número de habitantes de forma completa y exacta; de la misma manera que un físico que quiere contar el número de colisiones por segundo entre las moléculas de un gas debe empezar determinando con precisión la naturaleza de los objetos a contar. Los estadísticos se enfrentan a un complejo problema cuando, por ejemplo, toman una muestra para un sondeo de opinión o una encuesta electoral. El seleccionar una muestra capaz de representar con exactitud las preferencias del total de la población no es tarea fácil.

Para establecer una ley física, biológica o social, el estadístico debe comenzar con un conjunto de datos y modificarlo basándose en la experiencia. Por ejemplo, en los primeros estudios sobre crecimiento de la población, los cambios en el número de habitantes se predecían calculando la diferencia entre el número de nacimientos y el de fallecimientos en un determinado lapso. Los expertos en estudios de población comprobaron que la tasa de crecimiento depende sólo del número de nacimientos, sin que el número de defunciones tenga importancia. *Por tanto, el futuro crecimiento de la población se empezó a calcular basándose en el número anual de nacimientos por cada 1.000 habitantes.* Sin embargo, pronto se dieron cuenta que las predicciones obtenidas utilizando este método no daban resultados correctos. Los estadísticos comprobaron que hay otros factores que limitan el crecimiento de la población. Dado que el número de posibles nacimientos depende del número de mujeres, y no del total de la población, y dado que las mujeres sólo tienen hijos durante parte de su vida, el dato más importante que se ha de utilizar para predecir la población es el número de niños nacidos vivos por cada 1.000 mujeres en edad de procrear. El valor obtenido utilizando este dato mejora al combinarlo con el dato del porcentaje de mujeres sin descendencia. Por tanto, la diferencia entre nacimientos y fallecimientos sólo es útil para indicar el crecimiento de población en un determinado periodo de tiempo del *pasado*, el número de nacimientos por cada 1.000 habitantes sólo expresa la tasa de crecimiento en el mismo periodo, y sólo el número de nacimientos por cada 1.000 mujeres en edad de procrear sirve para predecir el número de habitantes en el *futuro.*

El Autor.

INDICE TEMATICO

TEMA	Introducción. HISTORIA, METODOS ESTADISTICOS		
1	Estadística descriptiva	12	
2	Distribución de Frecuencias	15	
3	Medidas de Posición Central	20	
	Medidas de posición no Central Cuartiles, Deciles, Percentiles	26	
4	Medidas de Dispersión	28	
	Medidas de forma, Grado de Concentración	30	
	Coeficiente de Asimetría	33	
5	Medidas de Forma Coeficiente de Curtosis	35	
6	Combinaciones, Variaciones y Permutaciones (I)	37	
	Combinaciones, Variaciones y Permutaciones (II y III)	39	
7	Teoría Elemental de Probabilidad: Introducción	42	
	Probabilidad: Relación entre sucesos	43	
	Cálculo de probabilidades	44	
	Probabilidad de sucesos	46	
	Probabilidad condicionada	51	
	Probabilidad compuesta	52	
	Teorema de la probabilidad total	53	
8	Distribuciones bidimensionales	56	
	Distribuciones marginales	57	
	Coeficiente de correlación lineal	59	
	Regresión lineal	62	
9	Teorema de Bayes	64	
10	Distribuciones discretas: Binomial	68	
	Distribuciones discretas: Poisson	69	
	Distribuciones discretas: Multinomial	73	
	Bibliografia	90	92

INTRODUCCION A LA ESTADÍSTICA BASICA

Estadística, rama de las matemáticas que se ocupa de reunir, organizar y analizar datos numéricos y que ayuda a resolver problemas como el diseño de experimentos y la toma de decisiones.

HISTORIA

Desde los comienzos de la civilización han existido formas sencillas de estadística, pues ya se utilizaban representaciones gráficas y otros símbolos en pieles, rocas, palos de madera y paredes de cuevas para contar el número de personas, animales o cosas. *Hacia el año 3000 a.C. los babilonios usaban pequeñas tablillas de arcilla* para recopilar datos sobre la producción agrícola y sobre los géneros vendidos o cambiados mediante trueque. *En el siglo XXXI a.C., mucho antes de construir las pirámides, los egipcios* analizaban los datos de la población y la renta del país. Los libros bíblicos de Números y Crónicas incluyen, en algunas partes, trabajos de estadística. El primero contiene dos censos de la población de Israel y el segundo describe el bienestar material de las diversas tribus judías. En China existían registros numéricos similares con anterioridad al año 2000 a.C. *Los griegos clásicos realizaban censos* cuya información se utilizaba hacia el 594 a.C. para cobrar impuestos.

El Imperio romano fue el primer gobierno que recopiló una gran cantidad de datos sobre la población, superficie y renta de todos los territorios bajo su control. Durante la edad media sólo se realizaron algunos censos exhaustivos en Europa. Los reyes caloringios Pipino el Breve y Carlomagno ordenaron hacer estudios minuciosos de las propiedades de la Iglesia en los años 758 y 762 respectivamente. Después de la conquista normanda de Inglaterra en 1066, el rey Guillermo I de Inglaterra encargó la realización de un censo. La información obtenida con este censo, llevado a cabo en 1086, se recoge en el Domesday Book. El registro de nacimientos y defunciones comenzó en Inglaterra a principios del siglo XVI, y en 1662 apareció el primer estudio estadístico notable de población, titulado *Observations on the London Bills of Mortality* (*Comentarios sobre las partidas de defunción en Londres*). Un estudio similar sobre la tasa de mortalidad en la ciudad de Breslau, en Alemania, realizado en 1691, fue utilizado por el astrónomo inglés Edmund Halley como base para la primera tabla de mortalidad. *En el siglo XIX, con la generalización del método científico para estudiar todos los fenómenos de las ciencias naturales y sociales,* los investigadores aceptaron la necesidad de reducir la información a valores numéricos para evitar la ambigüedad de las descripciones verbales.

En nuestros días, la estadística se ha convertido en un método efectivo para describir con exactitud los valores de datos económicos, políticos, sociales, psicológicos, biológicos o físicos, y sirve como herramienta para relacionar y analizar dichos datos. *El trabajo del experto estadístico no consiste ya sólo en reunir y tabular los datos, sino sobre todo en el proceso de "interpretación" de esa información.* El desarrollo de la teoría de la probabilidad ha aumentado el alcance de las aplicaciones de la estadística. Muchos conjuntos de datos se pueden aproximar, con gran exactitud, utilizando determinadas distribuciones probabilísticas; los resultados de éstas se pueden utilizar para analizar datos estadísticos. La probabilidad es útil para comprobar la fiabilidad de las inferencias estadísticas y para predecir el tipo y la cantidad de datos necesarios en un determinado estudio estadístico.

MÉTODOS ESTADÍSTICOS

La materia prima de la estadística consiste en conjuntos de números obtenidos al contar o medir elementos. Al recopilar datos estadísticos se ha de tener especial cuidado para garantizar que la información sea completa y correcta.

El primer problema para los estadísticos reside en determinar qué información y en que cantidad se ha de reunir. En realidad, la dificultad al compilar un censo está en obtener el número de habitantes de forma completa y exacta; de la misma manera que un físico que quiere contar el número de colisiones por segundo entre las moléculas de un gas debe empezar determinando con precisión la naturaleza de los objetos a contar. Los estadísticos se enfrentan a un complejo problema cuando, por ejemplo, toman una muestra para un sondeo de opinión o una encuesta electoral. El seleccionar una muestra capaz de representar con exactitud las preferencias del total de la población no es tarea fácil.

Para establecer una ley física, biológica o social, el estadístico debe comenzar con un conjunto de datos y modificarlo basándose en la experiencia. Por ejemplo, en los primeros estudios sobre crecimiento de la población, los cambios en el número de habitantes se predecían calculando la diferencia entre el número de nacimientos y el de fallecimientos en un determinado lapso. Los expertos en estudios de población comprobaron que la tasa de crecimiento depende sólo del número de nacimientos, sin que el número de defunciones tenga importancia. *Por tanto, el futuro crecimiento de la población se empezó a calcular basándose en el número anual de nacimientos por cada 1.000 habitantes.* Sin embargo, pronto se dieron cuenta que las predicciones obtenidas utilizando este método no daban resultados correctos. Los estadísticos comprobaron que hay otros factores que limitan el crecimiento de la población. Dado que el número de posibles nacimientos depende del número de mujeres, y no del total de la población, y dado que las mujeres sólo tienen hijos durante parte de su vida, el dato más importante que se ha de utilizar para predecir la población es el número de niños nacidos vivos por cada 1.000 mujeres en edad de procrear. El valor obtenido utilizando este dato mejora al combinarlo con el dato del porcentaje de mujeres sin descendencia. Por tanto, la diferencia entre nacimientos y fallecimientos sólo es útil para indicar el crecimiento de población en un determinado periodo de tiempo del *pasado*, el número de nacimientos por cada 1.000 habitantes sólo expresa la tasa de crecimiento en el mismo periodo, y sólo el número de nacimientos por cada 1.000 mujeres en edad de procrear sirve para predecir el número de habitantes en el *futuro.*

POBLACIÓN, INDIVIDUO, CARÁCTER

El primer campo de actuación de la estadística, como se ha visto, es la demografía. De esta ciencia ha tomado la nomenclatura (población, individuo…).

Se llama población al conjunto de todos los elementos cuyo conocimiento interesa. Cada uno de esos elementos es un *individuo*. Si se está estudiando el resultado de ciertos experimentos químicos, cada uno de esos experimentos será un individuo estadístico y el conjunto de todos los posibles experimentos en esas condiciones será la población.

Cada individuo puede ser descrito mediante uno o varios caracteres. Por ejemplo, si los individuos son personas, el *sexo, el estado civil, el número de hermanos o su estatura* son caracteres. Y si el individuo es una reacción química, el tiempo de reacción, la cantidad de producto obtenido o si éste es ácido o básico serán posibles caracteres que pueden analizarse.

Un carácter puede ser cuantitativo si es medible numéricamente o cualitativo si no admite medición numérica. El número de hermanos y la estatura son caracteres *cuantitativos* mientras que el sexo y el estado civil son caracteres *cualitativos.*

Los distintos valores que puede tomar un carácter cuantitativo configuran una variable estadística. La variable *estatura*, en cierta población estadística, toma valores en el intervalo 147-205; y la variable *número de hermanos* toma los valores 0, 1, 2, 3, 4, 5, 6, 7 y 8. Una *variable estadística como esta última es discreta*, ya que sólo admite valores aislados. *Una variable estadística es continua si admite todos los valores* de un intervalo, como ocurre con la *estatura*.

Una población puede ser finita o infinita ejemplo en todas *las galletas* producidas por una fábrica un cierto día es finita, mientras la determinación posible de caras cruces sucesivas de lanzamientos de una moneda es infinita.

ESTADÍSTICA DESCRIPTIVA

- La estadística descriptiva analiza, estudia y describe a la totalidad de individuos de una población. Su finalidad es obtener información, analizarla, elaborarla y simplificarla lo necesario para que pueda ser interpretada cómoda y rápidamente y, por tanto, pueda utilizarse eficazmente para el fin que se desee. El proceso que sigue la estadística descriptiva para el estudio de una cierta población consta de los siguientes pasos:

  • Selección de caracteres dignos de ser estudiados.
  • Mediante encuesta o medición, obtención del valor de cada individuo en los caracteres seleccionados.
  • Elaboración de tablas de frecuencias, mediante la adecuada clasificación de los individuos dentro de cada carácter.
  • Representación gráfica de los resultados (elaboración de gráficas estadísticas).
  • Obtención de parámetros estadísticos, números que sintetizan los aspectos más relevantes de una distribución estadística.

ESTADÍSTICA INFERENCIAL

La estadística descriptiva trabaja con todos los individuos de la población. La estadística inferencial, sin embargo, trabaja con muestras, subconjuntos formados por algunos individuos de la población. A partir del estudio de la muestra se pretende inferir aspectos relevantes de toda la población. Cómo se selecciona la muestra, cómo se realiza la inferencia, y qué grado de confianza se puede tener en ella son aspectos fundamentales de la estadística inferencial, para cuyo estudio se requiere un alto nivel de conocimientos de estadística, probabilidad y matemáticas.

VARIABLES DISCRETAS Y CONTINUAS

Una variables es un símbolo tal como X,Y,H, o B que puede tomar un conjunto de prefijado de valores llamado dominio de esa variable, si la variable toma un solo valor se llama constante.

Una variable que puede tomar cualquier valor entre dos valores dados se dice que es una variable continua en caso contrario diremos que es discreta. Ejemplo el número de hijos de una familia puede ser 0,1,2,3,4, pero no puede ser 2.5, 3.5 o sea es una *variable discreta.*

En cambio las alturas en pulgadas de los estudiantes en Ciencias Agrarias es una *variable continua*

REDONDEO DE DATOS

El resultado como redondear los datos como 22.5 en una unidad será 22 el más próximo, en cambio si se tiene un numero 72.546548 redondear a la centésima será 72.55.

NOTACION CIENTIFICA

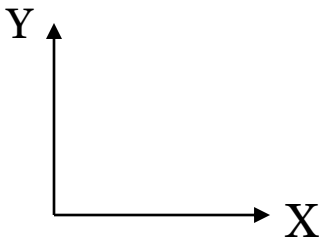
Al escribir números especialmente los que tienen muchos ceros y se utiliza un exponente de 10, ejemplo 10^2 será 100, 10^5 será 100.000 o al revés 10^{-2} será 0.01 , 10^{-5} será 0.00001

DATOS SIGNIFICATIVOS

Al escribir números que se anota con la precisión posible como 65.4 (tres datos significativos) lo que significa que esta entre 65.35 y 65.45

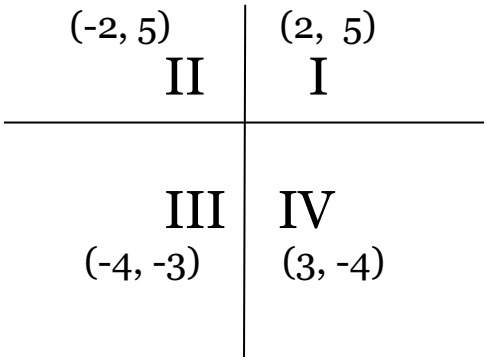
FUNCIONES

Es al escribir Y = f(x) la variable X se llama variable independiente y la Y dependiente



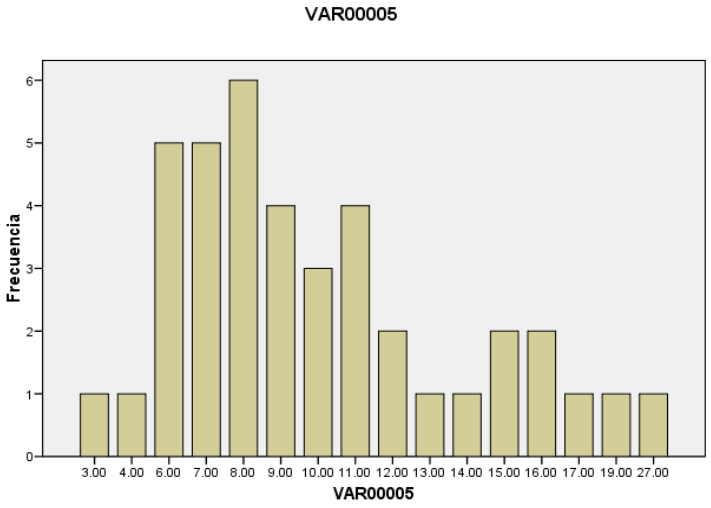
COORDENADAS RECTANGULARES

En el cuadro consideramos dos rectas perpendiculares X' o X e Y' o Y llamados ejes X e Y respectivamente estas rectas se dividen en un plano cartesiano de las cuatro regiones denotadas por I, II, III, IV cuadrantes llamados primero segundo, tercero y cuarto cuadrante.



GRAFICAS

Gráficas estadísticas, representaciones gráficas de los resultados que se muestran en una tabla estadística. Pueden ser de formas muy diversas, pero con cada tipo de gráfica se cumple un propósito. Por ejemplo, en los medios de comunicación, libros de divulgación y revistas especializadas se encuentran multitud de gráficas estadísticas en las que, con notable expresividad, se ponen de manifiesto los rasgos de la distribución que se pretende destacar. Los diagramas de barras, los diagramas de sectores, los histogramas y los polígonos de frecuencias son algunas de ellas.



ECUACIONES

Las ecuaciones son enunciados del tipo A =B Donde A se llama miembro (o lado) izquierdo y B miembro derecho de la ecuación. Por tanto se pueden sumar, restar, multiplicar o dividir ambos lados de las ecuaciones por el mismo número y se llegara a una ecuación equivalente.

Dado la Ecuación $2x + 3 = 9$ Ambos restamos (-3) en la ecuación. $2x +3 -3 = 9 – 3$; $2x =6$
Entonces: dividimos entre 2
Tenemos. $2x/2 = 6/2$
Finalmente: $x = 3$

Este resultado se reemplaza en la ecuación.

$2(3) + 3 = 9$

$9 = 9 = 0$

Es una identidad el proceso de resolver la ecuación.

DESIGUALDADES

Las desigualdades son símbolos > < que significan mayor y menor respectivamente son símbolos que sirven para expresar igualdades entre números, ejemplo 2 < 5 se lee dos es menor que cinco, 5>3. 5 =5 o viceversa.

2 < 5 se lee 3 menor que 5
5> 3 se lee 5 es mayor que 3

LOGARITMOS

Todo numero positivo N se puede expresar como una potencia de 10 siempre es posible encontrar p tal que $N = 10^p$ p se llama logaritmo de N en base 10.
Ce escribe así: $p = \log N$; $p = \log_{10} N$

Ej. Log de 1000 es 10^3
Entonces; log de 1000 = 3

GRAFICAS ESTADISTICAS

Gráficas estadísticas, representaciones gráficas de los resultados que se muestran en una tabla estadística. Pueden ser de formas muy diversas, pero con cada tipo de gráfica se cumple un propósito. Por ejemplo, en los medios de comunicación, libros de divulgación y revistas especializadas se encuentran multitud de gráficas estadísticas en las que, con notable expresividad, se ponen de manifiesto los rasgos de la distribución que se pretende destacar. Los diagramas de barras, los diagramas de sectores, los histogramas y los polígonos de frecuencias son algunas de ellas.

DIAGRAMA DE BARRAS

En este tipo de gráfica, sobre los valores de las variables se levantan barras estrechas de longitudes proporcionales a las frecuencias correspondientes. Se utilizan para representar variables cuantitativas discretas.

El diagrama de barras siguiente representa la distribución del número de hijos de 43 familias:

HISTOGRAMA Y POLÍGONO DE FRECUENCIAS

Los histogramas se utilizan para representar tablas de frecuencias con datos agrupados en intervalos. Si los intervalos son todos iguales, cada uno de ellos es la base de un rectángulo cuya altura es proporcional a la frecuencia correspondiente. El histograma que se muestra a continuación es el correspondiente a la tabla de frecuencias con intervalos adjunta (1.200 calificaciones distribuidas en 10 intervalos):

Si se unen los puntos medios de la base superior de los rectángulos se obtiene el polígono de frecuencias.

HISTOGRAMA Y POLÍGONO DE FRECUENCIAS ACUMULADAS

Si se representan las frecuencias acumuladas de una tabla de datos agrupados se obtiene el histograma de frecuencias acumuladas o su correspondiente polígono. He aquí los que se obtienen de la tabla de 1.200 calificaciones:

Tablas estadísticas

INTRODUCCIÓN

Tablas estadísticas, recopilaciones numéricas bien estructuradas y fáciles de interpretar de las que se vale el estadístico para sintetizar los datos obtenidos con el fin de hacer un uso sencillo de ellos o bien para darlos a conocer de forma comprensible.

Existen infinidad de tablas estadísticas, pero las más básicas son las tablas de frecuencias, las de frecuencias relativas y frecuencias acumuladas, las de frecuencias con datos agrupados en intervalos y las de doble entrada.

TABLAS DE FRECUENCIAS

Estas tablas constan de dos columnas. En la primera se escriben los valores de la variable, x_i. En la segunda las correspondientes frecuencias, f_i. Estas sencillas tablas se utilizan, únicamente, cuando la variable es discreta y admite pocos valores (a lo sumo, de 12 a 16).

La tabla siguiente da la distribución de la variable “número de hijos” correspondiente a un conjunto de 43 familias:

TABLAS DE FRECUENCIAS RELATIVAS Y FRECUENCIAS ACUMULADAS

Una tabla de frecuencias se puede ampliar con nuevas columnas con las frecuencias relativas y las frecuencias acumuladas. La tabla anterior con estos nuevos datos sería:

TABLAS DE FRECUENCIAS CON DATOS AGRUPADOS EN INTERVALOS

Cuando la variable es continua, o es discreta pero toma una gran cantidad de valores, conviene dividir el rango de la variable en unos pocos intervalos (entre 6 y 12) y repartir los valores en ellos. El resultado será una tabla de frecuencias en la cual la variable, en lugar de tomar valores numéricos concretos, varía dentro de intervalos.

Cuando se necesita (por ejemplo para el cálculo de parámetros) que cada intervalo quede representado por un único número, se toma su punto medio, al que se llama marca de clase.

En la tabla adjunta se muestra cómo se han repartido 1.200 calificaciones entre 0 y 10, en 10 intervalos iguales —columna (*a*). Las marcas de clase (centros de los intervalos) están en la columna (*b*), las frecuencias en la (*c*), las frecuencias relativas en la (*d*), las frecuencias acumuladas en la (*e*) y las frecuencias acumuladas relativas en la columna (*f*).

TABLAS DE DOBLE ENTRADA

En las distribuciones bidimensionales, en las que a cada individuo le corresponden dos valores, x_i, y_i, puede suceder que cada par de valores (x_m, y_n) ocurra varias veces, es decir, lleve apareada una frecuencia. En tal caso conviene disponer los resultados mediante una tabla de doble entrada como la que se muestra a continuación, correspondiente a los resultados de un colectivo de 125 personas puntuadas por su sensibilidad ecológica, x_i, y por sus conocimientos de biología, y_i:

DIAGRAMA DE SECTORES

En un diagrama de este tipo, los 360º de un círculo se reparten proporcionalmente a las frecuencias de los distintos valores de la variable. Resultan muy adecuados cuando hay pocos valores, o bien cuando el carácter que se estudia es cualitativo. El diagrama de sectores siguiente refleja el resultado de una encuesta (realizada a 300 personas) sobre los tipos de película preferidos por el público en general:

TEMA 1ª

Estadística Descriptiva

La ***estadística descriptiva*** es una ciencia que analiza series de datos (por ejemplo, edad de una población, altura de los estudiantes de una escuela, temperatura en los meses de verano, etc.) y trata de extraer conclusiones sobre el comportamiento de estas variables.

Las ***variables*** pueden ser de dos tipos:

Variables cualitativas o atributos: no se pueden medir numéricamente (por ejemplo: nacionalidad, color de la piel, sexo).

Variables cuantitativas: tienen valor numérico (edad, precio de un producto, ingresos anuales).

Las ***variables*** también se pueden clasificar en:

Variables unidimensionales: sólo recogen información sobre una característica (por ejemplo: edad de los alumnos de una clase).

Variables bidimensionales: recogen información sobre dos características de la población (por ejemplo: edad y altura de los alumnos de una clase).

Variables pluridimensionales: recogen información sobre tres o más características (por ejemplo: edad, altura y peso de los alumnos de una clase).

Por su parte, las ***variables cuantitativas*** se pueden clasificar en discretas y continuas:

Discretas: sólo pueden tomar valores enteros (1, 2, 8, -4, etc.). Por ejemplo: número de hermanos (puede ser 1, 2, 3...., etc, pero, por ejemplo, nunca podrá ser 3,45).

Continuas: pueden tomar cualquier valor real dentro de un intervalo. Por ejemplo, la velocidad de un vehículo puede ser 80,3 km/h, 94,57 km/h.. Cantidad de cabello en un cm^2 de área en la cabeza humana, el valor de π 3,14159265359 etc.

Cuando se estudia el comportamiento de una variable hay que distinguir los siguientes conceptos:

Individuo: cualquier elemento que porte información sobre el fenómeno que se estudia. Así, si estudiamos la altura de los niños de una clase, cada alumno es un individuo; si estudiamos el precio de la vivienda, cada vivienda es un individuo.

Población: conjunto de todos los individuos (personas, objetos, animales, etc.) que porten información sobre el fenómeno que se estudia. Por ejemplo, si estudiamos el precio de la vivienda en una ciudad, la población será el total de las viviendas de dicha ciudad.

Muestra: subconjunto que seleccionamos de la población. Así, si se estudia el precio de la vivienda de una ciudad, lo normal será no recoger información sobre todas las viviendas de la ciudad (sería una labor muy compleja), sino que se suele seleccionar un subgrupo (muestra) que se entienda que es suficientemente representativo.

La Estadística estudia los métodos científicos para recoger, organizar, resumir y analizar datos, así como para sacar conclusiones válidas y tomar decisiones razonables basadas con tal análisis.

En un sentido menos amplio, el término estadística se usa para denotar los propios datos, o número derivados de ellos, tales como los promedios. Así se habla de estadística de empleo, estadística de accidentes.

Estadístico:

Unidad de medida referente a la muestra. Se le llama estadístico también a la persona que trabaja con la estadística.

Frecuencia:

Número de veces en que se repite un dato.

La frecuencia absoluta f(xi) se determina como el número de veces que se repite un dato xi.

Frecuencia absoluta acumulada Fi: Para un determinado valor se considera como la frecuencia de cada dato xi mas la suma de los valores anteriores a dicha suma

Es el número de estudiantes con calificaciones iguales o menores que el rango de cada intervalo sucesivo. (Frecuencia) donde xi son el número de datos acumulados de las clases anteriores

Frecuencia Relativa:

La frecuencia relativa hi : es el cociente hi =fi/N , en algunas ocasiones se representa como ni/N, donde N es el número total de datos, corresponde a la suma de todas la frecuencias individuales de cada clase, y ni o fi los datos en cada clase. Las frecuencias relativas representan el porciento de veces en que ocurre un dato. Como veremos en teoría de probabilidades, el concepto de frecuencia relativa nos conduce a un concepto de probabilidad.

Es la proporción entre la frecuencia de un intervalo y el número total de datos.

Histograma:

Es una serie de rectángulos con bases iguales al rango de los intervalos y con área proporcional a sus frecuencias.

Inferir:

Es emitir juicios o conclusiones basados en algún conocimiento o experiencia sobre un evento o suceso.

Inferencia Estadística:

Es aplicar resultados de estudios de una muestra a las poblaciones y emitir juicios o conclusiones sobre esa población en general. (Estadística)

Ejemplo:

Estudio: Causas de la deserción estudiantil en una Facultad de Ciencias Agrarias
Población: 1200 alumnos.
Muestra: 10% de la población.
Resultados del Estudio de la Muestra: La situación económica, dificultad en el aprendizaje.

Conclusiones: ***Se puede inferir que las causas de deserción de los alumnos de la*** Facultad de Ciencias Agrarias y Veterinarias.
Son la situación económica y la dificultad en el aprendizaje

Intervalo de Clase:

Pequeña sección de la escala según la cual se agrupan las puntuaciones de una distribución de frecuencia, Tamaño o rango de la clase.

Límites del Intervalo:

Son los valores extremos que tiene el intervalo de clase, inferior y superior, entre los cuales van a estar los valores de los datos agrupados en el intervalo de clase.

TEMA 2ª
Distribución de frecuencia

DISTRIBUCIÓN DE FRECUENCIAS

Organización de datos agrupados

Definiciones

Clases o intervalos de clase: Grupo de valores que describen una característica. Deben incluir todas las observaciones y ser excluyentes. Los intervalos contienen los límites de clase que son los puntos extremos del intervalo. Se denominan intervalos cerrados, cuando contienen ambos límites e intervalos abiertos si incluyen solo un límite.

Limites Reales: Sirven para mantener la continuidad de las clases

Anchura o tamaño del intervalo: es la diferencia entre los límites reales de una clase

Número de clases: es el número total de grupos en que se clasifica la información, se recomienda que no sea menor que 5 ni mayor que 15

Marca de Clase: Es el punto medio del intervalo de clase, se recomienda observar que los puntos medios coincidan con los datos observados para minimizar el error.

Frecuencia: es el número de veces que aparece un valor

Frecuencia Acumulada: Indica cuantos casos hay por debajo o arriba de un determinado valor o límite de clase.

Frecuencia Relativa: Indica la proporción que representa la frecuencia de cada intervalo de clase en relación al total, es útil para comparar varias distribuciones con parámetros de referencia uniformes.

Frecuencia Acumulada Relativa: Indica la proporción de datos que se encuentra por arriba o debajo de cierto valor o límite de clase.

El manejo de la información requiere de la ordenación de datos de tal forma que permita la obtención de una forma más fácil la obtención de conclusiones acerca de la muestra. Una primera ordenación se realiza mediante el manejo de tablas, en las que se ordenan los datos de acuerdo a ciertas características de los datos.

El manejo de datos discretos permite la manipulación de tablas, sobre todo cuando el número de datos no es muy reducido. A continuación describiremos la manipulación de datos mediante el manejo de datos discretos y continuos.

Las variables a ser manejadas en el estudio, en el caso de valores discretos, se puede representar mediante una tabla en la que se representa los variables mediante variables con nombre como (xi) y el número de veces en que un dato se representara mediante frecuencias, frecuencias absolutas, frecuencias relativas.

La representación de esta tabla mediante frecuencias se le conoce como tabla estadística cuya función es la determinar la frecuencia de cada clase las cuales aparecen a un lado de cada clase.

La ***distribución de frecuencia*** es la representación estructurada, en forma de tabla, de toda la información que se ha recogido sobre la variable que se estudia.

Variable	**Frecuencias absolutas**		**Frecuencias relativas**	
(Valor)	**Simple**	**Acumulada**	**Simple**	**Acumulada**
X_1	n_1	n_1	$f_1 = n_1 / n$	f_1
X_2	n_2	$n_1 + n_2$	$f_2 = n_2 / n$	$f_1 + f_2$
...	...	...	...	...
X_{n-1}	n_{n-1}	$n_1 + n_2 + .. + n_{n-1}$	$f_{n-1} = n_{n-1} / n$	$f_1 + f_2 + .. + f_{n-1}$
X_n	N_n	Σ n	$f_n = n_n / n$	Σ f
Siendo **X** los distintos valores que puede tomar la variable.				
Siendo **n** el número de veces que se repite cada valor.				
Siendo **f** el porcentaje que la repetición de cada valor supone sobre el total				

Veamos ***un ejemplo***:

Medimos la altura de los estudiantes de una clase y obtenemos los siguientes resultados (cm):

Alumno	**Estatura**	**Alumno**	**Estatura**	**Alumno**	**Estatura**
Alumno 1	1,25	Alumno 11	1,23	Alumno 21	1,21
Alumno 2	1,28	Alumno 12	1,26	Alumno 22	1,29
Alumno 3	1,27	Alumno 13	1,30	Alumno 23	1,26
Alumno 4	1,21	Alumno 14	1,21	Alumno 24	1,22
Alumno 5	1,22	Alumno 15	1,28	Alumno 25	1,28

Alumno 6	1,29	Alumno 16	1,30	Alumno 26	1,27
Alumno 7	1,30	Alumno 17	1,22	Alumno 27	1,26
Alumno 8	1,24	Alumno 18	1,25	Alumno 28	1,23
Alumno 9	1,27	Alumno 19	1,20	Alumno 29	1,22
Alumno 10	1,29	Alumno 20	1,28	Alumno 30	1,21

Si presentamos esta información estructurada obtendríamos la siguiente ***tabla de frecuencia****:*

Variable	**Frecuencias absolutas**		**Frecuencias relativas**	
(Valor)	**Simple**	**Acumulada**	**Simple**	**Acumulada**
1,20	1	1	3,3%	3,3%
1,21	4	5	13,3%	16,6%
1,22	4	9	13,3%	30,0%
1,23	2	11	6,6%	36,6%
1,24	1	12	3,3%	40,0%
1,25	2	14	6,6%	46,6%
1,26	3	17	10,0%	56,6%
1,27	3	20	10,0%	66,6%
1,28	4	24	13,3%	80,0%
1,29	3	27	10,0%	90,0%
1,30	3	30	10,0%	100,0%

Si los valores que toma la variable son muy diversos y cada uno de ellos se repite muy pocas veces, entonces conviene agruparlos por intervalos, ya que de otra manera obtendríamos una tabla de frecuencia muy extensa que aportaría muy poco valor a efectos de síntesis. (Tal como se verá en la siguiente lección).

Gráficos de una Distribución de Frecuencias

Los gráficos son útiles porque ponen en relieve y aclaran las tendencias que no se captan fácilmente en la tabla, ayudan a estimar valores con una simple ojeada y brinda una verificación gráfica de la veracidad de las soluciones.

Histograma:

Esta formado por rectángulos cuya base es la amplitud del intervalo y tiene la característica que la superficie que corresponde a las barras es representativa de la cantidad de casos o frecuencia de cada tramo de valores, puede construirse con clases que tienen el mismo tamaño o diferente (intervalo variable). La utilización de los intervalos de amplitud variable se recomienda cuando en alguno de los intervalos, de amplitud constante, se presente la frecuencia cero o la frecuencia de alguno o algunos de los intervalos sea mucho mayor que

la de los demás, logrando así que las observaciones se hallen mejor repartidas dentro del intervalo.

Polígono de Frecuencias

Se puede obtener uniendo cada punto medio (marca de clase) de los rectángulos del histograma con líneas rectas, teniendo cuidado de agregar al inicio y al final marcas de clase adicionales, con el objeto de asegurar la igualdad de las áreas.

Curvas de frecuencia

No es más que la curva suavizada que se traza sobre el polígono y representa la asimetría y la curtosis que tiene la distribución, permite visualizar un esquema más claro del patrón de datos. Existen varios tipos de curva de frecuencia: Curvas J, Simétricas o Asimétricas (sesgada a la derecha o a la izquierda), Unimodales, Bimodales y Multimodales.

Ojivas: Cuando se trata de relacionar observaciones en un mismo aspecto para dos colectivos diferentes no es posible ejecutar comparaciones sobre la base de la frecuencia, es necesario tener una base estándar, la frecuencia relativa. La ojiva representa gráficamente la forma en que se acumulan los datos y permiten ver cuantas observaciones se hallan por arriba o debajo de ciertos valores. Es útil para obtener una medida de los cuartiles, deciles , percentiles.

Distribuciones de frecuencia agrupada

Supongamos que medimos la estatura de los habitantes de una vivienda y obtenemos los siguientes resultados (cm):

Habitante	Estatura	Habitante	Estatura	Habitante	Estatura
Habitante 1	1,15	Habitante 11	1,53	Habitante 21	1,21
Habitante 2	1,48	Habitante 12	1,16	Habitante 22	1,59
Habitante 3	1,57	Habitante 13	1,60	Habitante 23	1,86
Habitante 4	1,71	Habitante 14	1,81	Habitante 24	1,52
Habitante 5	1,92	Habitante 15	1,98	Habitante 25	1,48
Habitante 6	1,39	Habitante 16	1,20	Habitante 26	1,37
Habitante 7	1,40	Habitante 17	1,42	Habitante 27	1,16
Habitante 8	1,64	Habitante 18	1,45	Habitante 28	1,73
Habitante 9	1,77	Habitante 19	1,20	Habitante 29	1,62
Habitante 10	1,49	Habitante 20	1,98	Habitante 30	1,01

Si presentáramos esta información en una tabla de frecuencia obtendríamos una tabla de 30 líneas (una para cada valor), cada uno de ellos con una frecuencia absoluta de 1 y con una frecuencia relativa del 3,3%. Esta tabla nos aportaría escasa información

En lugar de ello, preferimos agrupar los datos por intervalos, con lo que la información queda más resumida (se pierde, por tanto, algo de información), pero es más manejable e informativa:

Estatura	Frecuencias absolutas		Frecuencias relativas	
Cm	Simple	Acumulada	Simple	Acumulada
1,01 - 1,10	1	1	3,3%	3,3%
1,11 - 1,20	3	4	10,0%	13,3%
1,21 - 1,30	3	7	10,0%	23,3%
1,31 - 1,40	2	9	6,6%	30,0%
1,41 - 1,50	6	15	20,0%	50,0%
1,51 - 1,60	4	19	13,3%	63,3%
1,61 - 1,70	3	22	10,0%	73,3%
1,71 - 1,80	3	25	10,0%	83,3%
1,81 - 1,90	2	27	6,6%	90,0%
1,91 - 2,00	3	30	10,0%	100,0%

El número de tramos en los que se agrupa la información es una decisión que debe tomar el analista: la regla es que mientras más tramos se utilicen menos información se pierde, pero puede que menos representativa e informativa sea la tabla.

BIBLIOGRAFÍA

Caballero , Wilfredo Introducción a la Estadística Serie Libros y Materiales Educativos N° 28 I edición. San José, Costa Rica IICA, 1981 Carrasquilla E. Pedro Manual para la confección de gráficos estadístico DEC-CGR, Dirección de Estadísticas y Censos Panamá. República de Panamá. DEC-CGR Manual para la elaboración y publicación de Cuadros estadísticos (tercera edición) Dirección de estadística y Censos. Panamá. República de Panamá. Nuñez del Benavente Estadística básica para planificación. 6ta.edición Arturo Siglo XX! Editores S.A, México. 1977

TEMA 3ª
Medidas de posición central

Medidas de Tendencia Central:

Un promedio es un valor típico o representativo de un conjunto de datos. Tales valores suelen situarse hacia al centro del conjunto de datos ordenados por magnitud.

Una vez que los datos han sido reunidos y tabulados, comienza el análisis con el objeto de calcular un número único, que represente o resuma todos los datos. Dado que por lo general la frecuencia de los intervalos centrales es mayor que el resto, este número se suele denominar valor o medida de la tendencia central.

$$\overline{X} = \frac{\sum_{i=1}^{N} X_i}{N}$$

Sean $X_1,X_2,...,X_n$ los datos de un estudio estadístico. El valor utilizado mas a menudo es la media aritmética o promedio aritmético que se escribe X , y que es igual a la suma de todos los valores dividida por n.

El símbolo $\sum_1$ o sumatoria, denota la suma de todos los datos. Si las X se agrupan en K intervalos, con puntos medios $m_1,m_2,..., m_k$ y frecuencias $f_1,f_2,...,f_k$, la media aritmética viene dada por:

$$Medis_Aritmetica = \frac{\sum(f_i * m_i)}{\sum f_i}$$

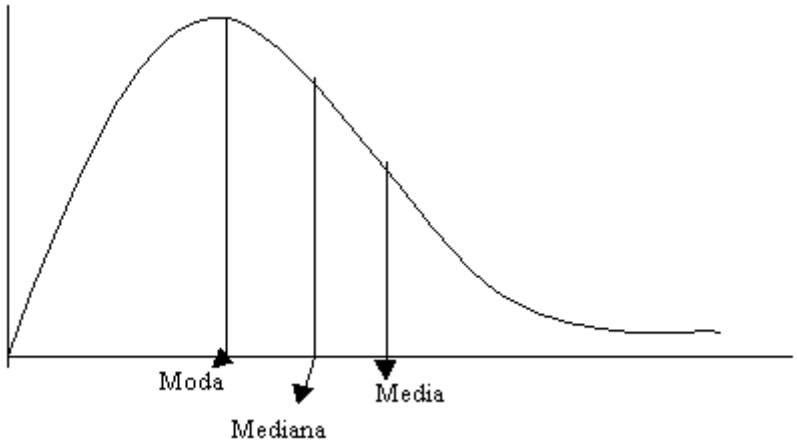
Donde i=1,2,...,k. La mediana y la moda son otros dos valores de la tendencia central. Si las x se ordenan según sus valores numéricos, si n es impar la mediana es la x que ocupa la posición central y si n es par la mediana es la media o promedio de las dos x centrales. La moda es la x que aparece con mayor frecuencia. Si dos o más x aparecen con igual máxima frecuencia, se dice que el conjunto de las x no tiene moda, o es bimodal, siendo la moda las dos x que aparecen con más frecuencia, o es trimodal, con modas las tres más frecuentes.

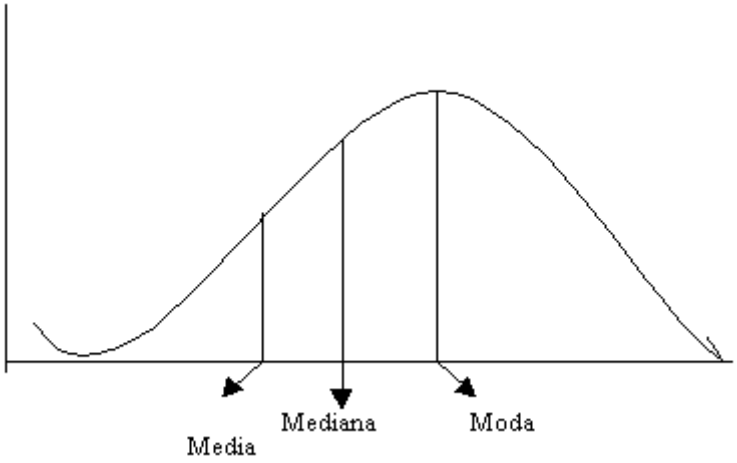
Relación Empírica entre Media, Mediana y Moda.

Para curvas de frecuencia unimodales que sean poco asimétricas tenemos la siguiente relación empírica

Media – Moda = 3(media- mediana)

Posiciones relativas de la media, la mediana y la moda para curvas de frecuencia asimétrica a derecha e izquierda, respectivamente, para curvas simétricas los tres valores coinciden.





Las medidas de posición nos facilitan información sobre la serie de datos que estamos analizando. Estas medidas permiten conocer diversas características de esta serie de datos.

Las ***medidas de posición*** son de dos tipos:

a) Medidas de posición central: informan sobre los valores medios de la serie de datos.

b) Medidas de posición no centrales: informan de como se distribuye el resto de los valores de la serie.

a) Medidas de posición central

Las principales medidas de posición central son las siguientes:

1.- Media: es el valor medio ponderado de la serie de datos. Se pueden calcular diversos tipos de media, siendo las más utilizadas:

a) Media aritmética: se calcula multiplicando cada valor por el número de veces que se repite. La suma de todos estos productos se divide por el total de datos de la muestra:

$$Xm = \frac{(X1 * n1) + (X2 * n2) + (X3 * n3) + ...+ (Xn\text{-}1 * nn\text{-}1) + (Xn * nn)}{N}$$

b) Media geométrica:

Se eleva cada valor al número de veces que se ha repetido. Se multiplican todo estos resultados y al producto final se le calcula la raíz "n" (siendo "n" el total de datos de la muestra).

Según el tipo de datos que se analice será más apropiado utilizar la media aritmética o la media geométrica.

c) Media Aritmética Ponderada:

A veces asociamos con los números X_1, X_2, X_3, .. X_k ciertos factores peso (o pesos) W_1, W_2, W_3, .. W_K dependientes de la relevancia asignada a cada número. En tal caso,

$$\overline{X} = \frac{w_1X_1 + w_2X_2 + K + w_kX_k}{w_1 + w_2 + K + w_k} = \frac{\sum wX}{\sum w}$$

Se llama la media aritmética ponderada con pesos f_1, f_2, ..., f_k.

Ejemplo: Si el examen final de un curso cuenta tres veces más que una evaluación parcial, y un estudiante tiene calificación 85 en el examen final y 70 y 90 en los dos parciales, la calificación media es:

$$\overline{X} = \frac{(1)(70) + (1)(90) + (3)(85)}{1+1+3} = 83$$

La media geométrica se suele utilizar en series de datos como tipos de interés anuales, inflación, etc., donde el valor de cada año tiene un efecto multiplicativo sobre el de los años anteriores. En todo caso, la media aritmética es la medida de posición central más utilizada.

Lo más positivo de la media es que en su cálculo se utilizan todos los valores de la serie, por lo que no se pierde ninguna información.

Sin embargo, presenta el problema de que su valor (tanto en el caso de la media aritmética como geométrica) se puede ver muy influido por valores extremos, que se aparten en exceso del resto de la serie. Estos valores anómalos podrían condicionar en gran medida el valor de la media, perdiendo ésta representatividad.

2.- Mediana: es el valor de la serie de datos que se sitúa justamente en el centro de la muestra (un 50% de valores son inferiores y otro 50% son superiores).

No presentan el problema de estar influido por los valores extremos, pero en cambio no utiliza en su cálculo toda la información de la serie de datos (no pondera cada valor por el número de veces que se ha repetido).

3.- Moda: es el valor que más se repite en la muestra.

Ejemplo: vamos a utilizar la tabla de distribución de frecuencias con los datos de la estatura de los alumnos que vimos en la lección 2ª.

Variable	Frecuencias absolutas		Frecuencias relativas	
(Valor)	Simple	Acumulada	Simple	Acumulada
1,20	1	1	3,3%	3,3%
1,21	4	5	13,3%	16,6%
1,22	4	9	13,3%	30,0%
1,23	2	11	6,6%	36,6%
1,24	1	12	3,3%	40,0%
1,25	2	14	6,6%	46,6%
1,26	3	17	10,0%	56,6%
1,27	3	20	10,0%	66,6%
1,28	4	24	13,3%	80,0%
1,29	3	27	10,0%	90,0%
1,30	3	30	10,0%	100,0%

Vamos a calcular los valores de las distintas posiciones centrales:

1.- Media aritmética:

La media de un conjunto de N números, X_1, X_2, X_3, .. X_N. Se define por:

$$\overline{X} = \frac{x_1 + x_2 + x_3 + \mathrm{K}\ x_n}{N} = \frac{\sum_{j=1}^{N} x_j}{N} = \frac{\sum x}{N}$$

$$Xm = \frac{(1,20*1) + (1,21*4) + (1,22 * 4) + (1,23 * 2) + ... + (1,29 * 3) + (1,30 * 3)}{30}$$

Luego:

Xm = 1,253

Por lo tanto, la estatura media de este grupo de alumnos es de 1,253 cm.

Propiedades de la *Media Aritmética*

1. La suma algebraica de las desviaciones de un conjunto de números respecto de su media aritmética es cero.
2. La suma de los cuadrados de las desviaciones de un conjunto de números X_j respecto de un cierto número "a" es mínima si y sólo si

   $a = \overline{X}$ SI f_1 números tienen media m_1, f_2 números tiene m_2,...,f_k números tienen media m_k, entonces la media de todos los números es:

   $$\overline{X} = \frac{f_1 m_1 + f_2 m_2 + \mathrm{K} + f_k m_k}{f_1 + f_2 + \mathrm{K} + f_k}$$

   Es decir, una media aritmética ponderada de todas las medias.

3. Si A es una media aritmética supuesta o conjeturada (que puede ser cualquier número) y si $d_j = X_j - A$ son las desviaciones de X_j respecto de A, las ecuaciones (1) y (2) se convierten, respectivamente, en lo siguiente.

$$\overline{X} = A + \frac{\sum_{j=1}^{k} f_j d_j}{\sum_{j=1}^{k} f_j} = A + \frac{\sum fd}{N} \qquad \overline{X} = A + \frac{\sum_{j=1}^{N} d_j}{N} = A + \frac{\sum d}{N}$$

Donde $N = \sum_{j=1}^{k} f_j = \sum f$

2.- Media geométrica:

X = ((1,20^ 1) * (1,21^4) * (1,22^ 4) * .....* (1,29^3)* (1,30^3)) ^ (1/30)

Luego:

Xm = 1,253

En este ejemplo la media aritmética y la media geométrica coinciden, pero no tiene siempre por qué ser así.

3.- Mediana:

La mediana de esta muestra es 1,26 cm, ya que por debajo está el 50% de los valores y por arriba el otro 50%. Esto se puede ver al analizar la columna de frecuencias relativas acumuladas.

En este ejemplo, como el valor 1,26 se repite en 3 ocasiones, la media se situaría exactamente entre el primer y el segundo valor de este grupo, ya que entre estos dos valores se encuentra la división entre el 50% inferior y el 50% superior.

La mediana de un conjunto de números ordenados en magnitud es o el valor central o la media de los dos valores centrales.

Ejemplo: El conjunto de números 3,4,4,5,6,8,8,8 y 10 tiene mediana 6.
Ejemplo: El conjunto de números 5,5,7,9,11,12,15 y18 tiene mediana.

$$Mediana = \frac{1}{2}(9+11) = 10$$

Para datos agrupados, la mediana obtenida por interpolación viene dada por:

Mediana= $$L_i + \left(\frac{\frac{N}{2} - \left(\sum f_i\right)}{f_mediana} \right) c$$

Donde:

L_i= frontera inferior de la clase de la mediana.
N= número de datos (frecuencia total).
$\left(\sum f_i\right)$= suma de frecuencia de las clases inferiores a la de la mediana
$f_{mediana}$= frecuencia de la clase de la mediana
c= anchura del intervalo de la clase de la mediana.

Geométricamente la mediana es el valor de X (abscisa) que corresponde a la recta vertical que divide un histograma en dos partes de igual área. Ese valor de X se suele denotar por $\tilde{X}$.

4.- Moda:

Hay 3 valores que se repiten en 4 ocasiones: el 1,21, el 1,22 y el 1,28, por lo tanto esta seria cuenta con 3 modas.

Es el valor que ocurre con mayor frecuencia; es decir, el valor más frecuente. La moda puede no existir, e incluso no ser única en caso de existir.

Ejemplo: El conjunto 2,2,5,7,9,9,9,10,10,11,12 y 18 tiene moda 9.
Ejemplo: El conjunto 3,5,8,10,12,15 y 16 no tiene moda.
Ejemplo: El conjunto 2,3,4,4,4,5,5,7,7,7 y 9 tiene dos modas, 4 y 7 y se llama bimodal.

Una distribución con moda única se dice *unimodal.*

En el caso de datos agrupados donde se haya construido una curva de frecuencias para ajustar los datos, la moda será el valor (o valores) de X correspondiente al máximo (o máximos) de la curva. Ese valor de X se denota por $\hat{X}$.

La moda puede deducirse de una distribución de frecuencia o de un histograma a partir de la fórmula:

$$Moda = L_1 + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2} \right) c$$

Donde:

L_1= frontera inferior de la clase modal (clase que contiene a la moda).
Δ_1 = exceso de la frecuencia modal sobre la de la clase inferior inmediata.
Δ_2 = exceso de la frecuencia modal sobre la de la clase superior inmediata.
c= anchura del intervalo de clase modal.

Medidas de posición no central

Medidas de posición no centrales

Cuartiles, Deciles y Percentiles:

Si un conjunto de datos está ordenado por magnitud, el valor central (o la media de los dos centrales) que divide al conjunto en dos mitades iguales, es la mediana. Extendiendo esa idea, podemos pensar en aquellos valores que dividen al conjunto de datos en cuatro partes iguales. Esos valores denotados Q_1, Q_2, y Q_3, se *llaman* ***primer cuartíl***, ***segundo cuartíl*** *y* ***tercer cuartíl***, *respectivamente.* EL Q_2 coincide con la mediana.

Análogamente, los valores que dividen a los datos en 10 partes iguales se llaman deciles, y se le denotan D_1, D_2,...,D_9, mientras que los valores que lo dividen en 100 partes iguales se llaman

percentiles, denotados por P_1, P_2,...,P_{99}. El 5º decil y el 50º percentil coinciden con l mediana. Los 25º y 75º percentiles coinciden con el primer y tercer cuartiles.

Colectivamente, cuartiles, deciles y percentiles se denominan cuartiles.

Las medidas de posición no centrales permiten conocer otros puntos característicos de la distribución que no son los valores centrales. Entre otros indicadores, se suelen utilizar una serie de valores que dividen la muestra en tramos iguales:

Cuartiles: son 3 valores que distribuyen la serie de datos, ordenada de forma creciente o decreciente, en cuatro tramos iguales, en los que cada uno de ellos concentra el 25% de los resultados.

Deciles: son 9 valores que distribuyen la serie de datos, ordenada de forma creciente o decreciente, en diez tramos iguales, en los que cada uno de ellos concentra el 10% de los resultados.

Percentiles: son 99 valores que distribuyen la serie de datos, ordenada de forma creciente o decreciente, en cien tramos iguales, en los que cada uno de ellos concentra el 1% de los resultados.

***Ejemplo**:* Vamos a calcular los cuartiles de la serie de datos referidos a la estatura de un grupo de alumnos (lección 2ª). Los deciles y centiles se calculan de igual manera, aunque haría falta distribuciones con mayor número de datos.

Variable	Frecuencias absolutas		Frecuencias relativas	
(Valor)	Simple	Acumulada	Simple	Acumulada
x	X	X	x	x
1,20	1	1	3,3%	3,3%
1,21	4	5	13,3%	16,6%
1,22	4	9	13,3%	30,0%
1,23	2	11	6,6%	36,6%
1,24	1	12	3,3%	40,0%
1,25	2	14	6,6%	46,6%
1,26	3	17	10,0%	56,6%
1,27	3	20	10,0%	66,6%
1,28	4	24	13,3%	80,0%
1,29	3	27	10,0%	90,0%
1,30	3	30	10,0%	100,0%

1º cuartil: es el valor 1,22 cm, ya que por debajo su valor del 25% de la frecuencia (tal como se puede ver en la columna de la frecuencia relativa acumulada).

2º cuartil: es el valor 1,26 cm, ya que entre este valor y el 1º cuartil se sitúa otro 25% de la frecuencia.

3º cuartil: es el valor 1,28 cm, ya que entre este valor y el 2º cuartil se sitúa otro 25% de la frecuencia. Además, por encima suya queda el restante 25% de la frecuencia.

Atención: cuando un cuartil recae en un valor que se ha repetido más de una vez (como ocurre en el ejemplo en los tres cuartiles) la medida de posición no central sería realmente una de las repeticiones.

TEMA 4ª
Medidas de Dispersión

Estudia la distribución de los valores de la serie, analizando si estos se encuentran más o menos concentrados, o más o menos dispersos.

Existen diversas ***medidas de dispersión***, entre las más utilizadas podemos destacar las siguientes:

1.- Rango: mide la amplitud de los valores de la muestra y se calcula por diferencia entre el valor más elevado y el valor más bajo.

2.- Varianza: Mide la distancia existente entre los valores de la serie y la media. Se calcula como sumatoria de las diferencias al cuadrado entre cada valor y la media, multiplicadas por el número de veces que se ha repetido cada valor. El sumatoria obtenido se divide por el tamaño de la muestra.

$$\sigma = \sqrt{\frac{\sum X^2 - \frac{(\sum X)^2}{N}}{N}}$$

La varianza siempre será mayor que cero. Mientras más se aproxima a cero, más concentrados están los valores de la serie alrededor de la media. Por el contrario, mientras mayor sea la varianza, más dispersos están.

3.- Desviación típica: Se calcula como raíz cuadrada de la varianza.

La desviación un elemento del conjunto es su diferencia con respecto a la medida; por ejemplo, en sucesión x_1, x_2, ..., x_n la desviación de x_1 es x_1-x, y el cuadrado de la desviación es $(x_1- x)^2$. La varianza es la medida del cuadro de las desviaciones. Por último, la desviación típica, representada por la letra griega sigma (σ), es la raíz cuadrada de la varianza, y se calcula de la siguiente manera:

$$\sigma = \sqrt{\frac{\sum X^2 - \frac{(\sum X)^2}{N}}{N}}$$

Si la desviación típica es pequeña, los datos están agrupados cerca de la media; si es grande, están muy dispersos.

Es una mediada de dispersión expresada por el espacio de la curva normal comprendido entre la media y +/- σ.

4.- Coeficiente de variación de Pearson: se calcula como cociente entre la desviación típica y la media.

Ejemplo: vamos a utilizar la serie de datos de la estatura de los alumnos de una clase (lección 2ª) y vamos a calcular sus medidas de dispersión.

Variable	Frecuencias absolutas		Frecuencias relativas	
(Valor)	Simple	Acumulada	Simple	Acumulada
1,20	1	1	3,3%	3,3%
1,21	4	5	13,3%	16,6%
1,22	4	9	13,3%	30,0%
1,23	2	11	6,6%	36,6%
1,24	1	12	3,3%	40,0%
1,25	2	14	6,6%	46,6%
1,26	3	17	10,0%	56,6%
1,27	3	20	10,0%	66,6%
1,28	4	24	13,3%	80,0%
1,29	3	27	10,0%	90,0%
1,30	3	30	10,0%	100,0%

1.- Rango: Diferencia entre el mayor valor de la muestra (1,30) y el menor valor (1,20). Luego el rango de esta muestra es 10 cm.

2.- Varianza: recordemos que la media de esta muestra es 1,253. Luego, aplicamos la fórmula:

$$C.V. = \frac{\sigma}{\mu} \qquad C.V. = 100 \cdot \frac{\sigma}{\mu}$$

Coeficiente de Variación = Desviación típica/Media

Por lo tanto, la varianza es 0,0010

3.- Desviación típica: es la raíz cuadrada de la varianza.

$$\delta = \sqrt{\sum (X_j - X)^2/N}$$

4.- Coeficiente de variación de Pearson: se calcula como cociente entre la desviación típica y la media de la muestra.

Cv = 0,0320 / 1,253

Luego,

Cv = 0,0255

El interés del coeficiente de variación es que al ser un porcentaje permite comparar el nivel de dispersión de dos muestras. Esto no ocurre con la desviación típica, ya que viene expresada en las mismas unidas que los datos de la serie.

Por ejemplo, para comparar el nivel de dispersión de una serie de datos de la altura de los alumnos de una clase y otra serie con el peso de dichos alumnos, no se puede utilizar las desviaciones típicas (una viene vienes expresada en cm y la otra en kg). En cambio, sus coeficientes de variación son ambos porcentajes, por lo que sí se pueden comparar.

LECCION 4b
Medidas de forma: Grado de concentración

Las ***medidas de forma*** permiten conocer que forma tiene la curva que representa la serie de datos de la muestra. En concreto, podemos estudiar las siguientes características de la curva:

a) Concentración: mide si los valores de la variable están más o menos uniformemente repartidos a lo largo de la muestra.

b) Asimetría: mide si la curva tiene una forma simétrica, es decir, si respecto al centro de la misma (centro de simetría) los segmentos de curva que quedan a derecha e izquierda son similares.

c) Curtosis: mide si los valores de la distribución están más o menos concentrados alrededor de los valores medios de la muestra.

a) Concentración

Para medir el nivel de concentración de una distribución de frecuencia se pueden utilizar distintos indicadores, entre ellos el **Índice de Gini**.

Este índice se calcula aplicando la siguiente fórmula:

$$IG = \frac{\Sigma\,(pi - qi)}{\Sigma\,pi}$$

(i toma valores entre 1 y n-1)

En donde pi mide el porcentaje de individuos de la muestra que presentan un valor igual o inferior al de xi.

$$p_i = \frac{n_1 + n_2 + n_3 + \ldots + n_i}{N} \times 100$$

Mientras que qi se calcula aplicando la siguiente fórmula:

$$q_i = \frac{(X_1 * n_1) + (X_2 * n_2) + \ldots + (X_i * n_i)}{(X_1 * n_1) + (X_2 * n_2) + \ldots + (X_n * n_n)} \times 100$$

El **Índice Gini** (IG) puede tomar valores entre 0 y 1:

IG = 0 : concentración mínima. La muestra está uniformemente repartida a lo largo de todo su rango.

IG = 1 : concentración máxima. Un sólo valor de la muestra acumula el 100% de los resultados.

Ejemplo*:* vamos a calcular el Índice Gini de una serie de datos con los sueldos de los empleados de una empresa (millones pesos).

Sueldos	**Empleados (Frecuencias absolutas)**		**Frecuencias relativas**	
(Millones)	**Simple**	**Acumulada**	**Simple**	**Acumulada**
3,5	10	10	25,0%	25,0%
4,5	12	22	30,0%	55,0%
6,0	8	30	20,0%	75,0%
8,0	5	35	12,5%	87,5%
10,0	3	38	7,5%	95,0%
15,0	1	39	2,5%	97,5%
20,0	1	40	2,5%	100,0%

Calculamos los valores que necesitamos para aplicar la fórmula del Índice de Gini:

X_i	n_i	Σn_i	p_i	$X_i * n_i$	$\Sigma X_i * n_i$	q_i	$p_i - q_i$
3,5	10	10	25,0	35,0	35,0	13,6	10,83
4,5	12	22	55,0	54,0	89,0	34,6	18,97
6,0	8	30	75,0	48,0	147,0	57,2	19,53
8,0	5	35	87,5	40,0	187,0	72,8	15,84
10,0	3	38	95,0	30,0	217,0	84,4	11,19
15,0	1	39	97,5	15,0	232,0	90,3	7,62
25,0	1	40	100,0	25,0	257,0	100,0	0
Σp_i (entre 1 y n-1) =			435,0	X	$\Sigma (p_i - q_i)$ (entre 1 y n-1) =		83,99

Por lo tanto:

IG = 83,99 / 435,0 = 0,19

Un Índice Gini de 0,19 indica que la muestra está bastante uniformemente repartida, es decir, su nivel de concentración no es excesivamente alto.

Ejemplo: Ahora vamos a analizar nuevamente la muestra anterior, pero considerando que hay más personal de la empresa que cobra el sueldo máximo, lo que conlleva mayor concentración de renta en unas pocas personas.

Sueldos	**Empleados (Frecuencias absolutas)**		**Frecuencias relativas**	
(Millones)	**Simple**	**Acumulada**	**Simple**	**Acumulada**
3,5	10	10	25,0%	25,0%
4,5	10	20	25,0%	50,0%
6,0	8	28	20,0%	70,0%
8,0	5	33	12,5%	82,5%
10,0	3	36	7,5%	90,0%
15,0	0	36	0,0%	90,0%
20,0	4	40	10,0%	100,0%

En este caso obtendríamos los siguientes datos:

X_i	n_i	Σn_i	p_i	$X_i * n_i$	$\Sigma X_i * n_i$	q_i	$p_i - q_i$
3,5	10	10	25,0	35	35	11,7	13,26
4,5	10	20	50,0	45	80	26,8	23,15
6,0	8	28	70,0	48	128	43,0	27,05
8,0	5	33	82,5	40	168	56,4	26,12

10,0	3	36	90,0	30	198	66,4	23,56
15,0	0	36	90,0	0	198	66,4	23,56
25,0	4	40	100,0	100	298	100,0	0,00
		Σ p_i (entre 1 y n-1) =	407,5	X	Σ (p_i - q_i) (entre 1 y n-1) =		136,69

El **Índice Gini** sería:

IG = 136,69 / 407,5 = 0,34

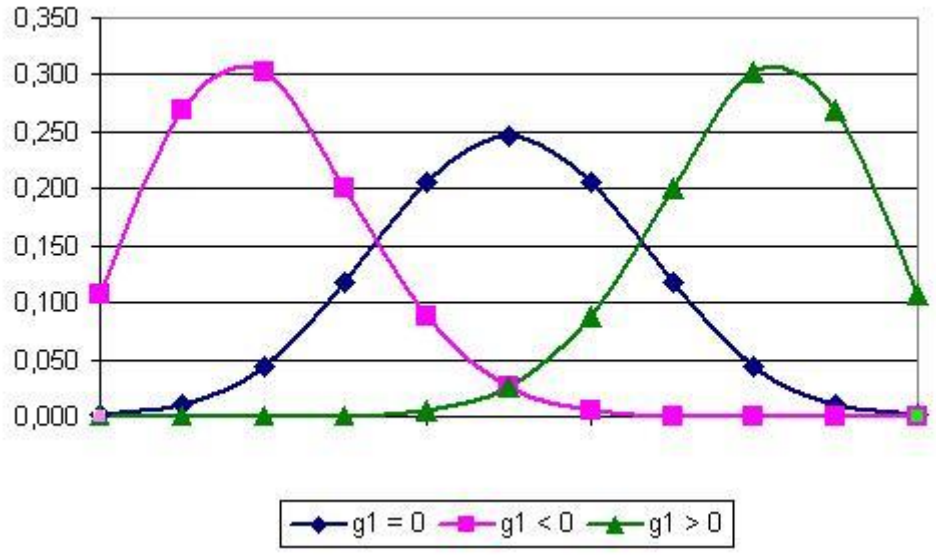
El Índice Gini se ha elevado considerablemente, reflejando la mayor concentración de rentas que hemos comentado.

LECCION 4c
Medidas de forma: Coeficiente de Asimetría

b) Asimetría

Hemos comentado que el concepto de asimetría se refiere a si la curva que forman los valores de la serie presenta la misma forma a izquierda y derecha de un valor central (media aritmética)

$$\mu_3 = E\left[(x-\mu)^3\right]$$



Para medir el nivel de asimetría se utiliza el llamado **Coeficiente de Asimetría de Fisher**, que viene definido:

Los resultados pueden ser los siguientes:

g1 = 0 (distribución simétrica; existe la misma concentración de valores a la derecha y a la izquierda de la media)

g1 > 0 (distribución asimétrica positiva; existe mayor concentración de valores a la derecha de la media que a su izquierda)

g1 < 0 (distribución asimétrica negativa; existe mayor concentración de valores a la izquierda de la media que a su derecha)

Ejemplo: Vamos a calcular el Coeficiente de Asimetría de Fisher de la serie de datos referidos a la estatura de un grupo de alumnos (lección 2ª):

Variable	Frecuencias absolutas		Frecuencias relativas	
(Valor)	Simple	Acumulada	Simple	Acumulada
x	x	x	x	x
1,20	1	1	3,3%	3,3%
1,21	4	5	13,3%	16,6%
1,22	4	9	13,3%	30,0%
1,23	2	11	6,6%	36,6%
1,24	1	12	3,3%	40,0%
1,25	2	14	6,6%	46,6%
1,26	3	17	10,0%	56,6%
1,27	3	20	10,0%	66,6%
1,28	4	24	13,3%	80,0%
1,29	3	27	10,0%	90,0%
1,30	3	30	10,0%	100,0%

Recordemos que la media de esta muestra es 1,253

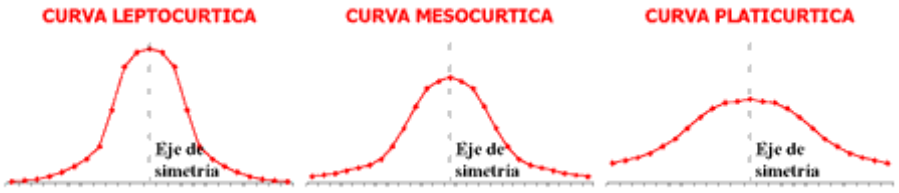
Σ((xi - x)^3)*ni	Σ((xi - x)^2)*ni
0,000110	0,030467

Luego:

$$g_1 = \frac{(1/30) * 0{,}000110}{(1/30) * (0{,}030467)^{(3/2)}} = -0{,}1586$$

Por lo tanto el **Coeficiente de Fisher de Simetría** de esta muestra es -0,1586, lo que quiere decir que presenta una distribución asimétrica negativa (se concentran más valores a la izquierda de la media que a su derecha).

TEMA 5°
Medidas de forma: Coeficiente de Curtosis



c) Curtosis

El **Coeficiente de Curtosis** analiza el grado de concentración que presentan los valores alrededor de la zona central de la distribución.

Se definen 3 tipos de distribuciones según su grado de curtosis:

Distribución mesocúrtica: presenta un grado de concentración medio alrededor de los valores centrales de la variable (el mismo que presenta una distribución normal).

Distribución leptocúrtica: presenta un elevado grado de concentración alrededor de los valores centrales de la variable.

Distribución platicúrtica: presenta un reducido grado de concentración alrededor de los valores centrales de la variable.

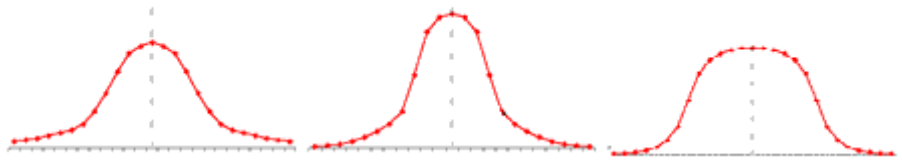
El **Coeficiente de Curtosis** viene definido por la siguiente fórmula:

$$\mu_4 = E\left[(x-\mu)^4\right]$$

El cuarto momento respecto de la media mide la curtosis de la distribución, es decir, la forma de la distribución de probabilidad. Al representar gráficamente variables con curtosis pequeña, platicúrticas, se observan curvas o histogramas con colas cortas y aspecto aplanado o en meseta; si la variable tiene curtosis grande, es decir, si es leptocúrtica, su gráfica ser alta y estilizada, con colas largas y pesadas.

La curtosis de una variable siempre es positiva y se mide en la unidades de la variable elevadas a potencia 4. Por tanto, nuevamente se nos plantean los problemas relacionados con las unidades de medida y las escalas y necesitamos una medida adimensional de la curtosis. Esta medida adimensional de la curtosis es el **coeficiente de curtosis**, g_2, que se calcula como el cociente entre el cuarto momento y el cuadrado de la varianza, al que se le resta 3 unidades. Esta corrección se debe a que, sin ella, las variables normales tendrían coeficiente de curtosis igual a 3; al restar 3 conseguimos que el coeficiente de curtosis de la variable normal sea 0 y que las variables platicúrticas tengan coeficiente de curtosis negativo y la leptocúrticas positivo, lo cual es más mnemotécnico que la distinción entre curtosis pequeña y grande.

$$g_2 = \frac{\mu_4}{\sigma^4} - 3$$



$g_2 = 0$ **$g_2 > 0$** **$g_2 < 0$**

El **Coeficiente de Curtosis** viene definido por la siguiente fórmula:

$$g_2 = \frac{(1/n) * \Sigma (x_i - xm)^4 * n_i}{((1/n) * \Sigma (x_i - xm)^2 * n_i))^2} - 3$$

Los resultados pueden ser los siguientes:

g2 = 0 (**distribución mesocúrtica)**.

g2 > 0 (distribución leptocúrtica).

g2 < 0 (distribución platicúrtica).

Ejemplo: Vamos a calcular el Coeficiente de Curtosis de la serie de datos referidos a la estatura de un grupo de alumnos (lección 2ª):

Variable	Frecuencias absolutas		Frecuencias relativas	
(Valor)	Simple	Acumulada	Simple	Acumulada
x	x	x	x	x
1,20	1	1	3,3%	3,3%
1,21	4	5	13,3%	16,6%
1,22	4	9	13,3%	30,0%
1,23	2	11	6,6%	36,6%
1,24	1	12	3,3%	40,0%

1,25	2	14	6,6%	46,6%
1,26	3	17	10,0%	56,6%
1,27	3	20	10,0%	66,6%
1,28	4	24	13,3%	80,0%
1,29	3	27	10,0%	90,0%
1,30	3	30	10,0%	100,0%

Recordemos que la media de esta muestra es 1,253

Σ((xi - xm)^4)*ni	**Σ((xi - xm)^2)*ni**
0,00004967	0,03046667

Luego:

$$g_2 = \frac{(1/30) * 0{,}00004967}{((1/30) * (0{,}03046667))^{\wedge}2} - 3 = -1{,}39$$

Por lo tanto, el **Coeficiente de Curtosis** de esta muestra es -1,39, lo que quiere decir que se trata de una distribución platicúrtica, es decir, con una reducida concentración alrededor de los valores centrales de la distribución.

TEMA 6ª
Combinaciones, Variaciones y Permutaciones (I)

Para aplicar la **Regla de Laplace**, el cálculo de los sucesos favorables y de los sucesos posibles a veces no plantea ningún problema, ya que son un número reducido y se pueden calcular con facilidad:

Por ejemplo: Probabilidad de que al lanzar un dado salga el número 2. Tan sólo hay un caso favorable, mientras que los casos posibles son seis.

Probabilidad de acertar al primer intento el horóscopo de una persona. Hay un caso favorable y 12 casos posibles.

Sin embargo, a veces calcular el número de casos favorables y casos posibles es complejo y hay que aplicar reglas matemáticas:

Por ejemplo: 5 matrimonios se sientan aleatoriamente a cenar y queremos calcular la probabilidad de que al menos los miembros de un matrimonio se sienten junto. En este caso, determinar el número de casos favorables y de casos posibles es complejo.

Las reglas matemáticas que nos pueden ayudar son el cálculo de **combinaciones**, el cálculo de **variaciones** y el cálculo de **permutaciones**.

a) Combinaciones:

Determina el número de subgrupos de 1, 2, 3, etc. elementos que se pueden formar con los "n" elementos de una nuestra. Cada subgrupo se diferencia del resto en los elementos que lo componen, sin que influya el orden.

Por ejemplo, calcular las posibles combinaciones de 2 elementos que se pueden formar con los números 1, 2 y 3.

Se pueden establecer 3 parejas diferentes: (1,2), (1,3) y (2,3). En el cálculo de combinaciones las parejas (1,2) y (2,1) se consideran idénticas, por lo que sólo se cuentan una vez.

b) Variaciones:

Calcula el número de subgrupos de 1, 2, 3, etc. elementos que se pueden establecer con los "n" elementos de una muestra. Cada subgrupo se diferencia del resto en los elementos que lo componen o en el orden de dichos elementos (es lo que le diferencia de las combinaciones).

Por ejemplo, calcular las posibles variaciones de 2 elementos que se pueden establecer con los números 1, 2 y 3.

Ahora tendríamos 6 posibles parejas: (1,2), (1,3), (2,1), (2,3), (3,1) y (3,3). En este caso los subgrupos (1,2) y (2,1) se consideran distintos.

c) Permutaciones:

Calcula las posibles agrupaciones que se pueden establecer con todos los elementos de un grupo, por lo tanto, lo que diferencia a cada subgrupo del resto es el orden de los elementos.

Por ejemplo, calcular las posibles formas en que se pueden ordenar los números 1, 2 y 3.

Hay 6 posibles agrupaciones: (1, 2, 3), (1, 3, 2), (2, 1, 3), (2, 3, 1), (3, 1, 2) y (3, 2, 1)

.

LECCION 6b
Combinaciones, Variaciones y Permutaciones (II)

¿Cómo se calculan?

a) Combinaciones:

Para calcular el número de combinaciones se aplica la siguiente fórmula:

$$"{}_{m}C_{,n}" = n!/n!(m-n)!$$

El término " **n !** " se denomina "factorial de n" y es la multiplicación de todos los números que van desde "n" hasta 1.

Por ejemplo: 4 ! = 4 * 3 * 2 * 1 = 24

La expresión **"mC,n"** representa las combinaciones de "m" elementos, formando subgrupos de "n" elementos.

Ejemplo: 10C,4 son las combinaciones de 10 elementos agrupándolos en subgrupos de 4 elementos:

$$"{}_{m}C_{,n}" = n!/n!(m-n)!$$

$${}_{10}C_{,4} = 10!/ 4!(10-4)! = 10.9.8.7.\cancel{6.5.4.3.2.1}/4.3.2.1(\cancel{6.5.4.3.2.1}) = 5040/24 = 210$$

Es decir, podríamos formar 210 subgrupos diferentes de 4 elementos, a partir de los 10 elementos.

b) Variaciones:

Para calcular el número de variaciones se aplica la siguiente fórmula:

$$"{}_{m}V_{,n}"$$

La expresión **"mV,n"** representa las variaciones de "m" elementos, formando subgrupos de "n" elementos. En este caso, como vimos en la lección anterior, un subgrupo se diferenciará del resto, bien por los elementos que lo forman, o bien por el orden de dichos elementos.

Ejemplo: 10V,4 son las variaciones de 10 elementos agrupándolos en subgrupos de 4 elementos:

$$"{}_{m}V_{,n}" \quad n!/(m-n)!$$

$${}_{10}C_{,4} = 10!/(10-4)! = 10.9.8.7.\cancel{6.5.4.3.2.1}/ (\cancel{6.5.4.3.2.1}) = 5040$$

Es decir, podríamos formar 5.040 subgrupos diferentes de 4 elementos, a partir de los 10 elementos.

c) Permutaciones:

Para calcular el número de permutaciones se aplica la siguiente fórmula:

$$\text{"}P_m\text{"}$$

La expresión **"Pm"** representa las permutaciones de "m" elementos, tomando todos los elementos. Los subgrupos se diferenciaran únicamente por el orden de los elementos.

Ejemplo: P10 son las permutaciones de 10 elementos:

$$\text{"}P_m\text{"} = 10! = 3.628.800,oo$$

Es decir, tendríamos 3.628.800 formas diferentes de agrupar 10 elementos.

LECCION 6c
Combinaciones, Variaciones y Permutaciones (III)

Vamos a analizar ahora que ocurriría con el cálculo de las combinaciones, de las variaciones o de las permutaciones en el **supuesto** de que al formar los subgrupos **los elementos pudieran repetirse**.

Por ejemplo: tenemos bolas de 6 colores diferentes y queremos formar subgrupos en los que pudiera darse el caso de que 2, 3, 4 o todas las bolas del subgrupo tuvieran el mismo color. En este caso no podríamos utilizar las fórmulas que vimos en la lección anterior.

a) Combinaciones con repetición:

Para calcular el número de combinaciones con repetición se aplica la siguiente fórmula:

Ejemplo: 10C',4 son las combinaciones de 10 elementos con repetición, agrupándolos en subgrupos de 4, en los que 2, 3 o los 4 elementos podrían estar repetidos:

$$nCr = (n-1)(n-2)….(n+1r)/r$$

$$= n! / r!(n-r)!$$

Es decir, podríamos formar 715 subgrupos diferentes de 4 elementos.

b) Variaciones con repetición:

Para calcular el número de variaciones con repetición se aplica la siguiente fórmula:

$$n!/ (n-r)!$$

Ejemplo: 10V',4 son las variaciones de 10 elementos con repetición, agrupándolos en subgrupos de 4 elementos:

$$n!/ (n-r)!$$

Es decir, podríamos formar 10.000 subgrupos diferentes de 4 elementos.

c) Permutaciones con repetición:

Para calcular el número de permutaciones con repetición se aplica la siguiente fórmula:

$$n!/ (n-r-1)!… n!/ (n-r-k)!$$

Son permutaciones de "m" elementos, en los que uno de ellos se repite " x1 " veces, otro " x2 " veces y así ... hasta uno que se repite " xk " veces.

Ejemplo: Calcular las permutaciones de 10 elementos, en los que uno de ellos se repite en 2 ocasiones y otro se repite en 3 ocasiones:

$$10!/ (10-1)! +10!/ (10-2)!$$

Es decir, tendríamos 302,400 formas diferentes de agrupar estos 10 elementos.

TEMA 7ª

Teoría Elemental de Probabilidad: Introducción

La **probabilidad** mide la frecuencia con la que aparece un resultado determinado cuando se realiza un experimento.

Ejemplo: tiramos un dado al aire y queremos saber cual es la probabilidad de que salga un 2, o que salga un número par, o que salga un número menor que 4.

El experimento tiene que ser aleatorio, es decir, que pueden presentarse diversos resultados, dentro de un conjunto posible de soluciones, y esto aún realizando el experimento en las mismas condiciones. Por lo tanto, a priori no se conoce cual de los resultados se va a presentar:

Ejemplos: lanzamos una moneda al aire: el resultado puede ser cara o cruz, pero no sabemos de antemano cual de ellos va a salir.

En la Lotería de Navidad, el "Gordo" (se llama "Gordo" al primer premio) puede ser cualquier número entre el 1 y el 100.000, pero no sabemos a priori cual va a ser (si lo supiéramos no estaríamos aquí escribiendo esta lección).

Hay experimentos que no son aleatorios y por lo tanto no se les puede aplicar las reglas de la probabilidad.

Ejemplo: en lugar de tirar la moneda al aire, directamente seleccionamos la cara. Aquí no podemos hablar de probabilidades, sino que ha sido un resultado determinado por uno mismo.

Antes de calcular las probabilidades de un experimento aleatorio hay que definir una serie de conceptos:

Suceso elemental: hace referencia a cada una de las posibles soluciones que se pueden presentar.

Ejemplo: al lanzar una moneda al aire, los sucesos elementales son la cara y la cruz. Al lanzar un dado, los sucesos elementales son el 1, el 2, .., hasta el 6.

Suceso compuesto: es un subconjunto de sucesos elementales.

Ejemplo: lanzamos un dado y queremos que salga un número par. El suceso "numero par" es un suceso compuesto, integrado por 3 sucesos elementales: el 2, el 4 y el 6

O, por ejemplo, jugamos a la ruleta y queremos que salga "menor o igual que 18". Este es un suceso compuesto formado por 18 sucesos elementales (todos los números que van del 1 al 18).

Al conjunto de todos los posibles sucesos elementales lo denominamos **espacio muestral**. Cada experimento aleatorio tiene definido su espacio muestral (es decir, un conjunto con todas las soluciones posibles).

Ejemplo: si tiramos una moneda al aíre una sola vez, el espacio muestral será cara o cruz.

Si el experimento consiste en lanzar una moneda al aire dos veces, entonces el espacio muestral estaría formado por (cara-cara), (cara-cruz), (cruz-cara) y (cruz-cruz).

LECCION 7b
Probabilidad: Relación entre sucesos

Entre los sucesos compuestos se pueden establecer distintas relaciones:

a) Un suceso puede estar contenido en otro: las posibles soluciones del primer suceso también lo son del segundo, pero este segundo suceso tiene además otras soluciones suyas propias.

Ejemplo: lanzamos un dado y analizamos dos sucesos: a) que salga el número 6, y b) que salga un número par. Vemos que el suceso a) está contenido en el suceso b).

Siempre que se da el suceso a) se da el suceso b), pero no al contrario. Por ejemplo, si el resultado fuera el 2, se cumpliría el suceso b), pero no el a).

b) Dos sucesos pueden ser iguales: esto ocurre cuando siempre que se cumple uno de ellos se cumple obligatoriamente el otro y viceversa.

Ejemplo: lanzamos un dado al aire y analizamos dos sucesos: a) que salga número par, y b) que salga múltiplo de 2. Vemos que las soluciones coinciden en ambos casos.

c) Unión de dos o más sucesos: la unión será otro suceso formado por todos los elementos de los sucesos que se unen.

Ejemplo: lanzamos un dado al aire y analizamos dos sucesos: a) que salga número par y b) que el resultado sea mayor que 3. El suceso unión estaría formado por los siguientes resultados: el 2, el 4, el 5 y el 6

d) Intersección de sucesos: es aquel suceso compuesto por los elementos comunes de dos o más sucesos que se intersectan.

Ejemplo: lanzamos un dado al aire, y analizamos dos sucesos: a) que salga número par, y b) que sea mayor que 4. La intersección de estos dos sucesos tiene un sólo elemento, el número 6 (es el único resultado común a ambos sucesos: es mayor que 4 y es número par).

e) Sucesos incompatibles: son aquellos que no se pueden dar al mismo tiempo ya que no tienen elementos comunes (su intersección es el conjunto vació).

Ejemplo: lanzamos un dado al aire y analizamos dos sucesos: a) que salga un número menor que 3, y b) que salga el número 6. Es evidente que ambos no se pueden dar al mismo tiempo.

f) Sucesos complementarios: son aquellos que si no se da uno, obligatoriamente se tiene que dar el otro.

Ejemplo: lanzamos un dado al aire y analizamos dos sucesos: a) que salga un número par, y b) que salga un número impar. Vemos que si no se da el primero se tiene que dar el segundo (y viceversa).

LECCION 7c
Cálculo de probabilidades

Probabilidad

Como hemos comentado anteriormente, la probabilidad mide la mayor o menor posibilidad de que se dé un determinado resultado (suceso) cuando se realiza un experimento aleatorio.

La probabilidad toma valores entre 0 y 1 (o expresados en tanto por ciento, entre 0% y 100%):

El valor cero corresponde al suceso imposible: lanzamos un dado al aire y la probabilidad de que salga el número 7 es cero (al menos, si es un dado certificado por la OMD, "Organización Mundial de Dados").

El valor uno corresponde al suceso seguro: lanzamos un dado al aire y la probabilidad de que salga cualquier número del 1 al 6 es igual a uno (100%).

El resto de sucesos tendrá probabilidades entre cero y uno: que será tanto mayor cuanto más probable sea que dicho suceso tenga lugar.

¿Cómo se mide la probabilidad?

Uno de los métodos más utilizados es aplicando la **Regla de Laplace:** define la probabilidad de un suceso como el cociente entre casos favorables y casos posibles.

P(A) = Casos favorables / casos posibles

Veamos algunos **ejemplos**:

a) Probabilidad de que al lanzar un dado salga el número 2: el caso favorable es tan sólo uno (que salga el dos), mientras que los casos posibles son seis (puede salir cualquier número del uno al seis). Por lo tanto:

P(A) = 1 / 6 = 0,166 (o lo que es lo mismo, 16,6%)

b) Probabilidad de que al lanzar un dado salga un número par: en este caso los casos favorables son tres (que salga el dos, el cuatro o el seis), mientras que los casos posibles siguen siendo seis. Por lo tanto:

$$P(A) = 3 / 6 = 0{,}50 \text{ (o lo que es lo mismo, 50\%)}$$

c) Probabilidad de que al lanzar un dado salga un número menor que 5: en este caso tenemos cuatro casos favorables (que salga el uno, el dos, el tres o el cuatro), frente a los seis casos posibles. Por lo tanto:

$$P(A) = 4 / 6 = 0{,}666 \text{ (o lo que es lo mismo, 66,6\%)}$$

d) Probabilidad de que nos toque el "Lotería" de Navidad: tan sólo un caso favorable, el número que jugamos (¡qué triste...¡), frente a 100.000 casos posibles. Por lo tanto:

$$P(A) = 1 / 100.000 = 0{,}00001 \text{ (o lo que es lo mismo, 0,001\%)}$$

Merece la pena ..... Por cierto, tiene la misma probabilidad el número 45.264, que el número 00001, pero ¿cuál de los dos comprarías?

Para poder aplicar la **Regla de Laplace** el experimento aleatorio tiene que cumplir **dos requisitos:**

a) El número de resultados posibles (sucesos) tiene que ser finito. Si hubiera infinitos resultados, al aplicar la regla "casos favorables / casos posibles" el cociente siempre sería cero.

b) Todos los sucesos tienen que tener la misma probabilidad. Si al lanzar un dado, algunas caras tuvieran mayor probabilidad de salir que otras, no podríamos aplicar esta regla.

A la regla de Laplace también se le denomina **"probabilidad a priori"**, ya que para aplicarla hay que conocer antes de realizar el experimento cuales son los posibles resultados y saber que todos tienen las mismas probabilidades.

¿Y si el experimento aleatorio no cumple los dos requisitos indicados, qué hacemos?, ¿ponemos una denuncia?

No, no va a ser necesario denunciar a nadie, ya que en este caso podemos acudir a otro modelo de cálculo de probabilidades que se basa en la experiencia **(modelo frecuentista)**:

Cuando se realiza un experimento aleatorio un número muy elevado de veces, las probabilidades de los diversos posibles sucesos empiezan a converger hacia valores determinados, que son sus respectivas probabilidades.

Ejemplo: si lanzo una vez una moneda al aire y sale "cara", quiere decir que el suceso "cara" ha aparecido el 100% de las veces y el suceso "cruz" el 0%.

Si lanzo diez veces la moneda al aire, es posible que el suceso "cara" salga 7 veces y el suceso "cruz" las 3 restantes. En este caso, la probabilidad del suceso "cara" ya no sería del 100%, sino que se habría reducido al 70%.

Si repito este experimento un número elevado de veces, lo normal es que las probabilidades de los sucesos "cara" y "cruz" se vayan aproximando al 50% cada una. Este 50% será la probabilidad de estos sucesos según el modelo frecuentista.

En este modelo ya no será necesario que el número de soluciones sea finito, ni que todos los sucesos tengan la misma probabilidad.

Ejemplo: si la moneda que utilizamos en el ejemplo anterior fuera defectuosa (o estuviera trucada), es posible que al repetir dicho experimento un número elevado de veces, la "cara" saliera con una frecuencia, por ejemplo, del 65% y la "cruz" del 35%. Estos valores serían las probabilidades de estos dos sucesos según el modelo frecuentista.

A esta definición de la probabilidad se le denomina **probabilidad a posteriori**, ya que tan sólo repitiendo un experimento un número elevado de veces podremos saber cual es la probabilidad de cada suceso.

LECCION 7d
Probabilidad de sucesos

Probabilidad de sucesos

Al definir los sucesos hablamos de las diferentes relaciones que pueden guardar dos sucesos entre sí, así como de las posibles relaciones que se pueden establecer entre los mismos. Vamos a ver ahora cómo se refleja esto en el cálculo de probabilidades.

a) Un suceso puede estar contenido en otro: entonces, la probabilidad del primer suceso será menor que la del suceso que lo contiene.

Ejemplo: lanzamos un dado y analizamos dos sucesos: a) que salga el número 6, y b) que salga un número par. Dijimos que el suceso a) está contenido en el suceso b).

$$P(A) = 1/6 = 0{,}166$$

$$P(B) = 3 / 6 = 0{,}50$$

Por lo tanto, podemos ver que la probabilidad del suceso contenido, suceso a), es menor que la probabilidad del suceso que lo contiene, suceso b).

b) Dos sucesos pueden ser iguales: en este caso, las probabilidades de ambos sucesos son las mismas.

Ejemplo: lanzamos un dado al aire y analizamos dos sucesos: a) que salga número par, y b) que salga múltiplo de 2. Las soluciones coinciden en ambos casos.

$$P(A) = 3 / 6 = 0{,}50$$

$$P(B) = 3 / 6 = 0,50$$

c) Intersección de sucesos: es aquel suceso compuesto por los elementos comunes de los dos o más sucesos que se intersectan. La probabilidad será igual a la probabilidad de los elementos comunes.

Ejemplo: lanzamos un dado al aire y analizamos dos sucesos: a) que salga número par, y b) que sea mayor que 3. La intersección de estos dos sucesos tiene dos elementos: el 4 y el 6.

Su probabilidad será por tanto:

$$P(A \wedge B) = 2 / 6 = 0,33$$

d) Unión de dos o más sucesos: la probabilidad de la unión de dos sucesos es igual a la suma de las probabilidades individuales de los dos sucesos que se unen, menos la probabilidad del suceso intersección

Ejemplo: lanzamos un dado al aire y analizamos dos sucesos: a) que salga número par, y b) que el resultado sea mayor que 3. El suceso unión estaría formado por los siguientes resultados: el 2, el 4, el 5 y el 6.

$$P(A) = 3 / 6 = 0,50$$

$$P(B) = 3 / 6 = 0,50$$

$$P\ (A \wedge B) = 2 / 6 = 0,33$$

Por lo tanto,

$$P\ (A \text{ u } B) = (0,50 + 0,50) - 0,33 = 0,666$$

e) Sucesos incompatibles: la probabilidad de la unión de dos sucesos incompatibles será igual a la suma de las probabilidades de cada uno de los sucesos (ya que su intersección es el conjunto vacío y por lo tanto no hay que restarle nada).

Ejemplo: lanzamos un dado al aire y analizamos dos sucesos: a) que salga un número menor que 3, y b) que salga el número 6.

La probabilidad del suceso unión de estos dos sucesos será igual a:

$$P(A) = 2 / 6 = 0,333$$

$$P(B) = 1 / 6 = 0,166$$

Por lo tanto,

$$P(A \text{ u } B) = 0,33 + 0,166 = 0,50$$

f) Sucesos complementarios: la probabilidad de un suceso complementario a un suceso (A) es igual a 1 - P(A)

Ejemplo: lanzamos un dado al aire. el suceso (A) es que salga un número par, luego su complementario, suceso (B), es que salga un número impar.
La probabilidad del suceso (A) es igual a :

$$P(A) = 3 / 6 = 0,50$$

Luego, la probabilidad del suceso (B) es igual a:

$$P(B) = 1 - P(A) = 1 - 0,50 = 0,50$$

Se puede comprobar aplicando la regla de "casos favorables / casos posibles":

$$P(B) = 3 / 6 = 0,50$$

g) Unión de sucesos complementarios: la probabilidad de la unión de dos sucesos complementarios es igual a 1.

Ejemplo: seguimos con el ejemplo anterior: a) que salga un número par, y b) que salga un número impar. La probabilidad del suceso unión de estos dos sucesos será igual a:

$$P(A) = 3 / 6 = 0,50$$

$$P(B) = 3 / 6 = 0,50$$

Por lo tanto,

$$P(A \cup B) = 0,50 + 0,50 = 1$$

LECCION 21ª
Ejercicios

1.- Ejercicio

Calcular la probabilidad de acertar los 14 signos de la lotería:

Solución:

Se aplica la **Regla de Laplace** (casos favorables / casos posibles). El **caso favorable** es tan sólo uno (acertar los 14 signos). Los casos posibles se calculan como variaciones con repetición de 3 elementos (1, X y 2), tomados de 14 en 14 (los signos que hay que rellenar).

Son variaciones y no combinaciones ya que el orden influye: no es lo mismo (1,1,X) que (1, X, 1). Y son con repetición, ya que cualquiera de los signos (1, X y 2) se puede repetir hasta 14 veces.

Por lo tanto, los **casos posibles** son:

Y la probabilidad de acertar los 14 resultados es:

No demasiado elevada....pero el que la sigue la consigue.

2.- Ejercicio

Y la probabilidad de acertar 12 signos de la lotería:

Solución:

Aplicamos nuevamente la **Regla de Laplace**. En este caso los **casos favorables** se calculan como combinaciones de 14 elementos tomados de 2 en 2, de esta manera obtenemos todas las posibles alternativas de fallar 2 resultados de 14 (lo que equivale a acertar 12 resultados). Utilizamos combinaciones y no variaciones ya que el orden no importa (da lo mismo fallar el 3º y el 6º, que el 6º y el 3º)

Los **casos posibles** siguen siendo los mismos:

Por lo que la probabilidad de acertar 12 resultados es:

Por lo tanto, tenemos más probabilidades de acertar 12 resultados que 14 (¿será por eso por lo que pagan menos?).

3.- Ejercicio

Calcular la probabilidad de, en una carrera de 12 motos, acertar los 3 que quedan primeros (sin importar cual de ellos queda primero, cual segundo y cual tercero).

Solución:

Se aplica la **Regla de Laplace**. El **caso favorable** es tan sólo uno: las 3 motos que entran en primer lugar. Los **casos posibles** se calculan como combinaciones de 12 elementos tomados de 3 en 3 (es decir, determinamos todas las posibles alternativas de 3 motos que pueden entrar en las 3 primeras posiciones). Como el orden de estas 3 primeras motos no importa, utilizamos combinaciones en lugar de variaciones.

Por lo tanto, los casos posibles son:

Por lo que la probabilidad de acertar los 3 caballos ganadores es:

Algo mayor que en las loterías.... Eso sí, se paga menos.

4.- Ejercicio

Y si hubiera que acertar, no sólo los 3 caballos que ganan, sino el orden de su entrada en meta.

Solución:

El **caso favorable** sigue siendo uno: los 3 caballos que entran en primer lugar, colocados en su orden correspondiente.

Los **casos posibles** se calculan ahora como variaciones (ya que el orden influye) de 12 elementos tomados de 3 en 3 (calculamos todas las posibles maneras en que los 12 caballos podrían ocupar las 3 primeras posiciones.

Por lo que la probabilidad de acertar los 3 caballos ganadores es:

Menor que en el ejemplo 3º. Ya no vale acertar que 3 caballos entran en primer lugar, sino que tenemos que acertar el orden de su entrada.

LECCION 7e
Probabilidad condicionada

Las **probabilidades condicionadas** se calculan una vez que se ha incorporado información adicional a la situación de partida:

Ejemplo: se tira un dado y sabemos que la probabilidad de que salga un 2 es 1/6 (probabilidad a priori). Si incorporamos nueva información (por ejemplo, alguien nos dice que el resultado ha sido un número par) entonces la probabilidad de que el resultado sea el 2 ya no es 1/6.

Las probabilidades condicionadas se calculan aplicando la siguiente **fórmula**:

Donde:

P (B/A) es la probabilidad de que se de el suceso B condicionada a que se haya dado el suceso A.

P (B Λ A) es la probabilidad del suceso simultáneo de A y de B

P (A) es la probabilidad a priori del suceso A

En el **ejemplo** que hemos visto:

P (B/A) es la probabilidad de que salga el número 2 (suceso B) condicionada a que haya salido un número par (suceso A).

P (B Λ A) es la probabilidad de que salga el dos y número par.

P (A) es la probabilidad a priori de que salga un número par.

Por lo tanto:

$$P (B \Lambda A) = 1/6$$

$$P (A) = 1/2$$

$$P (B/A) = (1/6) / (1/2) = 1/3$$

Luego, la probabilidad de que salga el número 2, si ya sabemos que ha salido un número par, es de 1/3 (mayor que su probabilidad a priori de 1/6).

2º ejemplo:

En un estudio sanitario se ha llegado a la conclusión de que la probabilidad de que una persona sufra problemas coronarios (suceso B) es el 0,10 (probabilidad a priori).

Además, la probabilidad de que una persona sufra problemas de obesidad (suceso A) es el 0,25 y la probabilidad de que una persona sufra a la vez problemas de obesidad y coronarios (suceso intersección de A y B) es del 0,05.

Calcular la probabilidad de que una persona sufra problemas coronarios si está obesa (probabilidad condicionada P(B/A)).

$$P (B \Lambda A) = 0,05$$

$$P (A) = 0,25$$

$$P (B/A) = 0,05 / 0,25 = 0,20$$

Por lo tanto, la probabilidad condicionada es superior a la probabilidad a priori. No siempre esto es así, a veces la probabilidad condicionada es igual a la probabilidad a priori o menor.

Por ejemplo: probabilidad de que al tirar un dado salga el número 2, condicionada a que haya salido un número impar.

La probabilidad condicionada es en este caso cero, frente a una probabilidad a priori de 1/6.

LECCION 7f
Probabilidad compuesta

La **probabilidad compuesta** (o regla de multiplicación de probabilidades) se deriva de la probabilidad condicionada:

La probabilidad de que se den simultáneamente dos sucesos (suceso intersección de A y B) es igual a la probabilidad a priori del suceso A multiplicada por la probabilidad del suceso B condicionada al cumplimiento del suceso A.

La **fórmula** para calcular esta probabilidad compuesta es:

Ejemplo 1º : Estudiamos el suceso A (porcentaje de varones mayores de 40 años casados) y el suceso B (varones mayores de 40 años con más de 2 hijos) y obtenemos la siguiente información:

Un 35% de los varones mayores de 40 años están casados.

De los varones mayores de 40 años y casados, un 30% tienen más de 2 hijos (suceso B condicionado al suceso A).

Calcular la probabilidad de que un varón mayor de 40 años esté casado y tenga más de 2 hijos (suceso intersección de A y B).

Por lo tanto:

$$P(A) = 0{,}35$$

$$P(B/A) = 0{,}30$$

$$P(A \wedge B) = 0{,}35 * 0{,}30 = 0{,}105$$

Es decir, un 10,5% de los varones mayores de 40 años están casados y tienen más de 2 hijos.

2º ejemplo: Estudiamos el suceso A (alumnos que hablan inglés) y el suceso B (alumnos que hablan francés) y obtenemos la siguiente información:

Un 50% de los alumnos hablan inglés.

De los alumnos que hablan inglés, un 20% hablan también francés (suceso B condicionado al suceso A).

Calcular la probabilidad de que un alumno hable inglés y francés (suceso intersección de A y B).

Por lo tanto:

$$P(A) = 0{,}50$$

$$P(B/A) = 0{,}20$$

$$P(A \wedge B) = 0{,}50 * 0{,}20 = 0{,}10$$

Es decir, un 10% de los alumnos hablan inglés y francés.

LECCION 7g
Teorema de la probabilidad total

El **Teorema de la probabilidad total** nos permite calcular la probabilidad de un suceso a partir de probabilidades condicionadas:

Ejemplo: supongamos que si llueve la probabilidad de que ocurra un accidentes es x% y si hace buen tiempo dicha probabilidad es y%. Este teorema nos permite deducir cuál es la probabilidad de que ocurra un accidente si conocemos la probabilidad de que llueva y la probabilidad de que haga buen tiempo.

La **fórmula** para calcular esta probabilidad es:

Es decir, **la probabilidad de que ocurra el suceso B** (en nuestro ejemplo, que ocurra un accidente) **es igual a la suma de multiplicar cada una de las probabilidades condicionadas** de este suceso con los diferentes sucesos A (probabilidad de un accidente cuando llueve y cuando hace buen tiempo) **por la probabilidad de cada suceso A.**

Para que este teorema se pueda aplicar hace falta cumplir **un requisito**:

Los sucesos A tienen que formar un sistema completo, es decir, que contemplen todas las posibilidades (la suma de sus probabilidades debe ser el 100%).

Ejemplo: al tirar una moneda, el suceso "salir cara" y el suceso "salir cruz" forman un sistema completo, no hay más alternativas: la suma de sus probabilidades es el 100%

Ejemplo: al tirar un dado, que salga el 1, el 2, el 3, o el 4 no forman un sistema completo, ya que no contempla todas las opciones (podría salir el 5 o el 6). En este caso no se podría aplicar el teorema de la probabilidad total.

Ejercicio 1º: En un saquito hay papeletas de tres colores, con las siguientes probabilidades de ser elegidas:

a) Amarilla: probabilidad del 50%.

b) Verde: probabilidad del 30%

c) Roja: probabilidad del 20%.

Según el color de la papeleta elegida, podrás participar en diferentes sorteos. Así, si la papeleta elegida es:

a) Amarilla: participas en un sorteo con una probabilidad de ganar del 40%.

b) Verde: participas en otro sorteo con una probabilidad de ganar del 60%

c) Roja: participas en un tercer sorteo con una probabilidad de ganar del 80%.

Con esta información, **¿qué probabilidad tienes de ganar el sorteo en el que participes?:**

1.- Las tres papeletas forman un sistema completo: sus probabilidades suman 100%

2.- Aplicamos la fórmula:

Luego,

$$P(B) = (0{,}50 * 0{,}40) + (0{,}30 * 0{,}60) + (0{,}20 * 0{,}80) = 0{,}54$$

Por tanto, la probabilidad de que ganes el sorteo es del 54%.

Ejercicio 2º: Van a cambiar a tu jefe y se barajan diversos candidatos:

a) Carlos, con una probabilidad del 60%

b) Juan, con una probabilidad del 30%

c) Luis, con una probabilidad del 10%

En función de quien sea tu próximo jefe, la probabilidad de que te suban el sueldo es la siguiente:

a) Si sale Carlos: la probabilidad de que te suban el sueldo es del 5%.

b) Si sale Juan: la probabilidad de que te suban el sueldo es del 20%.

c) Si sale Luis: la probabilidad de que te suban el sueldo es del 60%.

En definitiva, **¿cual es la probabilidad de que te suban el sueldo?**:

1.- Los tres candidatos forman un sistema completo

2.- Aplicamos la fórmula:

$$P(B) = (0{,}60 * 0{,}05) + (0{,}30 * 0{,}20) + (0{,}10 * 0{,}60) = 0{,}15$$

Por tanto, la probabilidad de que te suban el sueldo es del 15%. Lo llevas claro amigo...

TEMA 8
Distribuciones bidimensionales

Las distribuciones bidimensionales son aquellas en las que se estudian al mismo tiempo dos variables de cada elemento de la población: por ejemplo: peso y altura de un grupo de estudiantes; superficie y precio de las viviendas de una ciudad; potencia y velocidad de una gama de coches deportivos.

Para representar los datos obtenidos se utiliza una **tabla de correlación**:

X / Y	y_1	y_2	.....	y_{m-1}	y_m
x_1	$n_{1,1}$	$n_{1,2}$	X	$n_{1,m-1}$	$n_{1,m}$
x_2	$n_{2,1}$	$n_{2,2}$	X	$n_{2,m-1}$	$n_{2,m}$
.....	x	X	X	x	x
x_{n-1}	$n_{n-1,1}$	$n_{n-1,2}$	X	$n_{n-1,m-1}$	$n_{n-1,m}$
x_n	$n_{n,1}$	$n_{n,2}$	X	$n_{n,m-1}$	$n_{n,m}$

Las "x" representan una de las variables y las "y" la otra variable. En cada intersección de una valor de "x" y un valor de "y" se recoge el número de veces que dicho par de valores se ha presentado conjuntamente.

Ejemplo: Medimos el peso y la estatura de los alumnos de una clase y obtenemos los siguientes resultados:

Alumno	Estatura	Peso	Alumno	Estatura	Peso	Alumno	Estatura	Peso
Alumno 1	1,25	32	Alumno 11	1,25	31	Alumno 21	1,25	33
Alumno 2	1,28	33	Alumno 12	1,28	35	Alumno 22	1,28	32
Alumno 3	1,27	31	Alumno 13	1,27	34	Alumno 23	1,27	34
Alumno 4	1,21	34	Alumno 14	1,21	33	Alumno 24	1,21	34
Alumno 5	1,22	32	Alumno 15	1,22	33	Alumno 25	1,22	35
Alumno 6	1,29	31	Alumno 16	1,29	31	Alumno 26	1,29	31
Alumno 7	1,30	34	Alumno 17	1,30	35	Alumno 27	1,30	34
Alumno 8	1,24	32	Alumno 18	1,24	32	Alumno 28	1,24	33
Alumno 9	1,27	32	Alumno 19	1,27	31	Alumno 29	1,27	35
Alumno 10	1,29	35	Alumno 20	1,29	33	Alumno 30	1,29	34

Esta información se puede representar de un modo más organizado en la siguiente tabla de correlación:

Estatura / Peso	31 kg	32 kg	33 kg	34 kg	35 kg
1,21 cm	0	0	1	2	0
1,22 cm	0	1	1	0	1
1,23 cm	0	0	0	0	0
1,24 cm	0	2	1	0	0
1,25 cm	1	1	1	0	0
1,26 cm	0	0	0	0	0
1,27 cm	2	1	0	2	1
1,28 cm	0	1	1	0	1
1,29 cm	3	0	1	1	1
1,30 cm	0	0	0	2	1

Tal como se puede ver, en cada casilla se recoge el número de veces que se presenta conjuntamente cada par de valores (x,y).

Tal como vimos en las distribuciones unidimensionales si una de las variables (o las dos) presentan gran número de valores diferentes, y cada uno de ellos se repite en muy pocas ocasiones, puede convenir agrupar los valores de dicha variable (o de las dos) en tramos.

LECCION 8b
Distribuciones marginales

Al analizar una distribución bidimensional, uno puede centrar su estudio en el comportamiento de una de las variables, con independencia de como se comporta la otra. Estaríamos así en el análisis de una **distribución marginal**.

De cada distribución bidimensional se pueden deducir **dos distribuciones marginales**: una correspondiente a la variable x, y otra correspondiente a la variable y.

Distribución marginal de X

X	$n_{i.}$
x_1	$n_{1.}$
x_2	$n_{2.}$

.....	...
x_{n-1}	$n_{n-1.}$
X_n	$n_{n.}$

Distribución marginal de Y

Y	**$n_{.j}$**
y_1	$n_{.1}$
y_2	$n_{.2}$
.....	...
y_{m-1}	$n_{.m-1}$
Y_m	$n_{.m}$

Ejemplo: a partir del ejemplo que vimos en la lección anterior (serie con los pesos y medidas de los alumnos de una clase) vamos a estudiar sus distribuciones marginales.

Estatura / Peso	31 kg	32 kg	33 kg	34 kg	35 kg
1,21 cm	0	0	1	2	0
1,22 cm	0	1	1	0	1
1,23 cm	0	0	0	0	0
1,24 cm	0	2	1	0	0
1,25 cm	1	1	1	0	0
1,26 cm	0	0	0	0	0
1,27 cm	2	1	0	2	1
1,28 cm	0	1	1	0	1
1,29 cm	3	0	1	1	1
1,30 cm	0	0	0	2	1

Las variables marginales se comportan como variables unidimensionales, por lo que pueden ser representadas en tablas de frecuencias.

a) Distribución marginal de la variable X (estatura)

Obtenemos la siguiente tabla de frecuencia:

Variable	Frecuencias absolutas		Frecuencias relativas	
(Estatura)	Simple	Acumulada	Simple	Acumulada
1,21	3	3	10,0%	10,0%
1,22	3	6	10,0%	20,0%
1,23	0	6	0,0%	20,0%

1,24	3	9	10,0%	30,0%
1,25	3	12	10,0%	40,0%
1,26	0	12	0,0%	40,0%
1,27	6	18	20,0%	60,0%
1,28	3	21	10,0%	70,0%
1,29	6	27	20,0%	90,0%
1,30	3	30	10,0%	100,0%

b) Distribución marginal de la variable Y (peso)

Obtenemos la siguiente tabla de frecuencia:

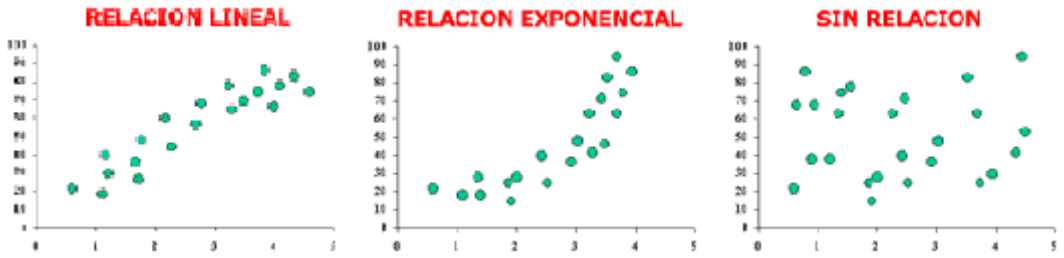
xVariable	Frecuencias absolutas		Frecuencias relativas	
(Peso)	Simple	Acumulada	Simple	Acumulada
31	6	6	20,0%	20,0%
32	6	12	20,0%	40,0%
33	6	18	20,0%	60,0%
34	7	25	23,3%	83,3%
35	5	30	16,6%	100,0%

LECCION 8c
Coeficiente de correlación lineal

En una distribución bidimensional puede ocurrir que las dos variables guarden algún tipo de relación entre si.

Por ejemplo, si se analiza la estatura y el peso de los alumnos de una clase es muy posible que exista relación entre ambas variables: mientras más alto sea el alumno, mayor será su peso.

El coeficiente de correlación lineal mide el grado de intensidad de esta posible relación entre las variables. Este coeficiente se aplica cuando la relación que puede existir entre las variables es lineal (es decir, si representaremos en un gráfico los pares de valores de las dos variables la nube de puntos se aproximaría a una recta).



No obstante, puede que exista una relación que no sea lineal, sino exponencial, parabólica, etc. En estos casos, el coeficiente de correlación lineal mediría mal la intensidad de la relación las variables, por lo que convendría utilizar otro tipo de coeficiente más apropiado.

Para ver, por tanto, si se puede utilizar el coeficiente de correlación lineal, lo mejor es representar los pares de valores en un gráfico y ver que forma describe.

El **coeficiente de correlación lineal** se calcula aplicando la siguiente fórmula:

$$r = \frac{1/n * \Sigma (x_i - xm) * (y_i - ym)}{\left((1/n * \Sigma (x_i - xm)^2) * (1/n * \Sigma (y_i - ym)^2) \right)^{1/2}}$$

Es decir:

Numerador: se denomina **covarianza** y se calcula de la siguiente manera: en cada par de valores (x,y) se multiplica la "x" menos su media, por la "y" menos su media. Se suma el resultado obtenido de todos los pares de valores y este resultado se divide por el tamaño de la muestra.

Denominador se calcula el producto de las varianzas de "x" y de "y", y a este producto se le calcula la raíz cuadrada.

Los valores que puede tomar el **coeficiente de correlación "r"** son: $-1 < r < 1$

Si "r" > 0, la correlación lineal es positiva (si sube el valor de una variable sube el de la otra). La correlación es tanto más fuerte cuanto más se aproxime a 1.

Por ejemplo: altura y peso: los alumnos más altos suelen pesar más.

Si "r" < 0, la correlación lineal es negativa (si sube el valor de una variable disminuye el de la otra). La correlación negativa es tanto más fuerte cuanto más se aproxime a -1.

Por ejemplo: peso y velocidad: los alumnos más gordos suelen correr menos.

Si "r" = 0, no existe correlación lineal entre las variables. Aunque podría existir otro tipo de correlación (parabólica, exponencial, etc.)

De todos modos, aunque el valor de "r" fuera próximo a 1 o -1, tampoco esto quiere decir obligatoriamente que existe una relación de causa-efecto entre las dos variables, ya que este resultado podría haberse debido al puro azar.

Ejemplo: vamos a calcular el coeficiente de correlación de la siguiente serie de datos de altura y peso de los alumnos de una clase:

Alumno	Estatura	Peso	Alumno	Estatura	Peso	Alumno	Estatura	Peso
Alumno 1	1,25	32	Alumno 11	1,25	33	Alumno 21	1,25	33
Alumno 2	1,28	33	Alumno 12	1,28	35	Alumno 22	1,28	34
Alumno 3	1,27	34	Alumno 13	1,27	34	Alumno 23	1,27	34
Alumno 4	1,21	30	Alumno 14	1,21	30	Alumno 24	1,21	31
Alumno 5	1,22	32	Alumno 15	1,22	33	Alumno 25	1,22	32
Alumno 6	1,29	35	Alumno 16	1,29	34	Alumno 26	1,29	34
Alumno 7	1,30	34	Alumno 17	1,30	35	Alumno 27	1,30	34
Alumno 8	1,24	32	Alumno 18	1,24	32	Alumno 28	1,24	31
Alumno 9	1,27	32	Alumno 19	1,27	33	Alumno 29	1,27	35
Alumno 10	1,29	35	Alumno 20	1,29	33	Alumno 30	1,29	34

Aplicamos la fórmula:

$$r = \frac{(1/30) * (0{,}826)}{(((1/30)*(0{,}02568)) * ((1/30)*(51{,}366)))^{(1/2)}}$$

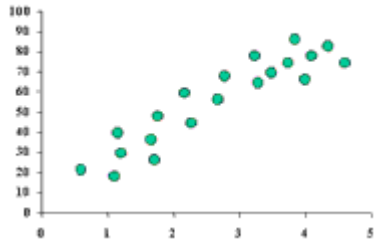
Luego,

$$r = 0{,}719$$

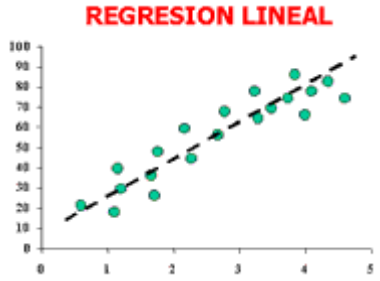
Por lo tanto, la correlación existente entre estas dos variables es elevada (0,7) y de signo positivo.

LECCION 8d
Regresión lineal

Representamos en un gráfico los pares de valores de una distribución bidimensional: la variable "x" en el eje horizontal o eje de abscisa, y la variable "y" en el eje vertical, o eje de ordenada. Vemos que la nube de puntos sigue una tendencia lineal:



El **coeficiente de correlación lineal** nos permite determinar si, efectivamente, existe relación entre las dos variables. Una vez que se concluye que sí existe relación, la **regresión** nos permite definir la recta que mejor se ajusta a esta nube de puntos.



Una recta viene definida por la siguiente fórmula:

$$y = a + bx$$

Donde "y" sería la variable dependiente, es decir, aquella que viene definida a partir de la otra variable "x" (variable independiente). Para definir la recta hay que determinar los valores de los parámetros "a" y "b":

El **parámetro "a"** es el valor que toma la variable dependiente "y", cuando la variable independiente "x" vale 0, y es el punto donde la recta cruza el eje vertical.

El **parámetro "b"** determina la pendiente de la recta, su grado de inclinación.

La **regresión lineal** nos permite calcular el valor de estos dos parámetros, definiendo la recta que mejor se ajusta a esta nube de puntos.

El **parámetro "b"** viene determinado por la siguiente fórmula:

$$b = \frac{1/n * \Sigma (x_i - xm) * (y_i - ym)}{1/n * \Sigma (x_i - xm)^2}$$

Es la covarianza de las dos variables, dividida por la varianza de la variable "x".

El **parámetro "a"** viene determinado por:

$$a = ym - (b * xm)$$

Es la media de la variable "y", menos la media de la variable "x" multiplicada por el parámetro "b" que hemos calculado.

Ejemplo: vamos a calcular la recta de regresión de la siguiente serie de datos de altura y peso de los alumnos de una clase. Vamos a considerar que la altura es la variable independiente "x" y que el peso es la variable dependiente "y" (podíamos hacerlo también al contrario):

Alumno	Estatura	Peso	Alumno	Estatura	Peso	Alumno	Estatura	Peso
Alumno 1	1,25	32	Alumno 11	1,25	33	Alumno 21	1,25	33
Alumno 2	1,28	33	Alumno 12	1,28	35	Alumno 22	1,28	34
Alumno 3	1,27	34	Alumno 13	1,27	34	Alumno 23	1,27	34
Alumno 4	1,21	30	Alumno 14	1,21	30	Alumno 24	1,21	31
Alumno 5	1,22	32	Alumno 15	1,22	33	Alumno 25	1,22	32
Alumno 6	1,29	35	Alumno 16	1,29	34	Alumno 26	1,29	34
Alumno 7	1,30	34	Alumno 17	1,30	35	Alumno 27	1,30	34
Alumno 8	1,24	32	Alumno 18	1,24	32	Alumno 28	1,24	31
Alumno 9	1,27	32	Alumno 19	1,27	33	Alumno 29	1,27	35
Alumno 10	1,29	35	Alumno 20	1,29	33	Alumno 30	1,29	34

El **parámetro "b"** viene determinado por:

$$b = \frac{(1/30) * 1,034}{(1/30) * 0,00856} = 40,265$$

Y el **parámetro "a"** por:

$$a = 33,1 - (40,265 * 1,262) = -17,714$$

Por lo tanto, la **recta** que mejor se ajusta a esta serie de datos es:

$$y = -17{,}714 + (40{,}265 * x)$$

Esta recta define un valor de la variable dependiente (peso), para cada valor de la variable independiente (estatura):

Estatura	Peso
1,20	30,6
1,21	31,0
1,22	31,4
1,23	31,8
1,24	32,2
1,25	32,6
1,26	33,0
1,27	33,4
1,28	33,8
1,29	34,2
1,30	34,6

TEMA 9ª
Teorema de Bayes

El **Teorema de Bayes** viene a seguir el **proceso inverso** al que hemos visto en el **Teorema de la probabilidad total**:

Teorema de la probabilidad total: a partir de las probabilidades del suceso A (probabilidad de que llueva o de que haga buen tiempo) deducimos la probabilidad del suceso B (que ocurra un accidente).

Teorema de Bayes: a partir de que ha ocurrido el suceso B (ha ocurrido un accidente) deducimos las probabilidades del suceso A (¿estaba lloviendo o hacía buen tiempo?).

La **fórmula** del Teorema de Bayes es:

Tratar de explicar estar fórmula con palabras es un galimatías, así que vamos a intentar explicarla con un ejemplo. De todos modos, antes de entrar en el ejercicio, recordar que este teorema también exige que el **suceso A forme un sistema completo.**

Ejercicio 1º: El parte meteorológico ha anunciado tres posibilidades para el fin de semana:

a) Que llueva: probabilidad del 50%.

b) Que nieve: probabilidad del 30%

c) Que haya niebla: probabilidad del 20%.

Según estos posibles estados meteorológicos, la posibilidad de que ocurra un accidente es la siguiente:

a) Si llueve: probabilidad de accidente del 10%.

b) Si nieva: probabilidad de accidente del 20%

c) Si hay niebla: probabilidad de accidente del 5%.

Resulta que efectivamente ocurre un accidente y como no estábamos en la ciudad no sabemos que tiempo hizo (nevó, llovió o hubo niebla). El teorema de Bayes nos permite calcular estas probabilidades:

Las probabilidades que manejamos antes de conocer que ha ocurrido un accidente se denominan **"probabilidades a priori"** (lluvia con el 60%, nieve con el 30% y niebla con el 10%).

Una vez que incorporamos la información de que ha ocurrido un accidente, las probabilidades del suceso A cambian: son probabilidades condicionadas P (A/B), que se denominan **"probabilidades a posteriori"**.

Vamos a aplicar la fórmula:

a) Probabilidad de que estuviera lloviendo:

La probabilidad de que efectivamente estuviera lloviendo el día del accidente (probabilidad a posteriori) es del 71,4%.

b) Probabilidad de que estuviera nevando:

La probabilidad de que estuviera nevando es del 21,4%.

c) Probabilidad de que hubiera niebla:

La probabilidad de que hubiera niebla es del 7,1%.

LECCION 9b
Independencia de sucesos

Dos **sucesos son independientes** entre sí, si la ocurrencia de uno de ellos no afecta para nada a la ocurrencia del otro:

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.

Para que dos sucesos sean independientes tienen que verificar al menos una de las siguientes **condiciones**:

P (B/A) = P (B) es decir, que la probabilidad de que se de el suceso B, condicionada a que previamente se haya dado el suceso A, es exactamente igual a la probabilidad de B.

Ejemplo: la probabilidad de que al tirar una moneda salga cara (suceso B), condicionada a que haga buen tiempo (suceso A), es igual a la propia probabilidad del suceso B.

P (A/B) = P (A) es decir, que la probabilidad de que se de el suceso A, condicionada a que previamente se haya dado el suceso B, es exactamente igual a la probabilidad de A.

Ejemplo: la probabilidad de que haga buen tiempo (suceso A), condicionada a que al tirar una moneda salga cara (suceso B), es igual a la propia probabilidad del suceso A.

P (A $\wedge$ B) = P (A) * P (B) es decir, que la probabilidad de que se de el suceso conjunto A y B es exactamente igual a la probabilidad del suceso A multiplicada por la probabilidad del suceso B.

Ejemplo: la probabilidad de que haga buen tiempo (suceso A) y salga cara al tirar una moneda (suceso B), es igual a la probabilidad del suceso A multiplicada por la probabilidad del suceso B

Si el **suceso A es independiente del suceso B**, entonces **el suceso B** también es **independiente del suceso A.**

Ejemplo 1º: analicemos dos sucesos:

Suceso A: la probabilidad de que haga buen tiempo es del 0,4

Suceso B: la probabilidad de tener un accidente es del 0,1

Suceso intersección: la probabilidad de que haga buen tiempo y tener un accidente es del 0,08

Veamos si se cumple alguna de las condiciones señaladas:

P (B/A) = P (A Λ B) / P (A) = 0,08 / 0,4 = 0,2 (que no es igual a P (B))

P (A/B) = P (A Λ B) / P (B) = 0,08 / 0,6 = 0,133 (que no es igual a P (A))

P (A Λ B) = 0,08 (que no es igual a P (A) multiplicado por P (B))

Por lo tanto, no se cumple ninguna de las tres condiciones señaladas por lo que **estos dos sucesos no son independientes**, sino que existe algún grado de dependencia entre ellos.

Ejemplo 2º: analicemos dos sucesos:

Suceso A: la probabilidad de que haga buen tiempo es del 0,4

Suceso B: la probabilidad de salir cara al lanzar una moneda es del 0,5

Suceso intersección: la probabilidad de que haga buen tiempo y que salga cara es 0,2

Veamos si se cumple alguna de las condiciones señaladas:

P (B/A) = P (A Λ B) / P (A) = 0,2 / 0,4 = 0,5 (igual que P (B))

P (A/B) = P (A Λ B) / P (B) = 0,2 / 0,6 = 0,4 (igual que P (A))

P (A Λ B) = 0,2 (igual a P (A) multiplicado por P (B))

Por lo tanto, **estos dos sucesos sí son independientes.**

TEMA 10ª
Distribuciones discretas: Binomial

La **distribución binomial** parte de la distribución de Bernouilli:

La **distribución de Bernouilli** se aplica cuando se realiza una sola vez un experimento que tiene únicamente dos posibles resultados (éxito o fracaso), por lo que la variable sólo puede tomar dos valores: el 1 y el 0

La **distribución binomial** se aplica cuando se realizan un número" de veces el experimento de Bernouilli, siendo cada ensayo independiente del anterior. La variable puede tomar valores entre:

0: si todos los experimentos han sido fracaso

n: si todos los experimentos han sido éxitos

Ejemplo: se tira una moneda 10 veces: ¿cuantas caras salen? Si no ha salido ninguna la variable toma el valor 0; si han salido dos caras la variable toma el valor 2; si todas han sido cara la variable toma el valor 10

La **distribución de probabilidad** de este tipo de distribución sigue el siguiente modelo:

¿Alguien entiende esta fórmula? Vamos a tratar de explicarla con un ejemplo:

Ejemplo 1: ¿Cuál es la probabilidad de obtener 6 caras al lanzar una moneda 10 veces?

" k " es el número de aciertos. En este ejemplo " k " igual a 6 (en cada acierto decíamos que la variable toma el valor 1: como son 6 aciertos, entonces k = 6)

" n" es el número de ensayos. En nuestro ejemplo son 10

" p " es la probabilidad de éxito, es decir, que salga **"cara"** al lanzar la moneda. Por lo tanto p = 0,5

La fórmula quedaría:

Luego,

P (x = 6) = 0,205

Es decir, se tiene una probabilidad del 20,5% de obtener 6 caras al lanzar 10 veces una moneda.

Ejemplo 2: ¿Cuál es la probabilidad de obtener cuatro veces el número 3 al lanzar un dado ocho veces?

" k " (número de aciertos) toma el valor 4

" n" toma el valor 8

" p " (probabilidad de que salga un 3 al tirar el dado) es 1 / 6 (= 0,1666)

La fórmula queda:

Luego,

P (x = 4) = 0,026

Es decir, se tiene una probabilidad del 2,6% de obtener cuatro veces el números 3 al tirar un dado 8 veces.

LECCION 10b
Distribuciones discretas: Poisson

La **distribución de Poisson** parte de la distribución binomial:

Cuando en una distribución binomial se realiza el experimento un número "n" muy elevado de veces y la probabilidad de éxito "p" en cada ensayo es reducida, entonces se aplica el modelo de **distribución de Poisson:**

Se tiene que cumplir que:

" p " < 0,10

" p * n " < 10

La **distribución de Poisson** sigue el siguiente modelo:

Vamos a explicarla:

El número **"e"** es 2,71828

" **λ** " = n * p (es decir, el número de veces " n " que se realiza el experimento multiplicado por la probabilidad " p " de éxito en cada ensayo)

" **k** " es el número de éxito cuya probabilidad se está calculando

Veamos un **ejemplo**:

La probabilidad de tener un accidente de tráfico es de 0,02 cada vez que se viaja, si se realizan 300 viajes, ¿cual es la probabilidad de tener 3 accidentes?

Como la probabilidad " p " es menor que 0,1, y el producto " n * p " es menor que 10, entonces aplicamos el modelo de distribución de Poisson.

Luego,

$$P\ (x = 3) = 0{,}0892$$

Por lo tanto, la probabilidad de tener 3 accidentes de tráfico en 300 viajes es del 8,9%

Otro **ejemplo**:

La probabilidad de que un niño nazca pelirrojo es de 0,012. ¿Cuál es la probabilidad de que entre 800 recien nacidos haya 5 pelirrojos?

Luego,

$$P\ (x = 5) = 4{,}602$$

Por lo tanto, la probabilidad de que haya 5 pelirrojos entre 800 recién nacidos es del 4,6%.

LECCION 10c
Distribuciones discretas: Hipergeométrica

La **distribución hipergeométrica** es el modelo que se aplica en experimentos del siguiente tipo:

En una urna hay bolas de dos colores (blancas y negras), ¿cuál es la probabilidad de que al sacar 2 bolas las dos sean blancas?

Son experimentos donde, al igual que en la distribución binomial, en cada ensayo hay tan sólo **dos posibles resultados**: o sale blanca o no sale. Pero se diferencia de la distribución binomial en que **los distintos ensayos son dependientes** entre sí:

Si en una urna con 5 bolas blancas y 3 negras en un primer ensayo saco una bola blanca, en el segundo ensayo hay una bola blanca menos por lo que las probabilidades son diferentes (hay dependencia entre los distintos ensayos).

La **distribución hipergeométrica** sigue el siguiente modelo:

Donde:

Vamos a tratar de explicarlo:

N: es el número total de bolas en la urna

N1: es el número total de bolas blancas

N2: es el número total de bolas negras

k: es el número de bolas blancas cuya probabilidad se está calculando

n: es el número de ensayos que se realiza

Veamos un **ejemplo**: en una urna hay 7 bolas blancas y 5 negras. Se sacan 4 bolas ¿Cuál es la probabilidad de que 3 sean blancas?

Entonces:

N = 12; N1 = 7; N2 = 5; k = 3; n = 4

Si aplicamos el modelo:

Por lo tanto, P (x = 3) = 0,3535. Es decir, la probabilidad de sacar 3 bolas blancas es del 35,3%.

Pero este modelo no sólo se utiliza con experimentos con bolas, sino que también se aplica con experimentos similares:

Ejemplo: en una fiesta hay 20 personas: 14 casadas y 6 solteras. Se eligen 3 personas al azar ¿Cuál es la probabilidad de que las 3 sean solteras?

Por lo tanto, P (x = 3) = 0,0175. Es decir, la probabilidad de que las 3 personas sean solteras es tan sólo del 1,75%.

LECCION 11ª
Distribuciones discretas: Multinomial

La **distribución multinomial** es similar a la distribución binomial, con la diferencia de que en lugar de dos posibles resultados en cada ensayo, puede haber múltiples resultados:

Ejemplo de distribución binomial: a unas elecciones se presentaron 2 partidos políticos: el POR obtuvo un 70% de los votos y el JEJE el 30% restante. ¿Cuál es la probabilidad de que al elegir 5 ciudadanos al azar, 4 de ellos hayan votado al JEJE?

Ejemplo de distribución multinomial: a esas elecciones se presentaron 4 partidos políticos: el POR obtuvo un 40% de los votos, el JEJE el 30%, el MUM el 20% y el ALK el 10% restante. ¿Cuál es la probabilidad de que al elegir 5 ciudadanos al azar, 3 hayan votado al POR, 1 al MUM y 1 al ALK?

La **distribución multinomial** sigue el siguiente modelo:

$$\frac{N!}{X_1!X_2!\cdots X_K!}p_1^{X_1}p_2^{X_2}\cdots p_K^{X_K}$$

Donde:

X1 = x1: indica que el suceso X1 aparezca x1 veces (en el ejemplo, que el partido POR lo hayan votado 3 personas)

n: indica el número de veces que se ha repetido el suceso (en el ejemplo, 5 veces)

n!: es factorial de n (en el ejemplo: 5 * 4 * 3 * 2 * 1)

p1: es la probabilidad del suceso X1 (en el ejemplo, el 40%)

Veamos el **ejemplo**:

Luego:

P = 0,0256

Es decir, que la probabilidad de que las 5 personas elegidas hayan votado de esta manera es tan sólo del 2,56%

Nota: 0! es igual a 1, y cualquier número elevado a 0 es también igual a 1

Veamos otro **ejemplo**:

En una fiesta, el 20% de los asistentes son paceños, el 30% cruceños, el 40% tarijeños y el 10% orureños. En un pequeño grupo se han reunido 4 invitados: ¿cual es la probabilidad de que 2 sean cruceños y 2 tarijeños?

Aplicamos el modelo:

Luego

P = 0,0384

Por lo tanto, la probabilidad de que el grupo esté formado por personas de estos paisanos es tan sólo del 3,84%.

LECCION 11b

Distribuciones discretas: Multihipergeométrica

La **distribución multihipergeométrica** es similar a la distribución hipergeométrica, con la diferencia de que en la urna, en lugar de haber únicamente bolas de dos colores, hay bolas de diferentes colores.

Ejemplo: en una urna hay 7 bolas blancas, 3 verdes y 4 amarillas: ¿cuál es la probabilidad de que al extraer 3 bolas sea cada una de un color distinto?

La **distribución multihipergeométrica** sigue el siguiente modelo:

Donde:

X1 = x1: indica que el suceso X1 aparezca x1 veces (en el ejemplo, que una de las bolas sea blanca)

N1: indica el número de bolas blancas que hay en la urna (en el ejemplo, 7 bolas)

N: es el número total de bolas en la urna (en el ejemplo, 14 bolas)

n: es el número total de bolas que se extraen (en el ejemplo, 3 bolas)

Veamos el **ejemplo**:

Luego:

P = 0,2307

Es decir, que la probabilidad de sacar una bola de cada color es del 23,07%.

Veamos otro **ejemplo**:

En una caja de lápices hay 10 de color amarillo, 3 de color azul y 4 de color rojo. Se extraen 7 lápices, ¿cual es la probabilidad de que 5 sean amarillos y 2 rojos?

Aplicamos el modelo:

Luego

P = 0,0777

Por lo tanto, la probabilidad de que los 5 lápices sean de los colores indicados es del 7,77%.

LECCION 11c

Distribuciones continuas: Uniforme

La **distribución uniforme** es aquella que puede tomar cualquier valor dentro de un intervalo, todos ellos con la misma probabilidad.

Es una **distribución continua** porque puede tomar cualquier valor y no únicamente un número determinado (como ocurre en las distribuciones discretas).

Ejemplo: el precio medio del litro de gasolina durante el próximo año se estima que puede oscilar entre 2.40 y 2.60 bs. Podría ser, por tanto, de 143 bs, o de 143,4 bs, o de 143,45 bs., o de 143,455 bs, etc. Hay infinitas posibilidades, todas ellas con la misma probabilidad.

Su **función de densidad,** aquella que nos permite conocer la probabilidad que tiene cada punto del intervalo, viene definida por:

Donde:

b: es el extremo superior (en el ejemplo, 2.60 bs.)

a: es el extremo inferior (en el ejemplo, 2.40 bs.)

Por lo tanto, la función de distribución del **ejemplo** sería:

Es decir, que el valor final esté entre 2.40 bs. y 2.41 bs. Tiene un 5% de probabilidad, que esté entre 2.41 y 2.42, otro 5%, etc.

El valor medio de esta distribución se calcula:

En el ejemplo:

Por lo tanto, el precio medio esperado de la gasolina para el próximo año es de 2.50 bs.

Veamos otro **ejemplo**:

El volumen de precipitaciones estimado para el próximo año en la ciudad de Trinidad va a oscilar entre 400 y 500 litros por metro cuadrado. Calcular la función de distribución y la precipitación media esperada:

Es decir, que el volumen de precipitaciones esté entre 400 y 401 litros tiene un 1% de probabilidades; que esté entre 401 y 402 litros, otro 1%, etc.

El valor medio esperado es:

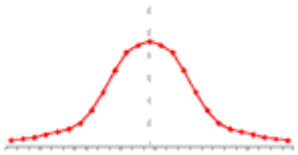
Es decir, la precipitación media estimada en Trinidad para el próximo año es de 450 litros.

LECCION 11d

Distribuciones continuas: Normal (I)

Es el **modelo de distribución más utilizado** en la práctica, ya que multitud de fenómenos se comportan según una distribución normal.

Esta distribución de caracteriza porque los valores se distribuyen formando una **campana de Gauss**, en torno a un valor central que coincide con el valor medio de la distribución:



Un 50% de los valores están a la derecha de este valor central y otro 50% a la izquierda

Esta distribución viene definida por **dos parámetros**:

$$X: N(\mu, \sigma^2)$$

μ **:** es el valor medio de la distribución y es precisamente donde se sitúa el centro de la curva (de la campana de Gauss).

σ^2 : es la varianza. Indica si los valores están más o menos alejados del valor central: si la varianza es baja los valores están próximos a la media; si es alta, entonces los valores están muy dispersos.

Cuando la media de la distribución es 0 y la varianza es 1se denomina **"normal tipificada**", y su ventaja reside en que hay tablas donde se recoge la probabilidad acumulada para cada punto de la curva de esta distribución.

Además, **toda distribución normal se puede transformar en una normal tipificada:**

Ejemplo: una variable aleatoria sigue el modelo de una distribución normal con media 10 y varianza 4. Transformarla en una normal tipificada.

$$X: N(10, 4)$$

Para transformarla en una normal tipificada **se crea una nueva variable** (Y) que será **igual a la anterior (X) menos su media y dividida por su desviación típica** (que es la raíz cuadrada de la varianza)

$$\sigma_x = +\sqrt{\sigma_x^2}$$

En el ejemplo, la nueva variable sería:

Esta nueva variable se distribuye como una normal tipificada, permitiéndonos, por tanto, conocer la probabilidad acumulada en cada valor.

Y: N (0, 1)

LECCION 11d

Distribuciones continuas: Normal (II)

La **distribución normal tipificada** tiene la ventaja, como ya hemos indicado, de que las probabilidades para cada valor de la curva se encuentran recogidas en una tabla.

X	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5723
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7090	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7813	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8416	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,97725	0,97778	0,97831	0,97882	0,97932	0,97982	0,98030	0,98077	0,98124	0,98169
2,1	0,98214	0,98257	0,98300	0,98341	0,98382	0,98422	0,98461	0,98500	0,98537	0,98574
2,2	0,98610	0,98645	0,98679	0,98713	0,98745	0,98778	0,98809	0,98840	0,98870	0,98899
2,3	0,98928	0,98956	0,98983	0,99010	0,99036	0,99061	0,99086	0,99111	0,99134	0,99158
2,4	0,99180	0,99202	0,99224	0,99245	0,99266	0,99286	0,99305	0,99324	0,99343	0,99361
2,5	0,99379	0,99396	0,99413	0,99430	0,99446	0,99461	0,99477	0,99492	0,99506	0,99520
2,6	0,99534	0,99547	0,99560	0,99573	0,99585	0,99598	0,99609	0,99621	0,99632	0,99643
2,7	0,99653	0,99664	0,99674	0,99683	0,99693	0,99702	0,99711	0,99720	0,99728	0,99736
2,8	0,99744	0,99752	0,99760	0,99767	0,99774	0,99781	0,99788	0,99795	0,99801	0,99807
2,9	0,99813	0,99819	0,99825	0,99831	0,99836	0,99841	0,99846	0,99851	0,99856	0,99861

¿Cómo se lee esta tabla?

La columna de la izquierda indica el valor cuya probabilidad acumulada queremos conocer. La primera fila nos indica el segundo decimal del valor que estamos consultando.

Ejemplo: queremos conocer la probabilidad acumulada en el valor 2,75.Entonces buscamos en la columna de la izquierda el valor 2,7 y en la primera fila el valor 0,05. La casilla en la que se interseccionan es su probabilidad acumulada (0,99702, es decir 99.7%).

Atención: la tabla nos da la probabilidad acumulada, es decir, la que va desde el inicio de la curva por la izquierda hasta dicho valor. No nos da la probabilidad concreta en ese punto. En una distribución continua en el que la variable puede tomar infinitos valores, la probabilidad en un punto concreto es prácticamente despreciable.

Ejemplo: Imaginemos que una variable continua puede tomar valores entre 0 y 5. La probabilidad de que tome exactamente el valor 2 es despreciable, ya que podría tomar infinitos valores: por ejemplo: 1,99, 1,994, 1,9967, 1,9998, 1999791, etc.

Veamos **otros ejemplos**:

Probabilidad acumulada en el valor 0,67: la respuesta es 0,7486

Probabilidad acumulada en el valor 1,35: la respuesta es 0,9115

Probabilidad acumulada en el valor 2,19: la respuesta es 0,98574

Veamos ahora, como podemos **utilizar esta tabla con una distribución normal**:

Ejemplo: el salario medio de los empleados de una empresa se distribuye según una distribución normal, con media 5 millones de bs. y desviación típica 1 millón de bs. Calcular el porcentaje de empleados con un sueldo inferior a 7 millones de bs.

Lo primero que haremos es transformar esa distribución en una normal tipificada, para ello se crea una nueva variable (Y) que será igual a la anterior (X) menos su media y dividida por la desviación típica

En el ejemplo, la nueva variable sería:

Esta nueva variable se distribuye como una normal tipificada. La variable Y que corresponde a una variable X de valor 7 es:

Ya podemos consultar en la tabla la probabilidad acumulada para el valor 2 (equivalente a la probabilidad de sueldos inferiores a 7 millones de bs.). Esta probabilidad es 0,97725

Por lo tanto, el porcentaje de empleados con salarios inferiores a 7 millones de bs. Es del 97,725%.

LECCION 12ª

Distribuciones continuas: Normal (III): Ejercicios

Ejercicio 1º: La renta media de los habitantes de un país es de 4 millones de bs/año, con una varianza de 1,5. Se supone que se distribuye según una distribución normal. Calcular:

a) Porcentaje de la población con una renta inferior a 3 millones de bs.

b) Renta a partir de la cual se sitúa el 10% de la población con mayores ingresos.

c) Ingresos mínimo y máximo que engloba al 60% de la población con renta media.

a) Porcentaje de la población con una renta inferior a 3 millones de bs.

Lo primero que tenemos que hacer es calcular la normal tipificada:

(*) Recordemos que el denominador es la desviación típica (raíz cuadrada de la varianza)

El valor de Y equivalente a 3 millones de bs es -0,816.

$$P (X < 3) = P (Y < -0,816)$$

Ahora tenemos que ver cuál es la probabilidad acumulada hasta ese valor. Tenemos un problema: la tabla de probabilidades (ver lección 35) sólo abarca valores positivos, no obstante, este problema tiene fácil solución, ya que la distribución normal es simétrica respecto al valor medio.

Por lo tanto:

$$P (Y < -0,816) = P (Y > 0,816)$$

Por otra parte, la probabilidad que hay a partir de un valor es igual a 1 (100%) menos la probabilidad acumulada hasta dicho valor:

$$P (Y > 0,816) = 1 - P (Y < 0,816) = 1 - 0,7925 \text{ (aprox.)} = 0,2075$$

Luego, el 20,75% de la población tiene una renta inferior a 3 millones bs.

b) Nivel de ingresos a partir del cual se sitúa el 10% de la población con renta más elevada.

Vemos en la tabla el valor de la variable tipificada cuya probabilidad acumulada es el 0,9 (90%), lo que quiere decir que por encima se sitúa el 10% superior.

Ese valor corresponde a Y = 1,282 (aprox.). Ahora calculamos la variable normal X equivalente a ese valor de la normal tipificada:

Despejando X, su valor es 5,57. Por lo tanto, aquellas personas con ingresos superiores a 5,57 millones de bs. Constituyen el 10% de la población con renta más elevada.

c) Nivel de ingresos mínimo y máximo que engloba al 60% de la población con renta media

Vemos en la tabla el valor de la variable normalizada Y cuya probabilidad acumulada es el 0,8 (80%). Como sabemos que hasta la media la probabilidad acumulada es del 50%, quiere decir que entre la media y este valor de Y hay un 30% de probabilidad.

Por otra parte, al ser la distribución normal simétrica, entre -Y y la media hay otro 30% de probabilidad. En definitiva, el segmento (-Y, Y) engloba al 60% de población con renta media.

El valor de Y que acumula el 80% de la probabilidad es 0,842 (aprox.), por lo que el segmento viene definido por (-0,842, +0,842). Ahora calculamos los valores de la variable X correspondientes a estos valores de Y.

Los valores de X son 2,97 y 5,03. Por lo tanto, las personas con ingresos superiores a 2,97 millones de ptas. e inferiores a 5,03 millones de bs. Constituyen el 60% de la población con un nivel medio de renta.

Ejercicio 2º: La vida media de los habitantes de un país es de 68 años, con una varianza de 25. Se hace un estudio en una pequeña ciudad de 10.000 habitantes:

a) ¿Cuántas personas superarán previsiblemente los 75 años?

b) ¿Cuántos vivirán menos de 60 años?

a) Personas que vivirán (previsiblemente) más de 75 años

Calculamos el valor de la normal tipificada equivalente a 75 años

Por lo tanto

$$P(X > 75) = (Y > 1,4) = 1 - P(Y < 1,4) = 1 - 0,9192 = 0,0808$$

Luego, el 8,08% de la población (808 habitantes) vivirán más de 75 años.

b) Personas que vivirán (previsiblemente) menos de 60 años

Calculamos el valor de la normal tipificada equivalente a 60 años

Por lo tanto

$$P (X < 60) = (Y < -1,6) = P (Y > 1,6) = 1 - P (Y < 1,6) = 0,0548$$

Luego, el 5,48% de la población (548 habitantes) no llegarán probablemente a esta edad.

LECCION 12b
Distribuciones continuas: Normal (IV): Ejercicios

Ejercicio 1º: El consumo medio anual de cerveza de los habitantes de un país es de 59 litros, con una varianza de 36. Se supone que se distribuye según una distribución normal.

a) Si usted presume de buen bebedor, ¿cuántos litros de cerveza tendría que beber al año para pertenecer al 5% de la población que más bebe?

b) Si usted bebe 45 litros de cerveza al año y su mujer le califica de borracho ¿qué podría argumentar en su defensa?

a) 5% de la población que más bebe.

Vemos en la tabla el valor de la variable tipificada cuya probabilidad acumulada es el 0,95 (95%), por lo que por arriba estaría el 5% restante.

Ese valor corresponde a Y = 1,645 (aprox.). Ahora calculamos la variable normal X equivalente a ese valor de la normal tipificada:

Despejando X, su valor es 67,87. Por lo tanto, tendría usted que beber más de 67,87 litros al año para pertenecer a ese "selecto" club de grandes bebedores de cerveza.

b) Usted bebe 45 litros de cerveza al año. ¿Es usted un bebedor?

Vamos a ver en que nivel de la población se situaría usted en función de los litros de cerveza consumidos.

Calculamos el valor de la normal tipificada correspondiente a 45 litros:

Por lo tanto

$$P (X < 45) = (Y < -2,2) = P (Y > 2,2) = 1 - P (Y < 2,2) = 0,0139$$

Luego, tan sólo un 1,39% de la población bebe menos que usted. Parece un argumento de suficiente peso para que dejen de catalogarle de "enamorado de la bebida"

Ejercicio 2º: A un examen de oposición se han presentado 2.000 aspirantes. La nota media ha sido un 5,5, con una varianza de 1,5.

a) Tan sólo hay 100 plazas. Usted ha obtenido un 7,7. ¿Sería oportuno ir organizando una fiesta para celebrar su éxito?

b) Va a haber una 2ª oportunidad para el 20% de notas más altas que no se hayan clasificados. ¿A partir de que nota se podrá participar en esta "revalida"?

a) Ha obtenido usted un 7,7

Vamos a ver con ese 7,7 en que nivel porcentual se ha situado usted, para ello vamos a comenzar por calcular el valor de la normal tipificada equivalente.

A este valor de Y le corresponde una probabilidad acumulada (ver tablas) de 0,98214 (98,214%), lo que quiere decir que por encima de usted tan sólo se encuentra un 1,786%.

Si se han presentado 2.000 aspirante, ese 1,786% equivale a unos 36 aspirantes. Por lo que si hay 100 plazas disponibles, tiene usted suficientes probabilidades como para ir organizando la "mejor de las fiestas".

b) "Revalida" para el 20% de los candidatos

Vemos en la tabla el valor de la normal tipificada que acumula el 80% de la probabilidad, ya que por arriba sólo quedaría el 20% restante.

Este valor de Y corresponde a 0,842 (aprox.). Ahora calculamos el valor de la normal X equivalente:

Despejamos la X y su valor es 6,38. Por lo tanto, esta es la nota a partir de la cual se podrá acudir a la "revalida".

LECCION 13ª

Teorema Central del Límite

El **Teorema Central del Límite** dice que si tenemos un grupo numeroso de variables independientes y todas ellas siguen el mismo modelo de distribución (cualquiera que éste sea), la suma de ellas se distribuye según una **distribución normal**.

Ejemplo: la variable "tirar una moneda al aire" sigue la distribución de Bernouilli. Si lanzamos la moneda al aire 50 veces, la suma de estas 50 variables (cada una independiente entre si) se distribuye según una distribución normal.

Este teorema se aplica tanto a suma de variables discretas como de variables continuas.

Los parámetros de la distribución normal son:

Media: **n * μ** (media de la variable individual multiplicada por el número de variables independientes)

Varianza: **n * σ2** (varianza de la variable individual multiplicada por el número de variables individuales)

Veamos un **ejemplo**:

Se lanza una moneda al aire 100 veces, si sale cara le damos el valor 1 y si sale cruz el valor 0. Cada lanzamiento es una variable independiente que se distribuye según el modelo de Bernouilli, con media 0,5 y varianza 0,25.

Calcular la probabilidad de que en estos 100 lanzamientos salgan más de 60 caras.

La variable suma de estas 100 variables independientes se distribuye, por tanto, según una distribución normal.

$$\text{Media} = 100 * 0{,}5 = 50$$

$$\text{Varianza} = 100 * 0{,}25 = 25$$

Para ver la probabilidad de que salgan más de 60 caras calculamos la variable normal tipificada equivalente:

(*) 5 es la raiz cuadrada de 25, o sea la desviación típica de esta distribución

Por lo tanto:

$$P(X > 60) = P(Y > 2{,}0) = 1 - P(Y < 2{,}0) = 1 - 0{,}9772 = 0{,}0228$$

Es decir, la probabilidad de que al tirar 100 veces la moneda salgan más de 60 caras es tan sólo del 2,28%

LECCION 13b
Teorema Central del Límite: Ejercicios (I)

Ejercicio 1.

La renta media de los habitantes de un país se distribuye uniformemente entre 4,0 millones ptas. y 10,0 millones ptas. Calcular la probabilidad de que al seleccionar al azar a 100 personas la suma de sus rentas supere los 725 millones ptas.

Cada renta personal es una variable independiente que se ditribuye según una función uniforme. Por ello, a la suma de las rentas de 100 personas se le puede aplicar el **Teorema Central del Límite.**

La **media** y **varianza** de cada variable individual es:

$$\mu = (4 + 10) / 2 = 7$$

$$\sigma 2 = (10 - 4)^2 / 12 = 3$$

Por tanto, la suma de las 100 variables se distribuye según una normal cuya **media** y **varianza** son:

Media: $n * \mu = 100 * 7 = 700$

Varianza: $n * \sigma 2 = 100 * 3 = 300$

Para calcular la probabilidad de que la suma de las rentas sea superior a 725 millones ptas, comenzamos por calcular el valor equivalente de la variable normal tipificada:

Luego:

$$P (X > 725) = P (Y > 1,44) = 1 - P (Y < 1,44) = 1 - 0,9251 = 0,0749$$

Es decir, la probabilidad de que la suma de las rentas de 100 personas seleccionadas al azar supere los 725 millones de pesetas es tan sólo del 7,49%

Ejercicio 2.

En una asignatura del colegio la probabilidad de que te saquen a la pizarra en cada clase es del 10%. A lo largo del año tienes 100 clases de esa asignatura. ¿Cuál es la probabilidad de tener que salir a la pizarra más de 15 veces?

Se vuelve a aplicar el **Teorema Central del Límite**.

Salir a la pizarra es una variable independiente que sigue el modelo de distribución de Bernouilli:

"Salir a la pizarra", le damos el valor 1 y tiene una probabilidad del 0,10

"No salir a la pizarra", le damos el valor 0 y tiene una probabilidad del 0,9

La **media** y la **varianza** de cada variable independiente es:

$$\mu = 0{,}10$$

$$\sigma^2 = 0{,}10 * 0{,}90 = 0{,}09$$

Por tanto, la suma de las 100 variables se distribuye según una normal cuya **media** y **varianza** son:

Media: $n * \mu = 100 * 0{,}10 = 10$

Varianza: $n * \sigma^2 = 100 * 0{,}09 = 9$

Para calcular la probabilidad de salir a la pizarra más de 15 veces, calculamos el valor equivalente de la variable normal tipificada:

Luego:

$$P(X > 15) = P(Y > 1{,}67) = 1 - P(Y < 1{,}67) = 1 - 0{,}9525 = 0{,}0475$$

Es decir, la probabilidad de tener que salir más de 15 veces a la pizarra a lo largo del curso es tan sólo del 4,75% (¡¡¡ ánimo !!!, no es tan grave)

LECCION 13c
Teorema Central del Límite: Ejercicios (II)

Ejercicio 1.

Un día visitamos un Casino y decidimos jugar en la ruleta. Nuestra apuesta va a ser siempre al negro y cada apuesta de 500 bs. Llevamos 10.000 bs. y queremos calcular que probabilidad tenemos de que tras jugar 80 veces consigamos doblar nuestro dinero.

Cada jugada es una variable independiente que sigue el modelo de distribución de Bernouilli.

"Salir negro", le damos el valor 1 y tiene una probabilidad del 0,485

"No salir negro", le damos el valor 0 y tiene una probabilidad del 0,515

(*) La probabilidad de "no salir negro" es mayor ya que puede salir rojo o el cero

La **media** y **varianza** de cada variable individual es:

$$\mu = 0{,}485$$

$$\sigma^2 = 0{,}485 * 0{,}515 = 0{,}25$$

A la suma de las 80 apuestas se le aplica el **Teorema Central del Límite**, por lo que se distribuye según una normal cuya **media** y **varianza** son:

Media: $n * \mu = 80 * 0{,}485 = 38{,}8$

Varianza: $n * \sigma^2 = 80 * 0{,}25 = 20$

Para doblar nuestro dinero el negro tiene que salir al menos 20 veces más que el rojo (20 * 500 = 10.000), por lo que tendrá que salir como mínimo 50 veces (implica que el rojo o el cero salgan como máximo 30 veces).

Comenzamos por calcular el valor equivalente de la variable normal tipificada:

Luego:

$$P(X > 50) = P(Y > 2{,}50) = 1 - P(Y < 2{,}50) = 1 - 0{,}9938 = 0{,}0062$$

Es decir, la probabilidad de doblar el dinero es tan sólo del 0,62% (así, que más vale que nos pongamos a trabajar).

Ejercicio 2.

El precio de una acción en bolsa se mueve aleatoriamente entre 10 bs. y 20 bs., con la misma probabilidad en todo el tramo. Hemos dado la orden a nuestro pariente de que nos compre paquetes de 1.000 acciones cada día durante las próximas 40 sesiones.

Una vez ejecutada la orden, tenemos un total de 40.000 acciones. A final de año vendemos todas las acciones al precio de 13 bs./acción, recibiendo 520.000 bs. Calcular la probabilidad de que ganemos dinero en esta operación.

El precio de cada paquete comprado es una variable aleatoria independiente que se distribuye uniformemente entre 10.000 bs y 20.000 bs. Su **media** y **varianza** son:

$$\mu = (10.000 + 20.000) / 2 = 15.000$$

$$\sigma 2 = (20.000 - 10.000)^2 / 12 = 833,3$$

El precio total de los 40 paquetes comprados se distribuye según una distribución normal cuya **media** y **varianza** son:

Media: $n * \mu = 40 * 15.000 = 600.000$

Varianza: $n * \sigma 2 = 40 * 833,3 = 33.333,3$

Para estimar la probabilidad de que ganemos dinero, calculamos el valor equivalente de la variable normal tipificada:

Luego:

$$P (X > 520.000) = P (Y > 2,40) = 1 - P (Y < 2,40) = 1 - 0,9918 = 0,0082$$

Por tanto, la probabilidad de que ganemos dinero con la "dichosa" operación es tan sólo del 0,82% (¡¡¡ lo llevamos claro, Gervasio !!!)

REFERENCIAS BIBLIOGRAFICAS

Barro, Robert J. *Economic Growth and Convergence*. 1994. Occasional Papers No. 46. San Francisco: International Center for Economic Growth.

Bishop, C. E. & W. D. Toussaint. *Introducción al Análisis de Economía Agrícola.1966* México: Limusa-Wiley.

Goldberger, Arthur S. *Topics in Regression Analysis*. 1968. New York: Macmillan.

Harberger, Arnold C. "The Dynamics of Inflation in Chile," 1963 en C. F. Christ, et al., *Measurement in Economics: Studies in Mathematical Economics and Econometrics in Memory of Yehuda Grunfeld*, pp. 219-50. Stanford University Press.

Johnston, J. *Análisis Estadístico de los Coste*s. 1966. Barcelona: Sagitario, S. A.

Little M. Thomas y Hills F. Jackson .1985. Métodos estadísticos para la Investigación en la agricultura. Editorial Trillas S. A. de C. V. Sexta reimpresión 270 p. México D.F.

Martínez Garza Ángel 1995. Diseños Experiméntales Métodos y elementos de teoría Editorial Trillas S.A. de C.V. Primera edición 756 p México D.F.

Murray R. Spiegel 1990. Estadística. Colección Mc Graw Hill /Interamericana de España S.A. Segunda Edición, Impreso en España. 555 p.

Okun, A. M. "Potential GNP: Its Measurement and Significance," *Proceedings* (Business and Economics Section), American Statistical Association, 1962, pp. 98-104.

Quiroga Víctor H. 1976. Manual práctico para el análisis de experimentos de campo. Editorial IICA - OIADIC Serie de publicaciones misceláneas No. 12 114 p. San José Costa Rica

Steel G.d. Robert y Torrie H. James. 1992 Bioestadística. Principios y Procedimientos. Cuarta Edición. Mc Graw Hill de México S.A. DE C.V. México 622 p.

_____ 1986. Bioestadística. Editorial impresiones S.A. de C.V. lago Chalco Segunda edición 622 p México D.F.

Suits, D. B. "The Demand for New Automobiles in the United States," *Review of Economics and Statistics*, 40 (1958): 273-80.

Wilfredo Caballero A. 1975. Introducción a la Estadística. Editorial Instituto Interamericano de Ciencias Agrícolas, Tercera Edición ,287 p. San José Costa Rica

.

Esta obra se terminó de imprimir.
1000 ejemplares Talleres
tipográficos COINFERDEZ
S.R.L. al 15 de Enero 2019
Pagador #7234
Oruro Bolivia