ESTADÍSTICA INFERENCIAL APLICADA A LA ADMINISTRACIÓN Y FINANZAS

© TOSCANO GUERRERO FRANCISCO EDUARDO
CEVALLOS RAMOS CARINA DEL ROCÍO
SANTILLÁN ESPINOZA DIEGO IVÁN
NARANJO VACA MYRIAM JOHANNA

CASA EDITORA DEL POLO

ESTADÍSTICA INFERENCIAL APLICADA A LA ADMINISTRACIÓN Y FINANZAS

© Toscano Guerrero, Francisco Eduardo

Cevallos Ramos, Carina Del Rocío

Santillán Espinoza, Diego Iván

Naranjo Vaca, Myriam Johanna

CASA EDITORA DEL POLO

Estadística inferencial aplicada a la administración y finanzas

Toscano Guerrero, Francisco Eduardo
https://orcid.org/0000-0002-3951-7774
Escuela Superior Politécnica de Chimborazo

Cevallos Ramos, Carina Del Rocío
https://orcid.org/0000-0001-6639-9577
Escuela Superior Politécnica de Chimborazo

Santillán Espinoza, Diego Iván
https://orcid.org/0000-0002-4213-1936
Escuela Superior Politécnica de Chimborazo

Naranjo Vaca, Myriam Johanna
https://orcid.org/0000-0002-4711-6575
Escuela Superior Politécnica de Chimborazo

CASA EDITORA DEL POLO

Casa Editora del Polo - CASEDELPO CIA. LTDA.
Departamento de Edición

Editado y distribuido por:

Editorial: Casa Editora del Polo
Sello Editorial: 978-9942-816
Manta, Manabí, Ecuador. 2019
Teléfono: (05) 6051775 / 0991871420
Web: www.casedelpo.com
ISBN: 978-9942-621-61-0
DOI: https://doi.org/10.23857/978-9942-621-61-0

Revisión, Ortografía y Redacción:
Lic. Jessica Mero Vélez

Diseño de Portada:
Michael Josué Suárez-Espinar

Diagramación:
Ing. Edwin Alejandro Delgado-Veliz

Director Editorial:
Dra. Tibisay Milene Lamus-García

Comité Científico Académico

Constancia de Arbitraje

La Casa Editora del Polo, hace constar que este libro proviene de una investigación realizada por los autores, siendo sometido a un arbitraje bajo el sistema de doble ciego (peer review), de contenido y forma por jurados especialistas. Además, se realizó una revisión del enfoque, paradigma y método investigativo; desde la matriz epistémica asumida por los autores, aplicándose las normas APA, Sexta Edición, proceso de anti plagio en línea Plagiarisma, garantizándose así la cientificidad de la obra.

Comité Editorial

Contenido

PROLOGO

En el apasionante cruce de la estadística, la economía y el análisis de series de tiempo, encontramos un conjunto de herramientas que permiten descifrar los misterios ocultos en los datos y tomar decisiones informadas en un mundo cada vez más impulsado por la información. Este libro es un faro que ilumina el camino a través de este intrincado terreno, ofreciendo una guía integral sobre cómo utilizar la regresión lineal, los modelos econométricos y el análisis de series de tiempo para comprender, predecir y modelar fenómenos económicos y financieros.

La estadística es el lenguaje de los datos, y la regresión lineal es una de sus herramientas más poderosas. En estas páginas, encontrarás una exploración profunda y clara de la regresión lineal, desde sus fundamentos hasta sus aplicaciones en la economía y las finanzas. Los autores han desglosado los conceptos complejos en explicaciones accesibles, utilizando ejemplos prácticos y ejercicios que te guiarán paso a paso en la construcción y evaluación de modelos de regresión.

Pero esto es solo el comienzo. Los modelos econométricos, que incorporan variables económicas y financieras, agregan un nivel adicional de sofisticación al análisis. Este libro te llevará a través de la teoría detrás de estos modelos y te mostrará cómo aplicarlos en situaciones del mundo real. Aprenderás a comprender y analizar la relación entre variables económicas, realizar pruebas de hipótesis y utilizar modelos econométricos para pronosticar tendencias y tomar decisiones estratégicas.

El análisis de series de tiempo, por su parte, es una herramienta esencial para comprender cómo los datos evolucionan a lo largo del tiempo. Este libro te llevará a través de la modelización de series temporales, desde la identificación de patrones estacionales y tendencias hasta la construcción de modelos que permitan realizar predicciones precisas en un mundo en constante cambio.

Los autores de esta obra han combinado con maestría la teoría con la práctica. A medida que avanzas en las páginas de este libro, encontrarás ejemplos del mundo real que ilustran cómo estas herramientas pueden utilizarse para abordar problemas económicos y financieros reales. Descubrirás cómo aplicar estos métodos para tomar decisiones más informadas en contextos económicos, desde la inversión hasta la política pública.

Este libro es una invitación a un emocionante viaje a través del análisis de datos económicos y financieros utilizando la estadística, la regresión lineal, los modelos econométricos y el análisis de series de tiempo. Tanto si eres un estudiante de economía en busca de una comprensión más profunda, un analista financiero que busca mejorar sus habilidades o un profesional de la estadística interesado en el mundo económico, aquí encontrarás una guía valiosa y completa.

En un mundo cada vez más impulsado por los datos,

la habilidad para entender y aprovechar estas poderosas herramientas es esencial. Este libro es tu brújula en este emocionante viaje. ¡Prepárate para embarcarte en un emocionante viaje hacia el mundo de la estadística aplicada y la toma de decisiones basada en datos!.

INTRODUCCIÓN

En un mundo cada vez más impulsado por datos y cifras, la estadística se ha convertido en el idioma universal para descifrar los misterios ocultos en los números y extraer conocimiento valioso. En este emocionante viaje a través de la estadística, exploraremos una de las herramientas más versátiles y poderosas que existen: la regresión lineal. Pero no nos detendremos ahí. También adentraremos en el vasto territorio de los modelos econométricos y el análisis de series de tiempo, donde los datos financieros y económicos se convierten en nuestro lienzo para la exploración y la predicción.

La regresión lineal es la base sobre la cual se construyen muchos de los métodos estadísticos más avanzados. En su esencia, es una técnica que nos permite comprender y modelar la relación entre una variable de interés y una o más variables predictoras. A través de esta relación, podemos hacer predicciones, evaluar hipótesis y tomar decisiones informadas en una amplia variedad de campos, desde la ciencia social hasta la ingeniería, y, como lo veremos aquí, en el apasionante mundo de la economía y las finanzas.

Los modelos econométricos, por su parte, son la culminación de la estadística aplicada a los datos económicos y financieros. Estos modelos permiten explorar y cuantificar las relaciones complejas que existen en estos dominios, teniendo en cuenta factores como la inflación, el crecimiento económico, las tasas de interés y muchas otras variables que afectan nuestra vida cotidiana. A medida que avanzamos en este libro, te mostraremos cómo utilizar modelos econométricos para comprender y predecir fenómenos económicos, evaluar políticas públicas y tomar decisiones estratégicas en el ámbito financiero.

El análisis de series de tiempo, por último, es la herramienta que nos permite explorar cómo los datos evolucionan a lo largo del tiempo. En un mundo donde el tiempo es un factor crítico, entender las tendencias, los ciclos y las estacionalidades en los datos temporales es esencial. Aquí, te guiaremos a través de la identificación de patrones en series de tiempo, la construcción de modelos y la realización de predicciones precisas en contextos económicos y financieros.

A lo largo de este libro, encontrarás un equilibrio entre la teoría y la práctica. Hemos diseñado ejemplos del mundo real y ejercicios que te ayudarán a aplicar estos métodos de manera efectiva. Nuestra meta es proporcionarte las habilidades y el conocimiento necesarios para abordar problemas reales con confianza y precisión.

En un mundo donde los datos son la moneda del siglo XXI, la estadística, la regresión lineal, los modelos econométricos y el análisis de series de tiempo se convierten en herramientas esenciales para la toma de decisiones, la investigación y la comprensión del mundo que nos rodea. Este libro es tu entrada a este emocionante

mundo, donde los números cobran vida y la estadística se convierte en tu aliada para explorar, entender y prever. ¡Bienvenidos a este viaje!

CAPÍTULO I

PROBABILIDADES

Definiciones. -

Probabilidad Valor que va de 0 hasta 1, inclusive, que describe la posibilidad relativa de que ocurra el evento.

Experimento Proceso que conduce a que ocurra una y solamente una de varias observaciones posibles.

Resultado
Un suceso particular proviene de un experimento.

Resultado Conjunto de uno o más resultados de un experimento.

Ejemplo:

Si nosotros tomamos un dado y deseamos tener números pares encontraríamos que:

Probabilidad: Que caiga una cara con un número par de puntos.

Experimento: Tirar el dado varias veces.

Resultado 1, 2, 3, 4, 5,6.

Espacio Muestral→S= [2, 4, 6]

Evento: A= [2, 4, 6]

Utilizaremos también el símbolo $\sum$ **"SUMATORIA"** que simboliza una serie de sumas en una lista de datos que de otra manera seria muy largo de enumerar.

Ejemplo:

Vamos a sumar datos obtenidos del 1 al 20 de modo que tendríamos que indicar la forma que sumaremos, es decir, $X_{1+}X_{2+}X_{3+}X_{4+}X_{5}$............$X_{20.}$

Entonces utilizando SUMATORIA quedaría $\sum_1^{20}xi$ donde i=1.

Variable — Es una función que asigna valores numéricos a los puntos muéstrales de un experimento.

Frecuencia — De un evento, en un número finito de repeticiones o de observaciones de un suceso, la frecuencia es el número de veces en que el evento se cumple.

Frecuencia — Es la proporción en que dicho evento se cumple y se obtiene compactándolo de la siguiente manera

$$fr(A) = \frac{f(A) * 100}{N}$$

Dónde:

- **fr(A)**=Frecuencia relativa del evento

- **f(A)**=Frecuencia del evento A.

- **N**=Número de observaciones

También:

- **M**=Características o atributos.

- **S**=Espacio Muestral.

- **A**=Conjunto de todas ellas.

Ejemplo:

Encontrar la frecuencia relativa en base a edades de estudiantes de 5to semestre de la Carrera de Finanzas de la ESPOCH.

$$fr(A) = \frac{f(A) * 100}{N}$$

X	f	fr.
1	20	8.81
2	20	8.81
3	20	8.81
4	21	9.25
5	22	9.69
6	21	9.25
7	20	8.81
8	22	9.69
9	21	9.25
10	20	8.81
11	20	8.81
Total	227	

$$fr = \frac{fx}{N}.100$$

$$fr = \frac{20}{227}.100$$

$$fr = 8.81$$

Supongamos que un salón de clases hay 20 alumnos de los cuales algunos son quiteños y otros no lo son.

¿Cuál es la probabilidad de que si se toma uno de ellos al azar sea quiteño, "véase que el espacio muestral tendrá un igual a 20 (suceso que puede ocurrir de 20 maneras; suponiendo que hubiera 16 quiteños?

Datos:

P(A)=…

N= 20

M= 16

$$P(A) = \frac{M}{N}$$

$$P(A) = \frac{16}{20}$$

$$P(A) = 0.8$$

1.1.Probabilidad Clásica

En este tema vamos a ordenar la variable y con los datos de frecuencia iremos sumando en el casillero que corresponda a las veces que se repita.

Ejemplo:

Con las notas de los estudiantes de 5to semestre de contabilidad vamos a organizar la tabla de frecuencia relativa.

NOTAS X	F	FR
5		
6		
7	3	21,43
8	2	14,29
9	5	35,71
10	4	28,57
	N=14	

DEFINICIÓN

$$Fr = \frac{fx}{N}.100$$

$$Fr = \frac{3}{14}.100$$

$$Fr = 21.43$$

Un suceso puede ocurrir de **N** maneras mutuamente exclusivas e igualmente verosímiles si **M** de ellas posee una característica o atributo de **A** el conjunto de todas ellas será el Espacio Muestral (**S**) y el conjunto que posee el atributo será el evento A.

Entonces podemos decir que la probabilidad del evento A= a la característica sobre el número de datos.

$$P(A) = \frac{M}{N} \quad \frac{(\#\ de\ casos\ Favorables)}{(\#\ \text{de casos Posibles})}$$

Ejemplo:

En el ejemplo anterior indicamos el proceso de los 20 estudiantes que son quiteños.

1.2.Definicion Impírica de Probabilidades

Un suceso puede ocurrir de diversas maneras mutuamente exclusivas (no necesariamente iguales verosímiles), y alguna de estas maneras posee un atributo a, entonces la probabilidad del evento p(a) será igual al límite de su frecuencia relativa.

Es decir tenemos: $PA = \frac{\lim.M}{N}$ en donde:

N=Número de repeticiones

n= Número de veces que han cumplido A en N repeticiones.

Lim.= La precisión en la administración del valor verdadero de PA

1.3.Aplicaciones:

Supóngase que se tiene un juego ordinario de barajas las cuales constan de 52 cartas repetidas en 113 de cada una de las figuras (4 figuras de diamante, trébol, corazón negro y corazón) que van desde el A, 2, 3,4…..J, K. Si queremos extraer una de ellas podemos decir.

a) ¿Cuál es la probabilidad de que la baraja extraída sea el AS de diamante?

b) ¿Cuál es la probabilidad de que la carta extraída sea AS?

a) $PA = \frac{1}{52} = 2\%$

b) $PA = \frac{1}{13} = 7.7\ \%$

1.4.Desarrollo

Observando lo anterior se puede ver que el segundo problema se tiene mayor probabilidad debido a la información adicional de la que se disponía es decir que se tiene una probabilidad condicional con lo que podemos hacer una intersección entre A y B y obtendremos la siguiente fórmula.

$$P(\frac{A}{B}) = \frac{P(A \cap B)}{PB}$$

Reemplazando:

$$P(\frac{A}{B}) = \frac{P(\text{AS de diamante})}{P(\text{Diamante})}$$

$$P(\frac{A}{B}) = \frac{1/52}{13/52}$$

$$P(\frac{A}{B}) = \frac{1}{13}$$

Conclusión

La probabilidad de sacar el AS en cualquiera de las 4 figuras.

La probabilidad juega un papel importante en la estimación en el número de unidades defectuosas en un proceso de fabricación, la probabilidad de recibir pagos sobre cuentas por cobrar y las ventas potenciales de un nuevo producto, incluso los apostadores profesionales, en eventos deportivos, debe tener una comprensión sólida de la teoría de la probabilidad.

Ejemplo

Ciclismo, peles de gallos campeonato de básquet de la NBA, tenis entre otros.

Una probabilidad se expresa como una fracción decimal tal como 0.70; 0.80; 0.25 también puede indicarse como una fracción 7/4; 8/3; 9/6.

Es decir, cuando más se aproxime a 0 (cero) una probabilidad es más improbable de que ocurra el evento respectivo; cuando más se acerca a 1 es más probable que suceda dicho evento, para lo cual indicaremos en la siguiente tabla:

0	0,25	0,50	0,75	1
Probabilidad de que el sol desaparezca este año.	Probabilidad de que el Equipo de MushucRuna le gane al Emelec en el campeonato 2014	Probabilidad de que al lanzar una moneda caiga “CARA” al tirar una vez.	La probabilidad de que descienda de categoría el Mushcu-Runa y el Cuenca	Probabilidad de que llueva en Ambato este año.

1.5.Enfoque de Probabilidad

Se analizaron 2 enfoques del análisis probabilístico, especialmente los puntos de vista **objetivo y subjetivo.**

La probabilidad objetiva se puede dividir en 2:

1) Probabilidad Clásica

2) Probabilidad Empírica

1.6.Probabilidad Clasica

Se basa en la consideración de que los resultados de un experimento son igualmente posibles. Empleando el punto de vista clásico la probabilidad de que suceda un evento se calcula dividiendo el número de resultados favorables entre el número total de resultados posibles, es decir:

$$P(A) = \frac{\#\ \text{Resultados Favorables}}{\#\ \text{Resultados Posibles}}$$

Ejemplo:

Considere el ejemplo de lanzar un dado común. ¿Cuál es la probabilidad del evento de que caiga un número par?

Análisis:

Hay 3 resultados favorables en el conjunto de 6 resultados posibles igualmente probables.

Aplicando fórmula tenemos:

$$P(A) = \frac{3}{6}$$
$$P(A) = \frac{1}{2}$$
$$P(A) = 0.5$$

Si uno de varios elementos puede ocurrir cada vez, se dice que los eventos son mutuamente excluyentes.

- MUTUAMENTE EXCLUYENTES: La ocurrencia de un evento implica que ninguno de los otros eventos puede ocurrir al mismo tiempo.

Si un experimento tiene un conjunto de eventos que comprenden a todos los resultados posibles, tales eventos como: "cae un número par" y "cae un número impar", cuando se lanza un dado entonces el conjunto de eventos es colectivamente exhaustivo.

- COLECTIVAMENTE EXHAUSTIVO: Por lo menos uno de los eventos puede ocurrir cuando se realiza el experimento.

Razonamiento

Si lanzamos varias veces una moneda en el evento puede caer varias caras y varios sellos.

1.7. Probabilidad Impírica

Otro modo de definir la probabilidad es basarse en las frecuencias relativas; la probabilidad de que un evento ocurra se determina observando en que fracción de tiempo sucedieron eventos semejantes en el pasado, es decir tendríamos la siguiente fórmula:

$$P(A) = \frac{\text{Nº DE VECES QUE OCURRIO EL EVENTO EN EL PASADO}}{\text{TOTAL DE OBSERVACIONES}}$$

P(A)=Probabilidad de que suceda el evento.

Ejemplo:

Se efectuó un estudio con 751 egresados de la carrera de Finanzas, en la ESPOCH. Este experimento reveló que 383 de los 751 egresados no estaban empleados de acuerdo con su principal área de estudio. Por ejemplo, Un egresado especializado en Finanzas es ahora gerente de Mercadotecnia en una empresa de tomates. ¿Cuál es la probabilidad de que los egresados en Finanzas laboren en un área distinta a la de sus estudios universitarios?

$$P(A) = \frac{\text{N° DE VECES QUE OCURRIO EL EVENTO EN EL PASADO}}{\text{TOTAL DE OBSERVACIONES}}$$

$$P(A) = \frac{383}{751}$$

$$\mathbf{P(A)} = 0.51 = 51\ \%$$

Conclusión

Los egresados pueden o no ejercer su profesión.

1.8.Probabilidad Subjetiva

Si existe poca o ninguna experiencia en la cual se puede basar en la probabilidad, puede determinarse en forma subjetiva. Fundamentalmente esto significa evaluar las opiniones disponibles y otra información para después estimar o asignar la probabilidad.

CONCEPTO

Es la probabilidad que suceda un evento especifico, que es asignado por una persona basándose en cualquier información

que esté disponible.

Ejemplo:

✓ Estimar la probabilidad que los estudiantes de 5to contabilidad paralelo "C" tenga la calificación de 10 en estadística.

✓ Evaluar la probabilidad de GOODYEAR pierda su primer lugar en ventas de neumáticos frente a GENERAL TIRE ECUADOR en el año.

✓ Pongamos algunas reglas de probabilidad.

1.9.Regla de Adicion

La regla especial de adición, los eventos deben de ser mutuamente excluyentes, recuerde que mutuamente excluyentes significa:

Que cuando ocurre un evento ninguno de los otros puede suceder al mismo tiempo.

Ejemplo:

Hagamos el experimento de tirar un solo dado con los eventos mayores que 4 y menor que 2.

Si el resultado se encuentra en el primer grupo (4, 5, 6) pero tampoco puede estar en el 2do grupo (2 y 1).

En un producto industrial en la producción de neumáticos no puede ser defectuoso y satisfactorio al mismo tiempo.

Si los 2eventos A y B son mutuamente excluyentes, la regla especial de la adición indica que la probabilidad que ocurra uno u otro evento es igual a la suma de sus probabilidades, es decir:

P(A) o (B)=P(A o B)=PA+PB

Para eventos de 3 condiciones es decir A o B o C tendríamos la siguiente regla:

P(A) o (B) o (C)=P¨(A)+P(B)+P(C)

Para entender mejor esta regla nos basaremos en el siguiente ejemplo:

Una máquina automática de nombre PIRELLI llena de bolsas de plástico con mezcla de frijoles, brócolis y otras legumbres la mayor parte de las bolsas contienen un peso correcto pero debido a las ligeras variaciones en el tamaño de las verduras, un paquete puede tener un peso ligeramente menor o mayor. Una verificación de 4.000 paquetes que se llenaron en el mes pasado reveló lo siguiente:

PESO	EVENTO	N° DE PAQUETES	PROBABILIDAD DE OCURRENCIA
Menor	A	100	0.025
Satisfactorio	B	3.600	0.900
Mayor	C	300	0.075
TOTAL	**4000**		**1,00**

$$\mathrm{P(A)} = \frac{100}{4000}$$

$$\mathrm{P(A)} = 0.025$$

Utilizando la Fórmula:

Conclusión

Estaremos hablando de un evento exhaustivo.

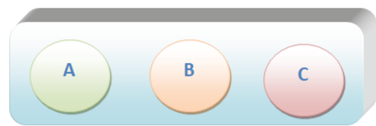
1.10. Diagrama de Venn

J. VENN ideo un diagrama para representar gráficamente el resultado de un evento.

Para elaborar un diagrama de Venn primero se delimita un espacio en el plano que representará todos los resultados posibles.

Un evento se representa mediante un círculo cuya área es proporcional a la probabilidad del evento y se dibuja dentro de un rectángulo.

Aplicando Venn en el ejercicio tenemos:



~=No es

P(~A)=No A

La probabilidad de que una bolsa de legumbres mixtas tenga menos peso la probabilidad del evento A más la probabilidad de que no sea una bolsa con peso menor; debe ser lógicamente igual a 1.

Esto se expresa mediante la siguiente fórmula:

P(A)=P(A)+P(~A)=1

En base a esta fórmula podemos expresar expresar que se transforma en una regla de complemento, es decir:

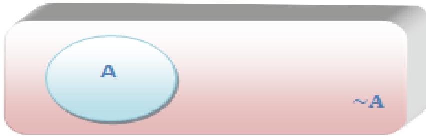
REGLA DE COMPLEMENTO

P(A)=1-P(~A)

Conclusión

Observamos que los eventos A y ~A son mutuamente excluyentes y colectivamente exhaustivos.

Utilizando el diagrama de Venn que ilustre la regla de complemento sería:



Para la resolución del ejercicio que estamos analizando tenemos la probabilidad de que el peso de la bolsa de legumbres no sea el correcto es igual a la probabilidad de que su peso sea mayor, más la probabilidad de que su peso sea menor.

P(A o C)=PA+PC

P(A y C)=0,025+0,075

P(AyC)=0,1

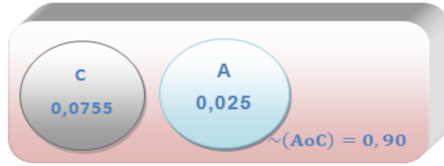
El peso satisfactorio

PB=1-(PA+PC)

PB=1-(0,025+0,075)

P(B)=0,90

APLICANDO VENN TENDRÍAMOS:



Los resultados de un experimento pueden no ser mutuamente excluyentes por ejemplo supongamos que la comisión de turismo de Florida seleccionó una muestra de 200 turistas. La muestra reveló que 120 fueron a Disney World y 100 fueron a Garden. 60 personas de cada 200 visitaron ambas atracciones.

¿Cuál es la probabilidad de que una persona seleccionada haya visitado Disney World o Garden?

Si empleamos la regla general de adición la probabilidad de seleccionar turistas a Disney world es:

$$\mathbf{P(DW)} = \frac{120}{200}$$

$$\mathbf{P(DW)} = 0.60$$

De manera similar la probabilidad de que un turista haya ido a Garden es:

$$P(G) = \frac{100}{200}$$

$$\boldsymbol{P(G)} = 0.50$$

- La suma de estas 2 probabilidades nos da 1,10
- Se sabe que la probabilidad no debe ser mayor que 1.
- La explicación puede darse de que muchos turistas visitaron ambas atracciones y se está contando 2 veces, para encontrar solución a este cálculo tenemos:

(DoG)=PD+PG-P(DoG)

P(DoG)=0,60+0,50-0,30

P(DoG)=0,80

Conclusión

Cuando 2 eventos ocurren simultáneamente a la probabilidad respectiva se le denomina conjunta. (60/200=0.30)

Probabilidad Conjunta: Es la medina de probabilidad que evalúa de que 2 o más eventos ocurran en forma simultánea.

1.11.Regla General De La Adicion

$$\mathbf{P(AoB)} = \mathrm{P(A)} + \mathrm{P(B)} - \mathrm{P}(\tfrac{\mathrm{B}}{\mathrm{A}})$$

U = o ∩ = y

$$\mathbf{P(A\ U\ B)} = \mathrm{P(A)} + \mathrm{P(B)} - \mathrm{P} * (\mathrm{A} \cap \mathrm{B})$$

"o" = INCLUSIVO

Aplicación

¿Cuál es la probabilidad de que una carta elegida al zar sea rey o reina de corazones?

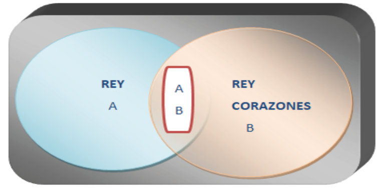
CARTA	PROBABILIDAD	EXPLICACIÓN
REY	PA	4/52
REINA CORAZONES	PB	13/52
REY DE CORAZONES	P(A y B)	1/52

$$\mathbf{P(AoB) = P(A) + P(B) - P(\tfrac{B}{A})}$$

$$\mathbf{P(AoB) = \frac{4}{52} + \frac{13}{52} - \frac{1}{52}}$$

$$\mathbf{P(AoB) = 0,31}$$

UTILIZANDO EL DIAGRAMA DE VENN



Aplicaciones

1.Algunas personas en Riobamba están a favor de la Reducción de los beneficios del Seguro Social, al fin de lograr un presupuesto Equilibrado, en tanto que otras están en Contra, se seleccionaron dos personas y se registraron sus opiniones. Mencione los resultados posibles.

$$\text{Clásica } A = \frac{O.P.F}{O.P.P}$$

$$A = \frac{1}{2}$$

$$A = 0.50$$

Conclusión: Puede ser favorable o no puede ser favorable

1. Un inspector de Control de Calidad elegía una pieza fabricada para probarla. Posteriormente se establece si la parte se aceptare para o se desecha. Después se prueba otra. Mencione todos los Resultados Posibles.

a) Acepta

a) Se repara

a) Se desecha

Conclusión: Se presenta tres eventos posibles.

1. Una encuesta de un grupo de estudiantes de 34 en una escuela de Administradores, se tiene la siguiente selección de Carreras Profesionales. Suponga que se selecciona un Estudiante y se considera su elección Profesional.

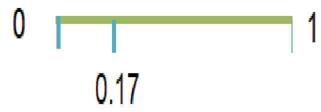
a) Cuál es la probabilidad de que él / ella estudie la carrera de Administración

b) Que concepto de probabilidad utiliza para ser la investigación.

Contaduría	10
Finanzas	5
Sistemas de Información	3
Administración	6
Mercadotecnia	10
TOTAL	34 estudiantes

a) $P = \frac{1}{6}$

$P = 0.17$

b) 

La probabilidad es clásica porque es excluyente ya que no puede suceder 2 eventos al mismo tiempo.

4. El departamento de vía pública, en la ciudad de Whitehouse, Illiois, está considerando ampliar la avenida Indiana a tres carriles. Antes de tomar una decisión, se preguntó a 500 ciudadanos si aprobaban la ampliación.

a) ¿Cuál es el experimento?

Consultar a 500 ciudadanos si apoyan la ampliación de la avenida Indiana a 3 carriles.

b) ¿Cuáles son algunos de los eventos posibles?

No se amplía, Si se amplía, No contestan.

b) Menciones 2 resultados posibles.

Contestan Si o contestan No.

5. El presidente del comité directivo de la empresa Rudd Industries pronunciará mañana un discurso ante los accionistas de la compañía, explicando su opinión en lo concerniente a que dicha corporación debe fusionarse con la empresa Zimmerman Plastics. Ha recibido seis cartas por correo respecto a este asunto, y está interesado en conocer el número de remitentes externos que están de acuerdo con él.

a) ¿Cuál es el experimento?

Fusión de la empresa Rudd Industries con la empresa Zimmerman.

b) ¿Cuáles son algunos de los posibles eventos?

Si es favorable, No es favorable.

c) ¿Mencione 2 resultados posibles?

5 remitentes que estén de acuerdo con él y 1 que no esté de acuerdo.

3 remitentes que estén de acuerdo con él y 3 que no esté de acuerdo.

6. En cada uno de los casos siguientes indique si se utiliza la probabilidad clásica, empírica y subjetiva.

a) Una jugadora de basquetbol realiza 30 canastas (o encestes) en 50 tiros por faltas.

La probabilidad de que efectúe bien el siguiente tiro es 0,6.

30/50=Clásica porque hay números de resultados favorables/ número de resultados de resultados posibles.

a) Se formó un comité de alumnos integrado por 7 miembros para estudiar asuntos ambientales. ¿Cuál es la probabilidad de que 1 de ellos sea elegido como el vocero?

1/7=0,14= Probabilidad subjetiva.

a) Considere que usted que compra uno de los 5 millones de billetes que se vendieron en el sorteo de lotería. ¿Cuál es la probabilidad de que gane el premio mayor de un millón de dólares?

1/5.000.000 =0.00002 % Probabilidad Clásica porque hay números de resultados favorables/número de resultados de resultados posibles.

a) La probabilidad de que ocurra un sismo en el norte de California en los próximos 10 años, es 0,80 80%.

Probabilidad Empírica

Conclusión:

Aplicando la probabilidad objetiva basándose en ejercicios prácticos.

7. Una empresa concederá un ascenso a 2 empleados de un grupo de 6 hombres y 3 mujeres.

a) Mencione los resultados de este experimento si hay interés especial relacionado con la igualdad de género sexual.

H=2/6

H=0,33

M=2/3

M=0,67

Conclusión:

Si hay interés especial relacionado con la igualdad de género

a) ¿Qué concepto de probabilidad utilizaría para calcular esas probabilidades?

El concepto de probabilidad clásica porque hay número de resultados favorables y números de resultados posibles.

8. Hay 52 cartas en una baraja americana.

a) ¿Cuál es la probabilidad de que la primera carta que se saque sea una As de espadas?

a. 1/52=0,02

b)¿Cuál es la probabilidad de que la primera carta seleccionada sea el sota de espadas?

c)1/52=0,02

Conclusión:

Existe la misma probabilidad de A.

a) ¿Qué concepto de probabilidad ilustran los incisos a y b.

Probabilidad clásica porque tiene eventos favorables y eventos posibles.

9. Se lanza un solo dado.

a) ¿Cuál es la probabilidad de que caiga un 2?

P(A)=1/6

P(A)=0,17 17 %

Conclusión:

Existe una 17% de probabilidad de que salga en 2 al lanzar un solo dado.

a) ¿Qué concepto de probabilidad se ilustra con esto?

Probabilidad clásica porque ésta se calcula dividiendo el número de resultados favorables para el número de resultados posibles.

a) ¿Los resultados para los números del 1 al 6 son igualmente probables y mutuamente excluyentes? Explique.

Si pueden ser mutuamente excluyentes porque la ocurrencia de un evento implica que ninguno de los otros eventos puede ocurrir al mismo tiempo.

10. Se seleccionó una muestra de 40 ejecutivos para que respondieran a un cuestionario de prueba. Una pregunta relacionada con aspectos ambientales requiere una respuesta de sí o no.

a) ¿Cuál es el experimento?

Ejecutivos que van a responder un cuestionario de prueba sobre aspectos ambientales

a) Mencione un evento posible.

La respuesta puede ser sí o no.

a) Diez de los 40 ejecutivos respondieron que “si”. Con base en las respuestas de la muestra, ¿Cuál es la probabilidad de que la respuesta de un ejecutivo sea afirmativa?

P(A)= 10/40

P(A)=0,25

Conclusión:

No hay la probabilidad de que la respuesta sea afirmativa.

a) ¿Qué concepto de probabilidad ilustra esto?

Probabilidad Clásica porque hay número de eventos favorables y numero de eventos posibles.

a) ¿Cada uno de los resultados posibles son igualmente probables y mutuamente excluyentes?

Si son mutuamente excluyentes porque no pueden suceder varios eventos a la vez, es decir, pueden ser sí o no pero no los dos al mismo tiempo.

Aplicación

Ejercicio 5.4: Como parte de un programa de salud para los empleados de la empresa General Concrete, se efectúan anualmente exámenes físicos de rutina. Se descubrió que 8% de los empleados necesitaban zapatos correctivos; 15%, un trabajo dental importante; y 3% requerían tanto zapatos correctivos como un trabajo dental mayor.

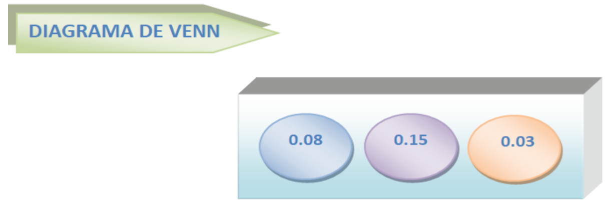
a) ¿Cuál es la probabilidad de que un empleado seleccionado al azar necesite calzado correctivo o un trabajo dental considerable?

b) Muestre esta situación con un diagrama de VENN

P (A o B)=PA+PB-(PA y B)

P (Ao B)=0.08+0.15+0.03

P (A)=0.20



Ejercicios

15. Los eventos A Y B son mutuamente excluyentes. Supóngase que P(A)=0.30 y P (B)=0.20. ¿Cuál es la probabilidad de que ocurra A o B? ¿Cuál es la probabilidad de que no suceda ni A ni B?

Conclusión:

Puede hacerse A o puede hacerse B.

P (A o B)=PA+PB

P (A o B)=0.30+0.20

P (A o B)=0.50

Conclusión:

Puede ocurrir B o puede ocurrir

16. Los eventos X y Y son mutuamente excluyentes. Supóngase que P(X)=0.05 y P (Y)=0.02. ¿Cuál es la probabilidad de que

ocurra X o Y? ¿Cuál es la probabilidad de que no suceda X ni Y?

$$P\,(A \text{ o } B) = 1-(PA+P\,B)$$

$$P\,(A\text{o } B) = 1-(0.05+0.02)$$

$$P\,(A \text{ o } B) = 1-0.50$$

$$P\,(A \text{ o } B) = 0.50$$

Conclusión:

No puede ocurrir ni X ni Y al mismo tiempo.

$$P\,(X \text{ o } Y) = 1-(PX+PY)$$

$$P\,(X \text{ o } Y) = 1-(0.05+0.02)$$

$$P\,(X \text{ o } Y) = 1-0.07$$

$$P\,(A) = 0.93$$

Conclusión:

Puede ocurrir X o puede ocurrir Y

17. Un estudio en 200 cadenas de tiendas de comestibles reveló estos ingresos (en dólares), después del pago de impuestos.

Ingreso (en dólares) Después de Impuestos	Cantidad de Empresas
Menos de 1 millón	102
De 1 millón a 20 millones	61
De 20 millones a mas	37

a) ¿Cuál es la probabilidad de que una cadena determinada tenga menos de 1 millón (de dólares), de ingresos después de pagar impuestos?

b) ¿Cuál es la probabilidad de que una cadena de tiendas seleccionada al azar tenga un ingreso entre 1 millón y 20 millones, o un ingreso de 20 millones o más? ¿Qué regla de probabilidad se aplicó?

$$P\,(A) = \frac{\text{N° RESULATDOS FAVORABLES}}{\text{N° RESULTADOS POSIBLES}}$$

$$P\,(A) = \frac{102}{200}$$

$$P\,(A) = 0.51$$

Como son mutuamente excluyentes o colectivamente exhaustivos

$$P\,(B \text{ o } C) = PB + PC$$

$$P\,(B \text{ o } C) = \frac{61}{200} + \frac{37}{200}$$

$$P\,(B \text{ o } C) = 0.49$$

$$P\,(B) = 1 - PA$$

$$P\,(A\text{o } B) = 1 - \frac{102}{200}$$

$$P\,(A) = 0.49$$

18. Un estudio de opiniones de diseñadores en lo referente al color primario más conveniente para aplicar en oficinas ejecutivas indico que:

Color primario	Numero de opiniones
Rojo	92
Naranja	86
Amarillo	46
Verde	91
Azul	37
Índigo	46
Violeta	2
	400

a) ¿Cuál es el experimento?

b) ¿Cuál es el evento posible?

c) ¿Cuál es la probabilidad de seleccionar una respuesta específica y descubrir que el diseñador prefiera rojo o verde?

d) ¿Cuál es la probabilidad de que un diseñador no prefiera el amarillo?

a) Estudio de opiniones de diseñadores para la aplicación de colores primarios.

b) Cualquier color primario que prefiere la gente.

Color primario	P(X)	Numero de opiniones
Rojo	A	92
Naranja	B	86
Amarillo	C	46
Verde	D	91
Azul	E	37
Índigo	F	46
Violeta	G	2
	TOTAL	400

c) La suma de todas las opiniones es de 400 por lo tanto para descubrir una respuesta específica y descubrir que el diseñador prefiere entre rojo y verde usamos la siguiente regla: P(A o D)=P(A)+P(D)

Son mutuamente excluyentes

$$P(A \text{ o } D)=\frac{92}{400}+\frac{91}{400}=0.46$$

Para la probabilidad de que el diseñador no prefiera el amarillo usamos la regla del complemento:

$$P(C)+P(\sim C)=1$$

$$P(\sim C)=1-\frac{46}{400}=0.86$$

19. El presidente de una junta de Directores dice: “Hay 50% de probabilidad de que esta compañía tenga utilidades, 30%de que

quede a nivel (Punto de Equilibrio), y 20%de que pierda dinero el siguiente trimestre”.

a) Utilice una regla de adición para encontrar la probabilidad de que no se pierda dinero el próximo trimestre.

b) Use la regla del complemento para encontrar la probabilidad de que no pierda dinero el próximo trimestre.

P (A o B)=PA+PB

P (A o B)=0.50+0.30

P (A o B)=0.80

P (A o B)=1-(PA+PB)

P (A o B)=1-(0.50+0.30)

P (A o B)=1-0.80

P (A o B)=0.20

Conclusión:

Es un evento exhaustivo.

20. Suponga que la probabilidad de que usted obtenga una calificación de A en el curso de esta materia es 0.25, y la de que tenga una B, es 0.50. ¿Cuál es la probabilidad de que su calificación sea mayor que una C?

P (A o B)=PA+PB

P (A o B)=0.25+0.50

P (A o B)=0.75

P (C)=1-(PA+PB)

P (C)=1-0.75

P (C)=0.25

Conclusión:

La probabilidad de una calificación mayor es favorable a 0.25

21. Se tira un solo dado. El evento A es "sale un 4", el evento B es "sale un numero par", y el evento C corresponde a "sale un número impar". Considere todas las parejas posibles de estos eventos e indique si son mutuamente excluyentes. Después identifique si son complementarias.

$$A = \frac{1}{6}$$

$$B = \frac{3}{6}$$

$$C = \frac{3}{6}$$

Conclusión:

Son mutuamente excluyentes porque no puede suceder una o más cuentas a la vez.

22. Se lanza dos monedas al aire. Si A es el evento "caen dos caras" y 6 es el evento "caen dos cruces", ¿Son A y 6 mutuamente excluyentes? ¿Son eventos complementarios?

Son eventos mutuamente excluyentes porque no pueden ocurrir dos eventos o más al mismo tiempo.

23. Las probabilidades de los eventos A y 6 son 0.20 y 0.30, respectivamente. La probabilidad de que tanto A como B

ocurran es 0.15. ¿Cuál es la probabilidad de que suceda A o B?

P (A o B)=PA+PB-P(A y B)

P (A o B)=0.20+0.30-0.15

P (A o B)=0.3

Conclusión:

La probabilidad de que ocurra A o B no es favorable.

24. Sea P(X)=0.55 y P (Y)=0.35. Supóngase que la probabilidad de que ambos ocurran es 0.20. ¿Cuál es la probabilidad de que ocurran X o Y?

P (X o Y)=PX+PY-P(X y Y)

P (X o Y)=0.55+0.35-0.20

P (X o Y)=0.70

Conclusión:

La probabilidad de que ocurra X o Y es favorable

25. Supóngase que los des eventos A y B son mutuamente excluyentes. ¿Cuál es la probabilidad de su ocurrencia conjunta?

La probabilidad es nula

6. Un estudiante está tomando dos cursos, Historia y Matemáticas. La probabilidad de que apruebe el curso de Historia es 0.60, y la de que apruebe el curso de Matemáticas,

es 0.70. La probabilidad de que apruebe ambos es 0.50. ¿Cuál es la probabilidad de que pase al menos uno?

P (X o Y)=PX+PY-(X y Y)

P (X o Y)=0.60+0.70-0.50

P (X o Y)=0.80

Conclusión:

La probabilidad es que pase una sola materia con un porcentaje del 80%.

27. Una encuesta a ejecutivos de alto nivel en EUA, reveló que 35% leen con regularidad la revista Time, 20% leen Newsweek, y 40% leen U.S.NewsWorldReport. Un 10% lee tanto Time como U.S.NewsWorldReport.

a) ¿Cuál es la probabilidad de que un ejecutivo determinado lea Time, o bien, U.S. News World Report con regularidad?

b) ¿Cómo se denomina a la probabilidad con valor de 0.10?

c) ¿Los eventos son mutuamente excluyentes? Explique la respuesta.

P (A y C)=PA+PC-P(A o C)

P(AyC)=0.35+0.40-0.10

P(AyC)=0.65

b) Se denomina probabilidad conjunta

c) No son excluyentes porque pueden leer las 2 revistas al mismo tiempo

28. Un estudio realizado por el servicio de Parque Nacionales (de Estados Unidos) reveló que 50% de los vacacionistas que viajan a la región de la Montañas Rocosas van al Parque Yellowstone, 40% visitan Tetons, y 35% van a ambos sitios.

a) ¿Cuál es la probabilidad de que un vacacionista visite al menos una de estas atracciones?

b) ¿Cómo se denomina a la probabilidad 0.35?

c) ¿Los eventos son mutuamente excluyentes? Explique su respuesta

P (X o Y)=PX+PY-P(X o Y)

P (X o Y)=0.50+0.40-0.35

P (X o Y)=0.55

Si hay probabilidad de que un vacacionista visite al menos de una de estas atracciones

a) Probabilidad conjunta

1.12.Regla de Multiplicación

La regla de multiplicación en este caso especial requiere que los eventos A y B sean independientes. Dos eventos son independientes si la ocurrencia de uno no altera la probabilidad de que suceda el otro.

INDEPENDIENTE La ocurrencia de un evento no tiene efecto en la probabilidad de la ocurrencia de cualquier otro evento.

- ✓ La Regla Especial quedaría:

P(A y B) = (PA) x (PB)

- ✓ Para tres eventos quedaría

P (A, B,C) = (PA)x(PB)x(PC)

Aplicación:

De una encuesta realizada por American Automóviles se encontraron que el 60% de sus socios hicieron alguna reservación en una línea aérea el año pasado. Se toman dos integrantes al azar.

¿Cuál es la probabilidad de que ambos hayan hecho una reservación en la misma línea aérea?

PA = 60% ----> Probabilidad de reservación de vuelo aerolínea

PB = 60% ----> Probabilidad de reservación de vuelo aerolínea

60% = 0.6

PA y B = (PA) x (PB)

= (0.6) (0.6)

= 0.36 36 %

Conclusión:

1. La probabilidad de que 2 integrantes no hicieron la reservación en la aerolínea.

Utilizando el diagrama de VEEN todos los resultados posibles

los podemos mostrar.

R = Reservar

NR = No reservar

2. En una encuesta realizada por Análisis Estudiantil a estudiantes de una escuela Dante Mars se encontró que el 45% de ellos se cambiara de escuela en el nuevo período electivo. Se tomaron al azar 3 estudiantes.

¿Cuál es la probabilidad de que los tres hayan ingresado a la misma escuela?

PA = 45% Probabilidad de ingreso a la escuela Milenium

PB = 45% Probabilidad de ingreso a la escuela Milenium

PC = 45% Probabilidad de ingreso a la escuela Milenium

P A, B, C =(PA) x (PB) x (PC)

= (0.45) (0.45) (0.45)

= 0.091 9.1 %

3.En una escuela de cocina se toma el 30% de estudiantes para que pongan en práctica la elaboración de un postre especial, si tres estudiantes son seleccionados al azar.

¿Cuál es la probabilidad de que los tres estudiantes desean utilizar gelatina en su postre especial?

30% = 0.30

PA = 30%

Probabilidad de preparar el postre con gelatina

PB = 30%

Probabilidad de preparar el postre con gelatina

PC = 30%

Probabilidad de preparar el postre con gelatina

P A, B, C = (PA) x (PB) x (PC)

= (0.3) (0.3) (0.3)

= 0.027 2.7 %

Conclusión:

La probabilidad que ninguno de los tres utiliza gelatina.

Aplicación:

Se va a realizar una rifa de una cocina y una licuadora en el centro comercial ambato se toma de todos los asistentes el 50% para dicho evento. Si seleccionamos 2 personas al azar.

¿Cuál es la probabilidad de que uno de los dos se haga acreedor a los dos premios?

PA1:50%= Probabilidad de ganarse el premio

PA2:50%= Probabilidad de ganarse el premio

P (A y B)=(PA) (PB)

P (A y B)=(0.5)(0.5)

P (A y B)=0.25

Conclusión:

La probabilidad es que no se pueda ganar los 2 premios.

NOTA: Cuando tenemos que dos eventos no son independientes, se dice que obviamente son dependientes.

Para demostrar la dependencia tenemos la siguiente aplicación.

Aplicación :

Supongamos que tenemos en una caja 10 rollos de película fotográfica y que se sabe que 3 están defectuosos se selecciona uno. Es obvio que la probabilidad de escoger un rollo con defecto es de 3/10, la probabilidad de seleccionar uno satisfactorio es de 7/10. Después se elige un segundo rollo de la caja sin devolver el primero a esta. La probabilidad de que sea defectuoso depende de que si el primer rollo seleccionado no fue aceptable.

La probabilidad de que el segundo rollo tenga defecto es 2/9; si el primer rollo fue defectuoso (quedaría solo dos rollos defectuosos en la caja que contiene nueve piezas), 3/9 si el primer rollo seleccionado fue bueno (los tres rollos defectuosos siguen estando en la caja que contienen nueve rollos).

A la fracción 2/9 o bien a la 3/9 se le denomina probabilidad condicional, porqué su valor está condicionado a este caso que el primer rollo que saco de la caja haya sido defectuoso o no.

1.13.Probabilidad Condicional

Es la probabilidad de que ocurra un evento determinado, dado que otro evento ya haya sucedido.

La regla general de la multiplicación se utiliza para determinar la probabilidad conjunta de que ocurra dos eventos, en este caso como seleccionar dos rollos fotográficos defectuosos con una caja de 10 rollos uno después de otro.

En general la regla establece que dados los eventos A y B, la probabilidad conjunta de que ambos ocurran se encuentra multiplicando la probabilidad que suceda el evento A por la probabilidad condicional que ocurra el evento B.

Con lo anotado anteriormente tenemos que

1.14. Regla General de Multiplicacion

P (A y B)=(PA) (P/B)

" / => dado que"

Una vez analizado aplicaremos en nuestro ejemplo anterior de la caja que contiene 10 rollos. Se va a seleccionar 2 rollos, uno después del otro.

¿Cuál es la probabilidad de escoger un rollo con defectos seguido de otro también de defectuosos?

1. El primer rollo seleccionado de la caja, que resulto con defectos es el evento A, de modo que tendríamos.

$$PA = \frac{3}{10} \longrightarrow \boxed{\text{DEFECTUOSOS}}$$

2. El segundo rollo que también era defectuosa, es el evento B, por lo tanto tendremos:

$$P\,{}^{B}/_{A} = \frac{2}{9} \longrightarrow \boxed{\text{DEFECTUOSOS}}$$

Porque después del primer objeto seleccionado fue rollo con defectos, solo quedaron 2 rollos "defectuosos", en la caja que contenía 9. La probabilidad para que cada fórmula de 2 rollo defectuosos sería.

$$PA \text{ y } PB = (PA)\,(P\,{}^{B}/_{A})$$

$$PAyPB = \left(\frac{3}{10}\right)\left(\frac{2}{9}\right)$$

$$PA \text{ Y } PB = 0.07 \qquad 7\%$$

Se supone que este experimento se realiza sin reposición (o reemplazo), es decir el primer rollo defectuoso de la película no se devolvió a la caja antes de seleccionar el siguiente rollo, también se puede observar que la regla general de la multiplicación puede ampliarse a más de 2 eventos.

Para el caso de 3 eventos tendríamos:

$$PA \text{ y } PB \text{ y } PC = (PA)\,(P\,{}^{B}/_{A})\,\left(P\frac{C}{AyB}\right)$$

$$PAyByC = \left(\frac{3}{10}\right)\left(\frac{2}{9}\right)\left(\frac{1}{8}\right)$$

$$PAyByC = \frac{1}{120}$$

$$PAyByC = 0.008$$

Conclusión:

La probabilidad de sacar 3 rollos defectuosos consecutivamente es mínima.

A continuación, se presenta otra aplicación de la regla general de multiplicación.

Una encuesta ejecutivos se enfocó a su lealtad de la empresa, una de las preguntas planteadas fue.

¿Si otra compañía le hiciera la oferta igual o ligeramente mejor a la propuesta, permanecería en la empresa o tomaría otro empleo?

Las respuestas de los 200 ejecutivos se clasificaron en forma cruzada, con su tiempo de servicio en la compañía.

Al tipo de tabla que resulta se le llama Tabla De Contingencia

TIEMPO DE SERVICIO

LEALTAD	MENOS DE 1 AÑO	1 A 5 AÑOS	6 A 10 AÑOS	MAS DE 10 AÑOS	TOTAL
SI PERMANECERIA	10	30	5	75	120
NO PERMANECERIA	25	15	10	30	80

En esta aplicación necesitamos saber:

¿Cuál es la probabilidad de seleccionar al azar un ejecutivo que sea leal a la empresa y que tenga más de 10 años de servicio?

- Solución

1

El evento ocurre si un ejecutivo seleccionado al azar a pesar de que otra compañía le hiciera una oferta igual o ligeramente mejor.

Para encontrar la probabilidad de que suceda el evento A, se observa que hay 120 ejecutivos, de los 200 que participaron en la encuesta, es decir:

$$P_A = \frac{120}{200}$$

$$P_A = \frac{3}{5}$$

$$P_A = 0.6$$

2

El evento B, ocurre si el ejecutivo seleccionado al azar tiene más de 10 años se servicio en la empresa. - De esta forma la probabilidad de B de que suceda P(B/A) es la forma con la probabilidad condicional de que un ejecutivo con más de 10 años de servicio permanezca en la empresa a pesar de que la otra compañía le haga una oferta igual o ligeramente mejor.

Al consultar la tabla de contingencia, de los 120 ejecutivos que se quedarían y que tienen más de 10 años de servicio tendíamos que:

$$P_{\frac{B}{A}} = \frac{75}{120}$$

$$P_{\frac{B}{A}} = 0,63$$

La probabilidad de que un ejecutivo seleccionado al azar sea uno de los que se quedarían en la compañía y en los que tienen más de 10 años de servicio.

Utilizando la regla general de la multiplicación tendríamos:

$$P_{AyB}=P_A P_{\frac{B_4}{A}}$$

$$P_{AyB}=\left(\frac{120}{200}\right)\left(\frac{75}{120}\right)$$

$$P_{AyB}=0{,}38$$

Conclusión:

La probabilidad de que un ejecutivo seleccionado al azar con más de 10 años de servicio no es favorable para que permanezca en la empresa.

Aplicaciones:

29.- Suponga que P(A)=0.40 Y P (B/A) = 0.30 ¿Cuál es la probabilidad conjunta de A y B.

PA y PB = PA* PB/A

= (0.40) (0.30)

= 0.12

Conclusión:

La probabilidad conjunta es mínima

30.- Considere que P(X1)=0.75 y P (Y2/X1)=0.40 ¿Cuál es la probabilidad conjunta de X1yY2

P(X)= 0.75

P (Y)=0.40

PX1y P Y2=PX1 * P Y2/X1

= 0.75*0.40

= 0.30

Conclusión:

La probabilidad conjunta es mínima

31.- Un banco local reporta que 80% de sus clientes tienen una cuenta de cheques, 60% una cuenta de ahorros, y 50% tienen ambas. Si se seleccionan una clienta al azar, ¿Cuál es la probabilidad de que este tenga una cuenta de cheques o una cuenta de ahorros? ¿Cuál es la probabilidad de que el cliente no tenga ninguna de las dos?

PA = 80% tienen cta. de cheques

PB= 60% tienen cta. Ahorros

PC= 50% tienen ambas

P(A o B)= (PA)+ (PB) ~ (PC)

= (0.80)+ (0.60) ~ (0.50)

= 0.90

Conclusión:

La probabilidad de que el cliente tenga una de las dos cuentas es de 0.90 es decir tenemos una probabilidad mínima.

32.- La empresa AllSeasonsPlumbing cuenta con dos camiones de servicios que se descomponen frecuentemente. Si la probabilidad de que el primer camión está disponible es 0.75 la de que el segundo camión también lo está es 0.50, y la probabilidad de que ambos camiones estén disponibles es 0.30 ¿Cuál es la probabilidad de que ningún vehículo esté disponible?

PA =0.75

PB=0.50

PC= 0.30

P(A o B) = (PA)+ (PB) ‿ (PC)

= (0.75) (0.50) ‿ (0.30)

= 0.95

Conclusión:

La probabilidad es mínima.

33.- considere la siguiente tabla

		PRIMER EVENTO			
Segundo evento		A1	A2	A3	Total
B1		2	1	3	6
B2		1	2	1	4
Total		3	3	4	10

a) Determinar P(A1)

a) Determinar P(B1/A2)

b) Determinar P(B2 y A3)

a) P (A) = 3/10

= 0.30

b) P (B1/A2)

PB1 = (6/10) (3/10)

= 18/100

0.18

c) P (B2/A3) = (4/10) (4/10)

=16/100

= 0.16

Conclusión:

La probabilidad es mínima.

34.- CleanbrushProduct envió por accidente a una farmacia tres cepillos dentales eléctricos, que estaban defectuosos, junto con 17 en buen estado.

a) ¿Cuál es la probabilidad de que los primeros dos cepillos vendidos se devuelvan a la farmacia por tener defectos?

b) ¿Cuál es la probabilidad de que los primeros dos cepillos vendidos no tengan defectos?

a) PA = (P B/A)

PA= 2/19

= 0.10

b) PB= 1– PB/A

PB=1 - 0.010

= 0.89

35.- Cada vendedor en la negociación Stiles- Compton se califica como "abajo del promedio" , " promedio" o " arriba del promedio" , con respecto a su actitud para las ventas Además , cada uno se clasifican respecto de su probabilidad de promoción en : regular , bien , o excelente . En la tabla que sigue se presenta la clasificación cruzada respecto a estos conceptos de los 500 vendedores.

	Posibilidades de promoción		
Aptitud en ventas	Regular	Buena	Excelente
Abajo del promedio	16	12	22
promedio	45	60	45
Arriba del promedio	93	72	135

a) ¿Cómo se denomina la tabla?

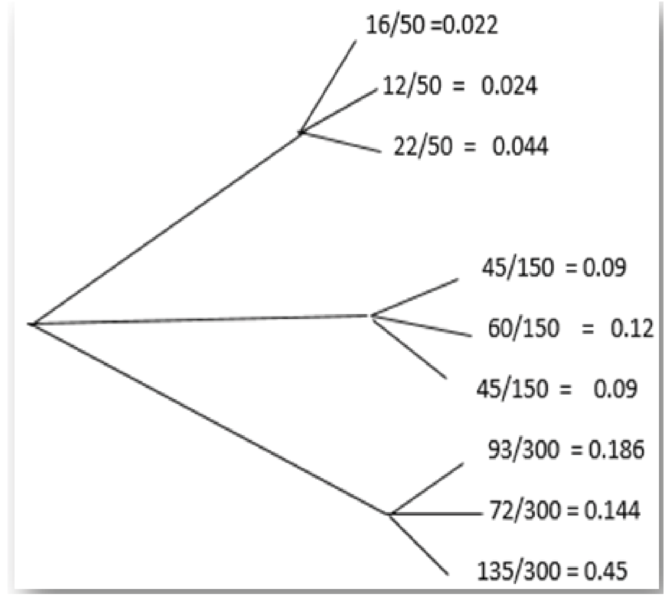
a) ¿Cuál es la probabilidad de que un vendedor seleccionado al azar tenga aptitud para las ventas por encima del promedio y excelente posibilidad de promoción?

a) Trace un diagrama de árbol que muestra las posibilidades normales, condicionales y conjuntas.

a) ¿Cómo se denomina la tabla?

Tabla de contingencia

a) ¿Cuál es la probabilidad de que un vendedor seleccionado al azar tenga aptitud para las ventas por encima del promedio y excelente posibilidad de promoción?



B) P (B)= 135/300

P (B)= 0.45

c) PA y B = PA* P B/A

PA y B0 300/500 * 135/300

PA y B = 0.27

Conclusión:

La probabilidad es de 0.27

36.- Un inversionista posee tres acciones comunes. Cada acción, independiente de las otras, tiene las mismas posibilidades de que (1) aumente su valor, (2) disminuya su valor, o (3) permanezca sin cambio. Mencione todos los posibles resultados de este experimento. Calcule la probabilidad de que al menos dos de las acciones aumenten de valor.

1.-Probabilidad

CLÁSICA $P(A) = \frac{1}{3}$ $P(B) = \frac{2}{3}$ $P(C) = \frac{3}{3}$

$P(A) = 0{,}333333$ $P(B) = 0{,}666667$ $P(C) = 1$

Probabilidad Clásica $P(A) = \frac{2}{3}$

$P(A) = 0{,}666667$

$P(A) = 67\%$

Conclusión:

Existe un 67% de que las acciones aumenten su precio.

37.- el comité directivo de una empresa pequeña está integrado por cinco personas. Tres son "lideres fuertes". Si aceptan un proyecto, lo aprobaran todos los demás miembros del comité. Los otros integrantes, "lideres débiles", no tiene influencia alguna. Se programa que tres vendedores harán presentaciones de ventas, uno después de otro ante un miembro del comité, elegido

por el vendedor. Los representantes de ventas son convincentes, pero no saben quiénes son los "líderes fuertes". Sin embargo, sabrán a quieres se dirigió el represéntate de ventas anterior. ¿El primer vendedor tiene la misma probabilidad de obtenerla? De lo contrario, evalué sus probabilidades respectivas de ganar dicha cuenta.

1.- Probabilidad

CLÁSICA $P(A) = \frac{3}{5}$

$P(A) = 0{,}60$

$P(A) = 60\%$

1.- Probabilidad

CLÁSICA $P(A) = \frac{1}{5}$ $P(B) = \frac{1}{4}$ $P(C) = \frac{1}{3}$

$P(A) = 0{,}2$ $P(B) = 0{,}25$ $P(C) = 0{,}33$

38.- Si en la universidad usted pregunta a tres personas desconocidas, ¿Cuál es la probabilidad de que: (a) todos hayan nacido en un día miércoles? (b) todos hayan nacido en días de la semana diferentes? (c) ninguna haya sido en sábado?

a.-Probabilidad

Clásica $P(A) = \frac{1}{8}$

$P(A) = 0{,}13$

$P(A) = 13\%$

b.-Probabilidad

Clásica $P(A) = \frac{1}{8}$

$P(A) = 0{,}38$

$P(A) = 38\%$

ARBORIGRAMA

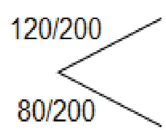
DIAGRAMA DEL ARBOL

Es una representación gráfica para organizar cálculos que abarcan varias etapas cada segmento en el árbol es una etapa del problema.

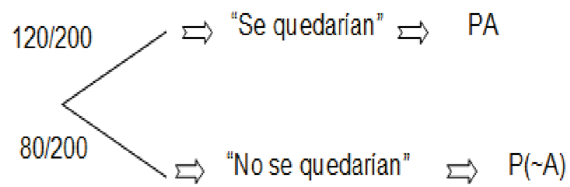
La probabilidad escrita cerca de las ramas son probabilidades condicionales del experimento. Para demostrar la elaboración de un diagrama de árbol utilizaremos los datos de la tabla anterior.

PASOS PARA LA ELABORACIÓN

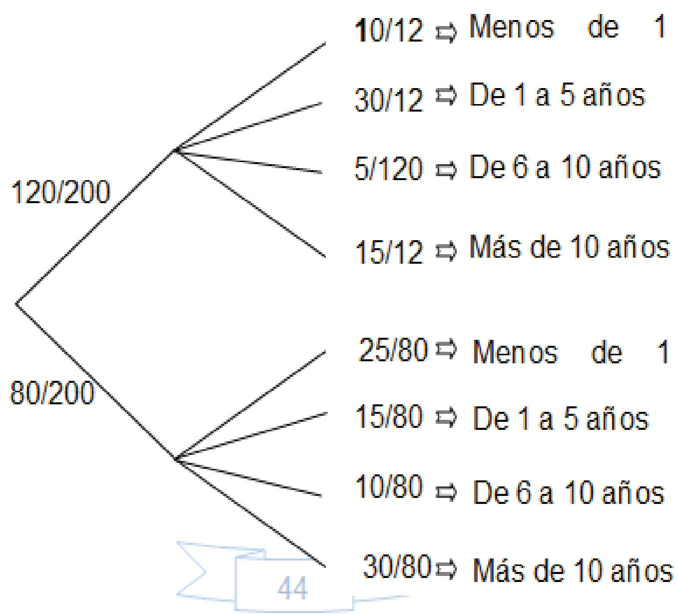
1. La elaboración de un diagrama del árbol se empieza trazando un pequeño punto a la izquierda, que representa el punto central de un tronco de árbol.



2. Para este problema salen dos ramas principales del tronco; la superior representa "se quedarían" y la inferior "no se quedarían". Sus probabilidades se indican en las ramas, específicamente 120/200 y 80/200. Se simboliza por PA y P(~A).



3.Cuatro ramas secundarias se desprenden de cada rama principal y corresponden a los tiempos de servicio: menos de un año, de 1 a 5 años, de 6 a 10 años y de más de 10 años. Las probabilidades que tendríamos son: …



Recordando que PB4/A sería:

PB1/A,PB2/A,PB3/A,PB4/A

De donde:
.B1 Se refiere a las personas que trabajan menos se un año.
.B2 A las personas que trabajan de 1 a 5 años.
.B3A las personas que trabajan de 6 a 10 años.
.B5 A las personas que trabajas más de diez años.

Se realiza el mismo proceso para los eventos de la parte interior, los empleados que no se quedarían.

4. Por último, las probabilidades conjuntas de que A y B ocurran al mismo tiempo se muestran al lado derecho. Por ejemplo, la probabilidad conjunta de seleccionar al azar un ejecutivo que permanecería en la empresa y que tiene menos de un año de servicio. Utilizando la formula sería

$$P(A \text{ y } B) = P(A)\,P(B1/A)$$

$$P(A \text{ y } B) = (120/200)\,(10/120)$$

$$P(A \text{ y } B) = 0.05$$

Conclusión:

La probabilidad de que un empleado que tenga menos de un año en la empresa se quede en el trabajo es casi nula.

El diagrama completo de árbol, en este caso nos quedaría así:

Si se suma las probabilidades de que se queden nos da un 0.60 y de que no se queden un 0.40.

Conclusión:

Existe la probabilidad de que la mayoría de los empleados de la empresa se queden en la empresa, con una probabilidad favorable del 60%.

1.15. Teorme de Bayes

La fórmula de Bayes permite calcular las probabilidades condicionales de los eventos a priori dado que el evento a posteriori ha ocurrido; sean A y C dos eventos tales que A es evento a priori y C es evento a posteriori, tal como indica en el árbol de probabilidad anterior; entonces, la probabilidad de que el evento A (a priori) ocurra, dado que el evento C (a posteriori) ha ocurrido se expresa de la siguiente manera:

$$P(A|C) = \frac{P(A \text{ y } C)}{P(C)} = \frac{P(A) \times P(C|A)}{(A) \times P(C|A) + (B) \times P(C|B)}$$

Asociando esta expresión con la definición clásica de probabilidad se puede indicar:

- Casos favorables: la probabilidad de que ocurra A y ocurra C dado que el evento A ha ocurrido.

- Total de casos: La probabilidad de que ocurra A o que ocurra B, dado que el evento C ha ocurrido.

1.16. Recomendaciones Metodológicas

- Identifique y describa los eventos que intervienen en el problema, señale claramente los eventos a priori y los eventos a posteriori; determine la probabilidad de ocurrencia de cada uno de ellos.

- Trace el árbol de probabilidad señalando los eventos a priori y los eventos a posteriori; en cada una de las ramas del árbol coloque la probabilidad de ocurrencia del evento; tome en cuenta que los eventos a posteriori están supeditados a la ocurrencia de los eventos a priori.

- Verifique que los eventos a priori y los eventos a posteriori sean colectivamente exhaustivos, es decir que la suma de sus probabilidades de ocurrencia sea igual a 1

- Para encontrar la probabilidad condicional de un evento a priori, dado que el evento a

a posteriori ha ocurrido debe aplicar la fórmula de Bayes.

1.17. Problemas Resueltos

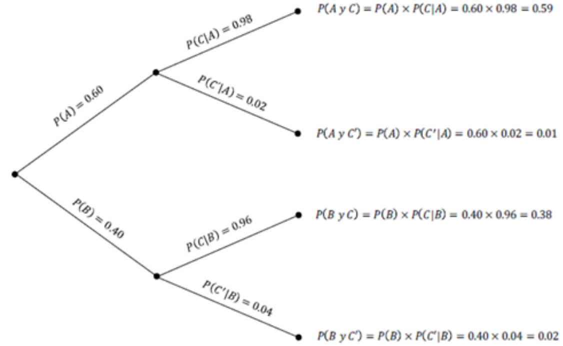
1.Los repuestos de computador se fabrican en dos máquinas, la máquina A fabrica el 60% de la producción total y la máquina

B fabrica el 40% restante de la demanda; existe un 98% de probabilidad de que los repuestos fabricados por la máquina A sean óptimos; mientras que existe un 96% de probabilidad que los repuestos fabricados con la máquina B sean óptimos; se toma un repuesto al azar, con esta información:

Eventos a priori		
Evento	Descripción	Probabilidad
A	Pieza fabricada en la máquina A	0.60
B	Pieza fabricada en la máquina B	0.40

Eventos a posteriori			
Evento	Descripción	Probabilidad ocurrencia	Probabilidad no ocurrencia
C|A	Pieza óptima máquina A	0.98	0.02
C|B	Pieza óptima máquina B	0.96	0.04

a. Construya el árbol de probabilidades.



b. Probabilidad de que se obtenga un repuesto óptimo que se fabricó en la máquina A

$P(C\ y\ A) = P(A) \times P(C|A) = 0.60 \times 0.98 = 0.59$

c. Probabilidad de que se obtenga un repuesto óptimo que se fabricó en la máquina B.

$P(C\ y\ B) = P(B) \times P(C|B) = 0.40 \times 0.96 = 0.38$

d. Probabilidad de que se obtenga un repuesto defectuoso que se fabricó en la máquina A.

$P(C'\ y\ A) = P(A) \times P(C'|A) = 0.60 \times 0.02 = 0.01$

e. Probabilidad de que se obtenga un repuesto defectuoso que se fabricó en la máquina B

$$P(C' \text{ y } B) = P(B) \times P(C'|B) = 0.40 \times 0.04 = 0.02$$

f. Probabilidad de que el repuesto obtenido sea óptimo.

$$P(C) = [P(C \text{ y } A) \text{ o } P(C \text{ y } B)]$$
$$P(C) = P(C \text{ y } A) + P(C \text{ y } B)$$
$$P(C) = P(A) \times P(C|A) + P(B) \times P(C|B)$$
$$P(C) = 0.59 + 0.38 = 0.97$$

g. Probabilidad de que el repuesto obtenido sea defectuoso.

$$P(C') = [P(C' \text{ y } A) \text{ o } P(C' \text{ y } B)]$$
$$P(C') = P(C' \text{ y } A) + P(C' \text{ y } B)$$
$$P(C') = P(A) \times P(C'|A) + P(B) \times P(C'|B)$$
$$P(C') = 0.01 + 0.02 = 0.03$$

h. Probabilidad de que el repuesto seleccionado sea de la máquina A dado que es óptimo.

$$P(A|C) = \frac{P(A \text{ y } C)}{P(C)} = \frac{P(A) \times P(C|A)}{P(A) \times P(C|A) + P(B) \times P(C|B)} = \frac{0.59}{0.59 + 0.38} = 0.61$$

i. Probabilidad de que el repuesto seleccionado sea de la máquina B dado que es óptimo.

$$P(B|C) = \frac{P(B \text{ y } C)}{P(C)} = \frac{P(B) \times P(C|B)}{P(A) \times P(C|A) + P(B) \times P(C|B)} = \frac{0.38}{0.59 + 0.38} = 0.39$$

j. Probabilidad de que el repuesto seleccionado sea de la máquina A, dado que es defectuoso.

$$P(A|C') = \frac{P(A \text{ y } C')}{P(C')} = \frac{P(A) \times P(C'|A)}{P(A) \times P(C'|A) + P(B) \times P(C'|B)} = \frac{0.01}{0.01 + 0.02} = 0.33$$

k. Probabilidad de que el repuesto seleccionado sea de la máquina B dado que es defectuoso.

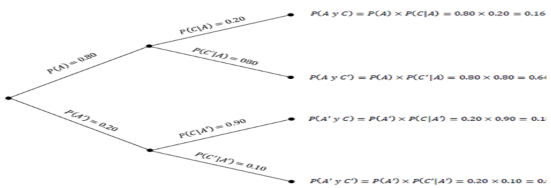
$$P(B|C') = \frac{P(B \text{ y } C')}{P(C')} = \frac{P(B) \times P(C'|B)}{P(A) \times P(C'|A) + P(B) \times P(C'|B)} = \frac{0.02}{0.01 + 0.02} = 0.67$$

5.2. El despertador de Javier no funciona muy bien y el 20% de las veces no suena. Cuando s
Javier llega tarde a clase con probabilidad 0.20; pero si no suena, la probabilidad de que
tarde a clase es 0.90.

a. Determine la probabilidad de que llegue tarde a clase y haya sonado el despertador. m
b. Halle la probabilidad de que llegue temprano.
c. Javier ha llegado tarde a clase, ¿cuál es la probabilidad de que haya sonado el desperta

Eventos a priori			
Evento	Descripción	Probabilidad ocurrencia	Probabilidad no ocurrencia
A	Despertador suena	0.80	0.20

Eventos a posteriori			
Evento	Descripción	Probabilidad ocurrencia	Probabilidad ocurrencia
C|A	Llega tarde dado que despertador suena	0.20	0.80
C|A'	Llega tarde dado que despertador no suena	0.90	0.10



a. La probabilidad de que llegue tarde a clase y haya sonado el despertador está dado por:

$$P(C \text{ y } A) = P(A \text{ y } C) = P(A) \times P(C|A) = 0.80 \times 0.20 = 0.16$$

b. La probabilidad de que llegue temprano, es decir que no llegue tarde está dado por:

$$P(C') = P[(C' \text{ y } A) \text{ o } (C' \text{ y } A')]$$
$$P(C') = P(A \text{ y } C') + P(A' \text{ y } C')$$
$$P(C') = P(A) \times P(C'|A) + P(A') \times P(C'|A')$$
$$P(C') = 0.64 + 0.02 = 0.66$$

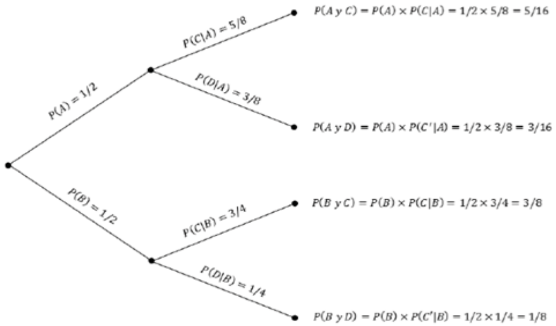
c. la probabilidad de que haya sonado el despertador, ya que Javier llegó tarde está dado por:

$$P(A|C) = \frac{P(A \text{ y } C)}{P(C)} = \frac{P(A) \times P(C|A)}{P(A) \times P(C|A) + P(A') \times P(C|A')} = \frac{0.16}{0.16 + 0.18} = 0.47$$

5.3. Se recibieron de la fábrica dos cajas de camisas para caballero marca Old Navy, la caja 1 contenía 25 camisas polo y 15 camisas Súper-T, la caja 2 contenía 30 camisas Polo y 10 camisas Súper-T. Una de las cajas se seleccionó al azar y se eligió una camisa de dicha caja, también en forma aleatoria, para revisarla. La camisa era Polo. Dada esta información, ¿cuál es la probabilidad de que la camisa polo provenga de la caja 1?

Eventos a priori		
Evento	Descripción	Probabilidad
A	Caja 1 ha sido seleccionada	0.50
B	Caja 2 ha sido seleccionada	0.50

Eventos a posteriori		
Evento	Descripción	Probabilidad ocurrencia
C|A	Camisa polo dada que fue de la caja 1	25/40 = 5/8
D|A	Camisa súper T dada que fue de la caja 1	15/40 = 3/8
C|B	Camisa polo dada que fue de la caja 2	30/40 = 3/4
D|B	Camisa súper T dada que fue de la caja 2	10/40 = 1/4



$$P(A|C)=\frac{P(A\,y\,C)}{P(C)}=\frac{P(A)\times P(C|A)}{P(A)\times P(C|A)+P(B)\times P(C|B)}=\frac{\frac{5}{16}}{\frac{5}{16}+\frac{3}{8}}=\frac{\frac{5}{16}}{\frac{11}{16}}=\frac{5}{11}$$

1.18. Permutacion

Se utiliza para determinar el número posibles de arreglos, cuando solo hay un grupo de objetos.

Ejemplo

1.
Se van a ensamblar tres partes electrónicos en una unidad modular para un receptor de televisión. Las partes se pueden ensamblar en cualquier lote. La pregunta relacionada con el conteo es: De cuantos modos diferentes se pueden ensamblarse.

2.Un operario debe realizar cuatro verificaciones de seguridad antes de actuar su máquina, no importa en qué orden lo realice.

¿De cuantas formas distintas pueden realizar las verificaciones?

A estos arreglos se les denomina permutaciones.

PERMUTACIÓN Un arreglo dispersión de "r" objetos seleccionada de un grupo

de "n" objetos posibles.

EN DONDE:

R: N° de objetos seleccionados

n: N° de objetos posibles.

n!: "factorial"

Obtendremos la siguiente fórmula: SI IMPORTA EL ORDEN

$$nPr=\frac{n!}{(n-r)!}$$

Aplicaciones:

1.- La empresa Cepeda S.A. tiene 8 tornos, pero solo hay disponible 3 espacios en la zona de producción. ¿En cuántas formas diferentes se puede colocar los 8 tornos en los tres espacios disponibles? Datos n=8; r=3 si es importante el orden.

$$nPr = \frac{n!}{(n-r)!}$$

$$8P3 = \frac{8!}{(8-3)!}$$

$$8P3 = \frac{8!}{(5)!}$$

$$8P3 = \frac{8*7*6*5*4*3*2*1}{5*4*3*2*1}$$

$$8P3 = 8*7*6$$

$$8P3 = 336$$

Conclusión:

Los tornos se pueden ubicar en 336 formas diferentes.

2.- Si tuviéramos una permutación de 6

DATOS

P=6

r= A*B*C

Solución:

ABC; BAC; CAB; ACB; BCA; CBA

Si el orden de los objetos seleccionados no es importante, a cualquier selección se lo llama combinación y la fórmula para contar el número de combinaciones de r objetos con un conjunto de n objeto es:

$$nPr = \frac{n!}{r!\,(n-r)!}$$

Ejercicios:

1. A un departamento de mercadotecnia se le ha solicitado que enseñe códigos de colores para 42 líneas de discos compactos “CD” que comercializa la empresa GOODSTUDENT. Se van a utilizar 3 colores en cada línea de “CD”, pero una combinación de 3 colores que se utiliza en una línea no puede reordenarse y utilizarse para identificar a otra línea diferente. Esto significa que si se usara los colores verde, amarillo y violeta para señalar una línea, entonces amarillo, verde y violeta (o cualquier otra combinación de estos 3 colores) no se podría emplear para identificar otra línea.

¿Serán adecuados 7 colores tomados 3 la vez para codificar adecuadamente las 42 líneas?

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

7C3= 7!/3! (7-3)!

7C3= 7.6.5.4.3.2.1/3.2.1 (4)!

7C3= 7.6.5.4.3.2.1/3.2.1(4.3.2.1)

7C3= 7.6.5/3.2.1

7C3= 35

Conclusión:

Utilizando 7 colores con 3 a la vez no satisfacen la necesidad de codificar.

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

8C3= 8!/3!(8-3)!

8C3= 8.7.6.5.4.3.2.1/3.2.1!(5)!

8C3= 8.7.6.5.4.3.2.1/3.2.1(5.4.3.2.1)

8C3= 8.7.6/3.2.1

8C3=56

Conclusión:

Utilizando 8 colores tomados de 3 en 3 darían 56 combinaciones distintas. Esto sería más que suficiente para codificar cromáticamente las 42 líneas.

2.Se va a utilizar los 10 primeros números del 0 al 9, para crear códigos de cuarto dígitos e identificar un artículo de ropa. El 1.083 podría identificar una blusa azul, talla mediana, 2.031 para unos pantalones talla 18 y así sucesivamente no se permiten repeticiones de los números. Es decir, el mismo número no puede ser utilizado dos veces (o más) en una secuencia.

¿Cuántos códigos diferentes se pueden establecer?

$$n\,C\,r = \frac{n!}{r!(n-r)!}$$

$$10\,C\,4 = \frac{10!}{4!\,(10-4)!}$$

$$10\,C\,4 = \frac{10!}{4(6)!}$$

$$10\,C\,4 = \frac{10.9.8.7.6.5.4.3.2.1}{4.3.2.1(6.5.4.3.2.1)}$$

$$10\,C\,4 = \frac{5040}{24}$$

$$10\,C\,4 = 210$$

Conclusión:

Se puede establecer 210 combinaciones

3. Un músico desea escribir una partitura basados solamente en 5 notas si bemol, do, re, mí, sol. Sin embargo, solo tres de los cinco se utilizan en la sucesión, como do, si bemol, mí. No se permitirán repeticiones, si bemol y mí.

a) ¿Cuántas permutaciones de 5 notas, tomados tres cada vez, son posibles?

b) ¿Utilizando la fórmula de permutaciones, cuantas permutaciones son posibles ahora?

A)

$$n\,P\,r = \frac{n!}{(n-r)!}$$

$$5\,P\,3 = \frac{5!}{(5-3)!}$$

$$5\,P\,3 = \frac{5!}{(2)!}$$

$$5\,C\,3 = \frac{5.4.3.2.1}{(2.1)}$$

$$10\,C\,4 = 60$$

B)

$$n\,P\,r = \frac{n!}{(n-r}$$

$$3\,P\,3 = \frac{3!}{(3-3}$$

$$3\,P\,3 = \frac{3!}{(0)!}$$

$$3\,C\,3 = 6$$

Conclusión :

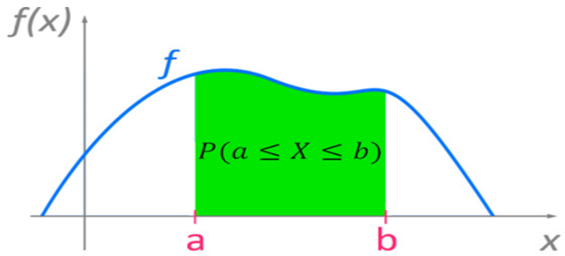
Se puede obtener 6 permutaciones posibles.

1.19.Función de densidad de probabilidad

Sea X una variable aleatoria continua. Entonces, una función de densidad de probabilidad de X es una función f(x) tal que para dos números cualesquiera a y b con $a \leq b$,

$$P(a \leq X \leq b) = \int_{a}^{b} f(x)dx$$

La probabilidad de que X asuma un valor en el intervalo [a, b] es el área sobre este intervalo y bajo la gráfica de la función de densidad. Se puede apreciar mejor en la siguiente gráfica:



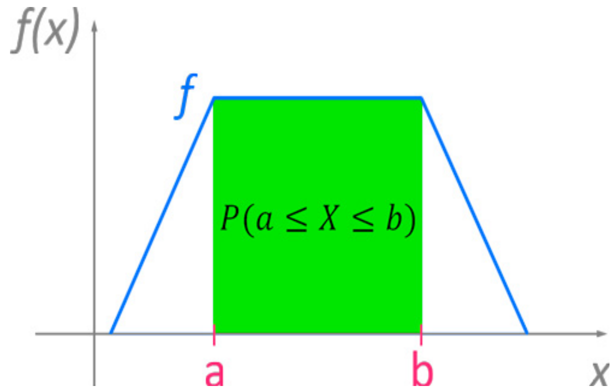
$P(a \leq X \leq b) = \int_a^b f(x)dx \ = \text{ área bajo la curva de densidad entre a y b}$

La gráfica de f(x) se suele llamar curva de densidad.

La función de probabilidad de una variable aleatoria continua siempre cumplirá con estas condiciones:

$$\begin{cases} 1. \ f(x) \geq 0 \quad \text{con todas las } x. \\ 2. \ \displaystyle\int_{-\infty}^{\infty} f(x)dx = 1 = \text{área bajo toda la gráfica de } f(x) \end{cases}$$

No te asustes con las integrales, dado que estamos en el curso de estadística, las integrales no serán difíciles y en muchos de los casos podemos usar la calculadora para encontrar el valor de estas integrales definidas. En otros casos, el área bajo la curva tendrá forma de triángulo o rectángulo y será muy sencilla de calcular.



Ejemplo 1

A partir de la función de densidad de probabilidad f(x), calcular $P(1 \leq X \leq 3)$

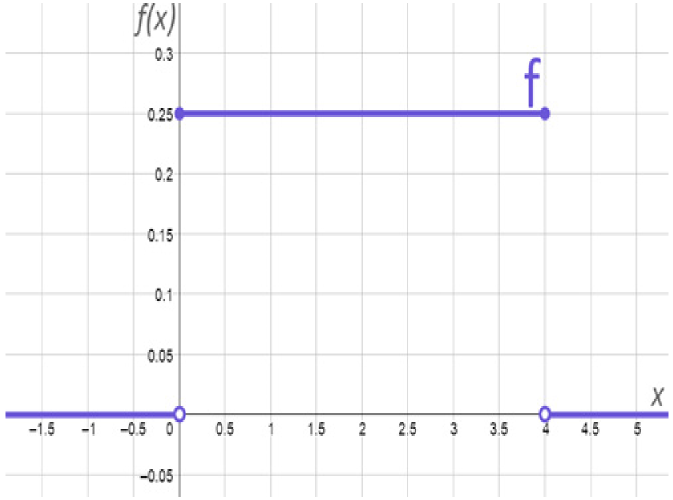
$$f(x) = \begin{cases} 0{,}25 \ ; & si \ \ 0 \leq x \leq 4 \\ 0 & de\ lo\ contrario \end{cases}$$

Solución:

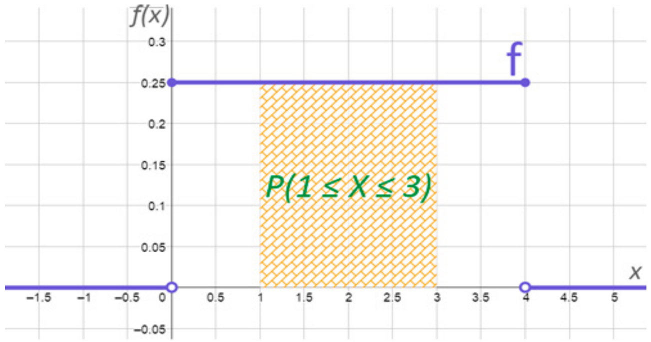
En este problema, nos piden calcular $P(1 \leq X \leq 3)$. Y con la función de probabilidad de una variable aleatoria continua, las probabilidades se calculan mediante el área bajo la curva, por ello:

$$P(1 \leq X \leq 3) = \text{área bajo la curva de densidad entre 1 y 3}$$

Graficamos la función f(x) para que se vea mucho mejor:

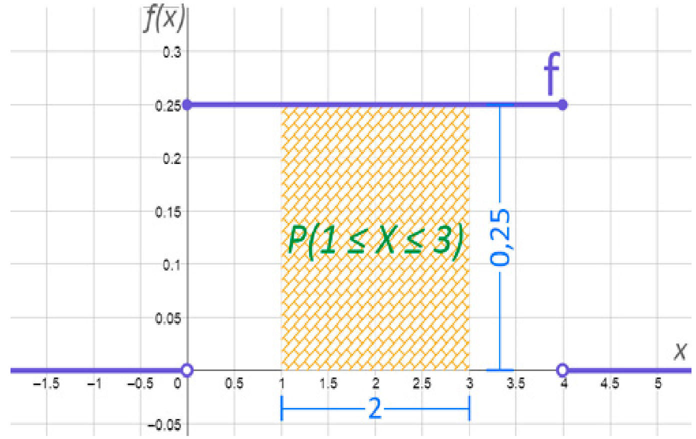


Como queremos calcular la probabilidad de que nuestra variable aleatoria continua X tome un valor entre 1 y 3, entonces sombreamos el área bajo la función f(x) en ese intervalo:



Solo nos queda calcular el valor del área sombreada y en este caso se puede realizar de 2 formas diferentes: mediante la fórmula del rectángulo y mediante la integral definida de f(x) desde x igual a 1 hasta 3.

son áreas:



$$P(1 \leq X \leq 3) = A_{\blacksquare} = b \cdot h = 2 \cdot 0{,}25 = 0{,}5$$

Con integrales:

$$P(1 \leq X \leq 3) = \int_1^3 f(x)dx$$

$$P(1 \leq X \leq 3) = \int_1^3 0{,}25\, dx = 0{,}25x \Big|_1^3$$

$$P(1 \leq X \leq 3) = 0{,}25(3) - 0{,}25(1) = 0{,}75 - 0{,}25$$

$$P(1 \leq X \leq 3) = 0{,}5$$

Como verás, obtuvimos el mismo resultado con ambos métodos, una probabilidad de 0,5 o 50%.

Propiedad importante:

En la función de probabilidad de una variable aleatoria continua sucede algo bien interesante si queremos calcular la probabilidad de que la variable aleatoria sea igual a un valor puntual c.

$$P(X = c)$$

La probabilidad se calcula mediante el área bajo la curva pero el área bajo una curva de densidad situada sobre cualquier valor único es 0:

$$P(X = c) = \int_{c}^{c} f(x)dx = 0$$

El hecho de que P(X = c) = 0 cuando X es continua nos permite afirmar que la probabilidad de que X quede en algún intervalo entre a y b no depende de si límite inferior a o el límite superior b está incluido en el cálculo de la probabilidad:

$$\boxed{P(a \le X \le b) = P(a < X < b) = P(a < X \le b) = P(a \le x < b)}$$

1.20.Ejercicios de Aplicación

1) Calcular k para que la función dada sea una función de densidad.

$$f(x) = \begin{cases} k, & x \in [1, 4] \\ 0, & x \notin [1, 4] \end{cases}$$

Hallar las probabilidades:

a) $P(2 \le x < 3)$

b) $P(2 \le x < 4)$

c) $P(2 \le x < 9)$

El área total bajo la curva ha de ser igual a 1.

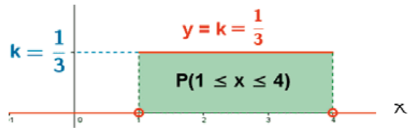
$$P(-\infty < x < +\infty) = P(1 < x < 4) = 3 \cdot k = 1$$

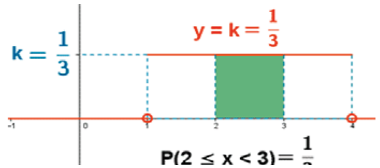
$$k = \frac{1}{3}$$

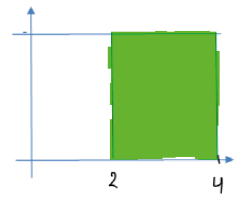
a) $P(2 \le X < 3) = (3-2) \cdot \frac{1}{3} = \frac{1}{3}$

b) $P(2 \le X < 4) = (4-2) \cdot \frac{1}{3} = \frac{2}{3}$

c) $P(2 \le X < 9) =$ Pues en el tramo $[4, 9]$ la probabilidad es cero







2) Calcular m para que la función dada sea una función de densidad.

$$f(x) = \begin{cases} mk, & x \in [1, 4] \\ 0, & x \notin [1, 4] \end{cases}$$

Hallar P(2<x<3)

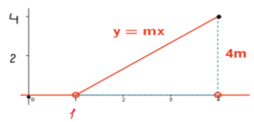
SOLUCIÓN:

El área total bajo la curva ha de ser igual a 1.

Área del triángulo : $\frac{4 \cdot 3m}{2} = 6m$ luego:

$$P(-\infty < x < +\infty) = P(1 < x < 4) = 6 \cdot m = 1$$

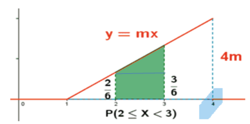
$$m = \frac{1}{6}$$



La función de densidad es: $f(x) = \begin{cases} \frac{x}{8} & , x \in [1, 4] \\ 0 & , x \notin [1, 4] \end{cases}$

$P(2 \le X < 3)$ Es el área del trapecio verde $= \frac{B+b}{2} \cdot h$

$$P(2 \le X < 3) = \frac{\left(\frac{2}{6} + \frac{3}{6}\right) \cdot 1}{2} = \frac{5}{12}$$



3) Hallar la función de distribución de la variable aleatoria cuya función de densidad es:

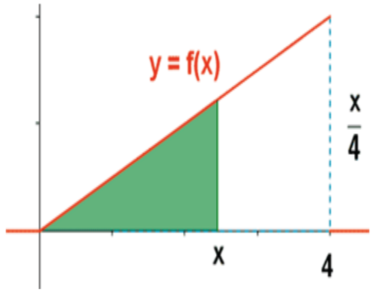
$$f(x) = \begin{cases} \dfrac{x}{4}, & x \in [0, 4] \\ 0, & x \notin [0, 4] \end{cases}$$

La función F(x) describe los valores que toma la probabilidad acumulada hasta la abcisa x : $F(x) = P(t \le x)$

$$P(t \le x) = \text{área del triángulo verde} = \frac{1}{2}x \cdot \frac{1}{4}x = \frac{x^2}{8}$$

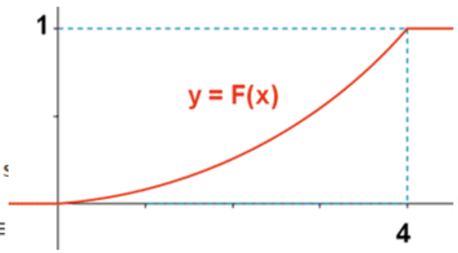
Luego:

$$F(x) = \begin{cases} 0 & \text{si } x \le 0 \\ \dfrac{x^2}{8} & \text{si } x \in [0, 4] \\ 1 & \text{si } x \ge 4 \end{cases}$$



Por tanto se trata de una función continua y creciente, que verifica que:

$F(x) = 0$ para $x \le 0$
$F(x) = 1$ para $x \ge 4$



4) Considera la s

$$f(x) = \begin{cases} \dfrac{1}{4} & \text{si } x \in \\ 0 & \text{si } x \notin [2, 6] \end{cases}$$

a) Comprueba que es una función de densidad.

b) Halla la función de distribución F de la variable aleatoria X cuya función de densidad es f y represéntala gráficamente.

c)Calcula $P(3{,}6 \le X \le 5{,}2)$

a)

Para comprobar que f es una función de densidad, debemos comprobar si cumple las condiciones de la definición :

– $f(x) \ge 0 \ \forall x \Rightarrow$ Está claro que se cumple, ya que f sólo toma los valores 0 y $\frac{1}{4}$.

– El área bajo la función f tiene que ser 1 :

$$\int_{-\infty}^{\infty} f(x)\,dx = 1 \Rightarrow \int_{-\infty}^{2} 0\,dx + \int_{2}^{6} \frac{1}{4}\,dx + \int_{6}^{\infty} 0\,dx = 0 + \left[\frac{x}{4}\right]_2^6 + 0 = \left(\frac{6}{4}\right) - \left(\frac{2}{4}\right) = 1$$

b)

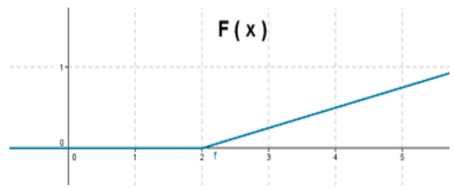
$$\text{Si } x < 2 \Rightarrow P(X \le x) = \int_{-\infty}^{\infty} 0\,dx = 0$$

$$\text{Si } 2 \le x \le 6 \Rightarrow P(X \le x) = \int_{2}^{x} \frac{1}{4}\,dx = \left[\frac{x}{4}\right]_2^x = \frac{x-2}{4}$$

$$\text{Si } x > 6 \Rightarrow P(X \le x) = \int_{-\infty}^{2} 0\,dx + \int_{2}^{6} \frac{1}{4}\,dx + \int_{6}^{x} 0\,dx = 0 + \left[\frac{x}{4}\right]_2^6 + 0 = 1$$

Por lo tanto, la función de distribución sería la siguiente :

$$F(x) = \begin{cases} 0 & \text{si } x < 2 \\ \dfrac{x-2}{4} & \text{si } 2 \le x \le 6 \\ 1 & \text{si } x > 4 \end{cases}$$



Taller Individual

La variable aleatoria continua X tiene la siguiente

función de distribución acumulativa:

$F(x) = \{0;\ \text{si } x < -2$

$0{,}25x + 0{,}5;\ \text{si} - 2 \leq x < 2$

$1;\ \text{si } 2 \leq x$

Encontrar la función de densidad de probabilidad.

1.21. Distribución de Porobabilidad Discreta

Introducción

Una distribución de probabilidad indica toda la gama de valores que puede presentarse como resultado de un experimento.

Ejemplo:

Un fabricante de medicamentos afirma que un tratamiento causará la pérdida de peso del 80% de la población.

Una agencia de protección al consumidor puede probar este medicamento en 6 personas si la declaración del fabricante es verdadera es casi imposible tener un resultado en el que ninguna persona de la muestra pierda peso y es muy probable que 5 de las 6 pierda peso.

1.22.Distribución Probabilidad

“Indica todos los resultados posibles de un experimento junto con la probabilidad correspondiente a cada uno de ellos”.

Función de probabilidad de una variable aleatoria discreta

Veamos la función de probabilidad de una variable aleatoria discreta con ejemplos y ejercicios.

Una función de probabilidad de una variable aleatoria discreta es una función que asigna probabilidades a los valores de la variable aleatoria.

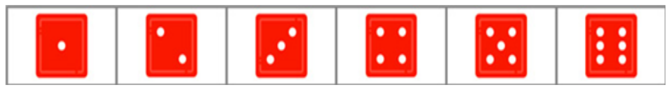
La fórmula que usaremos para la función de probabilidad de una variable aleatoria discreta es la siguiente:

$$f(x) = P(X = x)$$

Como suena complicado, mejor lo aprendemos con algunos ejemplos.

Ejemplo 1.-

Vamos a realizar un experimento muy sencillo que consiste en lanzar un dado. Los resultados que podemos obtener los colocamos de forma gráfica:



Definimos la variable aleatoria discreta X:

X = resultado de lanzar un dado.

Ahora colocamos en una tablita todos los valores posibles del rango de esta variable aleatoria discreta X:

x	1	2	3	4	5	6

Vamos a asociar cada valor con una probabilidad. En un dado, las probabilidades son bastante sencillas de calcular usando la fórmula de probabilidad:

$$\text{Probabilidad de que ocurra } A = \frac{\text{número de casos favorables de } A}{\text{número total de casos posibles}}$$

$$\text{Probabilidad de obtener } 1 = \frac{1}{6}$$

$$\text{Probabilidad de obtener } 2 = \frac{1}{6}$$

$$\vdots$$

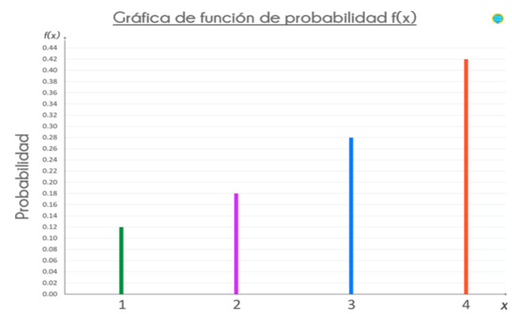
$$\text{Probabilidad de obtener } 6 = \frac{1}{6}$$

Ahora colocamos las probabilidades en la tablita:

x	1	2	3	4	5	6
$f(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Y listo, ya tenemos la función de probabilidad de la variable aleatoria X representada en forma de tabla.

Además, una función de probabilidad también se puede representar de manera gráfica mediante un diagrama de barras. En el eje horizontal se colocan los valores de la variable aleatoria y en el eje vertical las probabilidades. En el caso de nuestra función de probabilidad, la gráfica de barras sería la siguiente:



Una función de probabilidad de una variable aleatoria discreta se puede expresar de manera fácil como una tabla o lista que empareja los valores de la variable con sus probabilidades o también mediante una gráfica de barras. Sin embargo, una función de probabilidad de expresa con mayor frecuencia mediante una fórmula.

Por ejemplo, a partir de la tabla de probabilidades de la variable aleatoria X, vamos a encontrar la ecuación de la función de probabilidad de la variable aleatoria X en forma de ecuación:

x	1	2	3	4	5	6
$f(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Como todas las probabilidades son iguales a 1/6, entonces la ecuación de la función de probabilidad de X sería la siguiente:

$$f(x) = \frac{1}{6}$$

Y así ya tenemos nuestra función de probabilidad expresada mediante fórmula. Recordemos un poco de teoría:

Una función de probabilidad de una variable aleatoria discreta es una función que asigna probabilidades a los valores de la variable aleatoria. Su fórmula es la siguiente:

$$f(x) = P(X = x)$$

Además, la función de probabilidad tiene que cumplir con 2 propiedades:

1) La probabilidad de cada valor de la variable aleatoria debe estar entre 0 y 1.

$$0 \leq f(x) \leq 1$$

2) La suma de las probabilidades asignadas a todos los valores de la variable aleatoria debe ser 1.

$$\sum f(x) = 1$$

toda x

Estas propiedades seguro te sonarán conocidas, ya que son propiedades que hemos estudiado antes en el capítulo de probabilidades.

Es importante indicar que la función de probabilidad en una variable aleatoria discreta también es llamada función de masa de probabilidad o distribución de probabilidad. En algunos libros indican que la distribución y la función son cosas diferentes, pero la mayoría de los autores modernos no realizan distinción entre estos 2 conceptos en el caso de variables aleatorias discretas.

Ejemplo 2.-

Determine si la siguiente función es una función de probabilidad. Si lo es, encuentra la función de probabilidad en forma de tabla y el diagrama de barras de la función.

$$f(x) = \frac{x^2+5}{50}\ ,\ para\ x = 1,2,3,4$$

Desarrollo:

Para que f(x) sea una función de probabilidad, se tienen que cumplir las 2 propiedades que mencionamos líneas arriba, vamos a comprobarlo.

Empezamos colocando una tabla con los valores de x:

x	1	2	3	4

A continuación, calculamos los valores de f(x), teniendo en cuenta la fórmula del enunciado:

x	1	2	3	4
$f(x) = \frac{x^2+5}{50}$	$\frac{1^2+5}{50} = \frac{6}{50}$	$\frac{2^2+5}{50} = \frac{9}{50}$	$\frac{3^2+5}{50} = \frac{14}{50}$	$\frac{4^2+5}{50} = \frac{21}{50}$

1) Los valores de la función de probabilidad se encuentran entre 0 y 1.

$$0 \leq f(x) \leq 1$$

Se cumple la primera propiedad, pues todos los valores de f(x) se encuentran entre 0 y 1.

2) La sumatoria de los valores de f(x) tiene que ser 1:

$$\sum f(x) = 1$$

Veamos si se cumple realizando la sumatoria:

$$\sum f(x) = f(1) + f(2) + f(3) + f(4) = \frac{6}{50} + \frac{9}{50} + \frac{14}{50} + \frac{21}{50} = \frac{6+9+14+21}{50}$$

$$\sum f(x) = \frac{\cancel{50}}{\cancel{50}} = 1$$

Se cumple la segunda propiedad, ya que la sumatoria de los valores de f(x) es 1.

Como se cumplen las 2 propiedades, entonces podemos concluir que f(x) es una función de probabilidad.

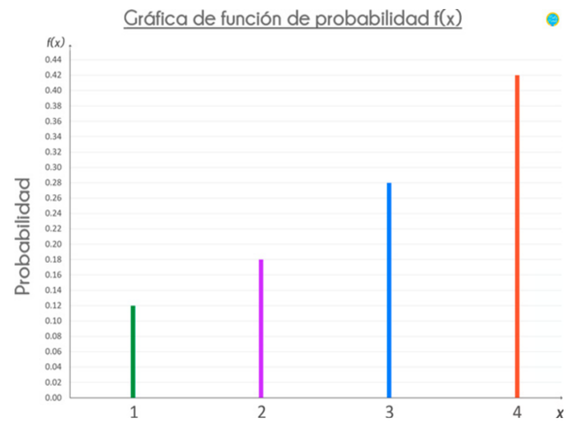
Ahora vamos a encontrar la función f(x) en forma de tabla, partiendo de la tabla que realizamos líneas arriba y simplificando las fracciones:

x	1	2	3	4
$f(x) = \frac{x^2+5}{50}$	$\frac{1^2+5}{50} = \frac{6}{\cancel{50}} = \frac{3}{25}$	$\frac{2^2+5}{50} = \frac{9}{50}$	$\frac{3^2+5}{50} = \frac{14}{\cancel{50}} = \frac{7}{25}$	$\frac{4^2+5}{50} = \frac{21}{50}$

A continuación, colocamos la misma tabla, pero ya ordenada y pasaremos las fracciones a decimales:

x	1	2	3	4
$f(x)$	0.12	0.18	0.28	0.42

Y terminamos con el diagrama de barras de la función:



Ejemplo 3.-

En base al siguiente ejercicio se puede determinar el número de caras (H) que se obtiene al lanzar 3 veces una moneda al aire, este es el experimento. Los posibles resultados 1, 2,3 caras. ¿Cuál es la distribución de probabilidad de número de caras?

Hay 8 posibles resultados.

En el primer lanzamiento puede caer cruz (T), otra cruz en el segundo y así sucesivamente las probabilidades.

Resultados Posibles	LANZAMIENTOS DE MONEDAS			Números de Caras
	PRIMERO	SEGUNDO	TERCERO	
1	T	T	T	0
2	T	T	H	1
3	T	H	T	1
4	T	H	H	2
5	H	T	T	1
6	H	T	H	2
7	H	H	T	2
8	H	H	H	3

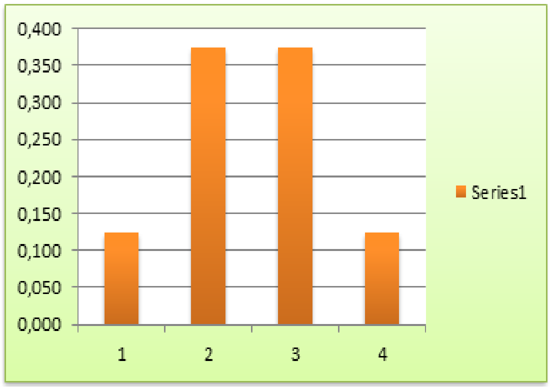
Observando en la tabla podemos decir que tenemos resultados de 0, 1, 2 y 3.

De modo que la probabilidad de 0 caras es 1/8 de la cara de 3 es 3/8 y así sucesivamente.

Distribución de probabilidad de número de caras

NÚMERO DE CARAS	PROBABILIDAD DE RESULTADO
0	1/8=0,125
1	3/8=0,375
2	3/8=0,375
3	1/8=0,125
TOTAL	8/8=1,000

El total de las probabilidades siempre será 1.



Taller Grupal

1.- Se lanza un par de dados. Se define la variable aleatoria como la suma de las puntuaciones obtenidas. Hallar la función de probabilidad.

2.- Un jugador lanza un dado corriente. Si sale o número primo, gana tantos cientos de euros como marca el dado, pero si no sale número primo, pierde tantos cientos de euros como marca el dado. Determinar la función de probabilidad.

3.- Sea una variable aleatoria discreta cuya función de probabilidad es:

X	p_i
0	0.1
1	0.2
2	0.1
3	0.4
4	0.1
5	0.1

a) Calcular la función de distribución.

b) Calcular las siguientes probabilidades:

$$P(X < 4.5)$$

$$P(X \geq 3)$$

$$P(3 \leq X < 4.5)$$

https://www.superprof.es/apuntes/escolar/matematicas/probabilidades/distribucion-binomial/ejercicios-de-distribuciones-discretas.html

1.23.Variable Aleatoria

"Cantidad que es el resultado de un experimento, y debido al azar, puede tomar valores diferentes".

En cualquier experimento aleatorio los resultados se presentan al azar y en consecuencia se habla de una variable aleatoria por ejemplo tirar un dado es un experimento.

Se puede presentar cualquiera de los resultados posibles. Algunos experimentos dan como resultados que son cuantitativos, como dólares, peso corporal, número de hijos, y otros dan resultados que son cualitativos: color, preferencia religiosa.

Ejemplo de variable Aleatotia

Se cuenta el número de empleados ausentes de un turno de trabajo del día lunes, el número puede ser 1, 2, 3,…….M. El número de inasistencias es la variable aleatoria.

Si se pesa un lingote de acero el resultado (en libras) puede ser 500; 500.2; 500.3; 500.4 dependiendo de la báscula.

Si se tiran al aire 2 monedas y se cuentan el número de caras el mismo puede ser 0, 1, 2. Puesto que el número de caras se debe al azar dicho número de caras es aleatorio.

Otras variables aleatorias podrían ser:

El número de lámparas defectuosas producidas durante una semana, las estaturas de los integrantes de un equipo de básquet femenino, la cantidad de corredores de un maratón, el número de automovilistas que cometie4ron una infracción por manejar bajo la influencia del alcohol.

1.24.Variable Aleatoria Discreta

"Una variable aleatoria discreta solo puede asumir cierto número de valores específicos y no admite ningún valor entre

valores consecutivos o claramente separados"

Ejemplo:

Si hay 100 empleados en una empresa, la cantidad de alumnos ausentes el día lunes. Una familia tiene 5 hijos 3 hombres y 2 mujeres.

Nota:

Una variable discreta en algunos casos puede tener valores enteros o fraccionarios estos valores deben estar separados, es decir, deben existir cierta distancia entre ellos.

Ejemplo:

Las puntuaciones otorgadas por jueces a aspectos técnicos y de forma artística son cifras decimales, como 6.8; 9.2; 9.9. Estos valores son discretos porque existen una distancia entre las calificaciones por lo que una calificación o puntuación no puede ser 7.25; 09.328, etc.

1.25.Variable Aleatoria Continua

"Se dice a la cual puede tomar un valor de una cantidad infinitamente grande de valores, dentro de ciertas limitaciones".

Ejemplos:

Si se mide algo como el ancho de una habitación, la altura de una persona, la presión de un neumático, la distancia entre ciudades, etc.

Las herramientas empleadas, así como las interpretaciones de probabilidad son diferentes según se trate de variables aleatorias discretas o continuas.

Estudiaremos por consiguiente temas importantes de distribución de probabilidad continua.

Media, varianza, desviación estándar de una distribución de probabilidad.

La media indica la ubicación central de los datos, la varianza describe su dispersión, de manera semejante una distribución de probabilidad se resume indicando su media y varianza.

La media de una distribución de probabilidad se denota con la letra griega μ(mu) minúscula y la desviación estándar con la letra griega sigma (σ).

1.26.Media

"La media es un valor típico que sirve para representar una distribución de probabilidad, también es el valor promedio a largo plazo de la variable aleatoria".

A la media de una distribución de probabilidad se le conoce como su valor esperado.

Esta media es un promedio ponderado en lo que los valores posibles se ponderan mediante sus probabilidades correspondientes de ocurrencia.

La media de una distribución de probabilidad discreta se calcula con la formula.

$$\mu = \sum [xP(x)]$$

VARIANZA

"Describe el grado de dispersión o variación en una distribución".

FÓRMULA

$$\sigma^2 = \sum [(x-\mu)^2 P(x)]$$

PASOS PARA CALCULAR LA VARIANZA

1) Restar la media por cada valor y elevar la diferencia al cuadrado.

2) Multiplicar el cuadrado de cada diferencia por su probabilidad.

3) Sumar los productos resultantes para obtener finalmente la varianza.

Nota:

$$\sigma = \sqrt{\sigma^2}$$

La desviación estándar "σ" se determina tomando la raíz cuadrada de la σ^2es decir:

Aplicación:

Andrés Analuisa vende automóviles nuevos de la agencia Chevrolet generalmente los sábados vende el mayor número de vehículos. E sr. Analuisa tiene la siguiente distribución de probabilidad de número de vehículos que espera vender en un día sábado en particular.

NÚMERO DE AUTOS VENDIDOS	PROBABILIDAD
	Px
0	0,10
1	0,20
2	0,30
3	0,30
4	0,10
TOTAL	1

1) ¿Qué tipo de distribución es esta?

2) ¿En un sábado común ¿Cuántos vehículos espera vender?

3) ¿Cuál es la varianza de la distribución?

DESARROLLO DEL PROBLEMA

Este es un ejemplo de distribución de probabilidad discreta porque observe que el sr. Anualiza solo espera vender una cantidad determinada de vehículos, no espera vender 5 o 50. Además no puede vender de 0, 1, 2, 3, 4 automóviles.

Así mismo los resultados son mutuamente excluyentes; no puede vender en total 3 o 4 vehículos en el mismo día sábado.

El número medio de automóviles vendidos se calcula ponderando la cantidad de autos vendidos con la probabilidad de vender ese número y luego se suman todos los productos aplicando la fórmula de la media, es decir:

$$\mu = \sum [xP(x)]$$

NÚMERO DE AUTOS VENDIDOS	Px	xP(x)
0	0,10	0,00
1	0,20	0,20
2	0,30	0,60
3	0,30	0,90
4	0,10	0,40
TOTAL	**1**	**2,10**

μ=0+0,20+0,60+0,90+0,40

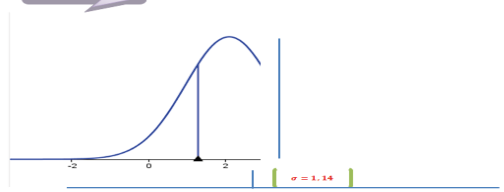
μ=2,10

Este valor indica que un gran número de sábados el sr. Analuisa espera vender un promedio de 2,1 vehículos por día (desde luego no es posible vender 2,1 autos en un sábado en particular). Por lo tanto a la media a veces se le denomina valor esperado.

Es un tipo para sistematizar los pasos para calcular la varianza.

N° DE AUTOS VENDIDOS	PROBABILIDAD				
	Px	xP(x)	x-μ	$(X-2)^2$	
0	0,10	0,00	-2,10	4,41	0,441
1	0,20	0,20	-1,10	1,21	0,242
2	0,30	0,60	-0,10	0,01	0,003
3	0,30	0,90	0,90	0,81	0,243
4	0,10	0,40	1,90	3,61	0,361
TOTAL	1	2,10			

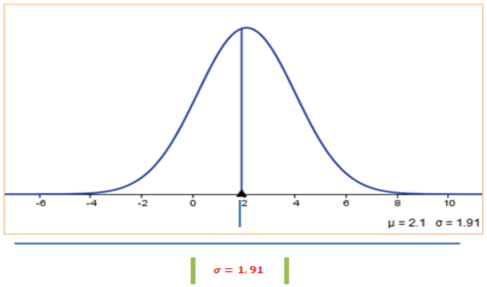
GRÁFICO



Comparando, si Jonathan también se dedica a la venta de automóviles quien vendió una cantidad media de autos fue de 1.91 de vehículos.

En base a los datos realice la comparación y saque la conclusión.

GRÁFICO



Conclusión:

Podemos deducir que Jonathan tiene más variedades de modelos de vehículos que el señor Analuisa.

Se recomienda incrementar publicidad, atención al público y que tenga más variedades de vehículos en stock.

A continuación, se presenta una fórmula alternativa para la varianza de una distribución de probabilidad discreta. Esta fórmula tiene la ventaja de evitar la mayor parte de restas.

$$\sigma^2 = \sum x^2 P(x) - \mu^2$$

DISTRIBUCIÓN DE PROBABILIDAD BINOMIAL

La distribución de probabilidad binomial es una distribución discreta que tiene muchísimas aplicaciones. Se asocia con un experimento de múltiples pasos que se llama experimento binomial. La palabra binomial viene de otra palabra que significa “dos nombres” y esto nos

hará recordar que en cada ensayo que veremos en este tema siempre habrá dos resultados: éxito y fracaso. Si es que respondes una pregunta de alternativas al azar, la respuesta es correcta o incorrecta. Si es que realizas un control de calidad a un producto, este será defectuoso o no defectuoso. Si es que lanzas una moneda, sale cara o sale cruz. Veamos a detalle la distribución binomial, su fórmula y por supuesto, muchos ejercicios.

Para estudiar esta distribución binomial, empezaremos conociendo los experimentos binomiales, como, por ejemplo, el lanzamiento de una moneda.

1)Experimento binomial

Un experimento binomial es un experimento que cumple las siguientes condiciones:

i) El experimento consta de una secuencia de n ensayos idénticos.

ii) En cada ensayo hay dos resultados posibles. A uno de ellos se le llama éxito y al otro, fracaso.

iii) La probabilidad de éxito es constante de un ensayo a otro, nunca cambia y se denota por p. Por ello, la probabilidad de fracaso será 1 – p. Esto se debe a que la probabilidad de éxito más la probabilidad de fracaso suman 1.

iv) Los ensayos son independientes, de modo que el resultado de cualquiera de ellos no influye en el resultado de cualquier otro ensayo.

Antes de resolver un ejercicio aplicando la fórmula de probabilidad binomial, tenemos que verificar siempre que se cumplen estas cuatro condiciones, pues esta fórmula solo funciona para experimentos binomiales.

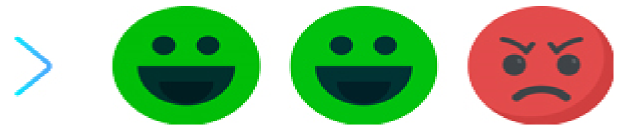
Esta ya se está poniendo muy aburrido, mejor veamos un ejemplo.

2) Las hamburguesas de Jorge

Últimamente con las clases de matemática no me va muy bien, por eso, he puesto mi cafetería. Preparo unas cosas riquísimas, pasteles, pizza, agua helada, pero lo más vendido son las matehamburguesas, las únicas hamburguesas que se venden con papas y leche chocolatada.

La probabilidad de que a un cliente nuevo le guste la matehamburguesa es de 0,8. Si vienen 3 nuevos clientes a mí cafetería ¿cuál será la probabilidad de que solo a dos de ellos les guste la hamburguesa?

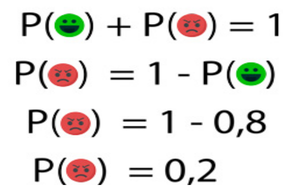
Si al cliente le gusta la hamburguesa tendrá carita feliz y si no le gusta, tendrá carita molesta. Por ejemplo, si a los dos primeros les gusta y al tercero no, colocaré esta gráfica:



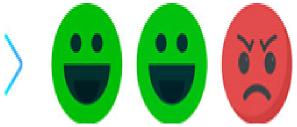
Pero no es la única opción, en total son tres opciones. Que a los dos primeros les guste y al tercero no. Que al primero y al tercero les guste, y al segundo no. O que les guste al segundo y tercero y al primero no. Graficamos estas tres opciones:

Si la probabilidad de que a un cliente le guste la matehamburguesa es de 0,8, entonces, la probabilidad de que no le gusten será:

$$P(\text{guste}) + P(\text{no guste}) = 1$$

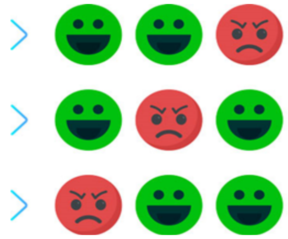


Ahora calcularemos la probabilidad de que ocurra cada evento. Por ejemplo, en la primera línea, calcularemos la probabilidad de que al primer y al segundo cliente les guste la hamburguesa y que al tercero no. Aplicaremos la regla de la multiplicación de probabilidades para eventos independientes.



$0{,}8 \times 0{,}8 \times 0{,}2$

Coloco ahora todas las opciones.

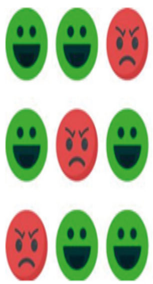


$0{,}8 \times 0{,}8 \times 0{,}2$

$0{,}8 \times 0{,}2 \times 0{,}8$

$0{,}2 \times 0{,}8 \times 0{,}8$

Esas mismas probabilidades, las colocaré en forma de potencias:

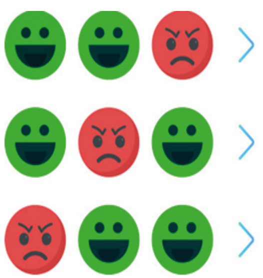


$0{,}8 \times 0{,}8 \times 0{,}2 = (0{,}8)^2 \times (0{,}2)^1$

$0{,}8 \times 0{,}2 \times 0{,}8 = (0{,}8)^2 \times (0{,}2)^1$

$0{,}2 \times 0{,}8 \times 0{,}8 = (0{,}8)^2 \times (0{,}2)^1$

Me quedaré solo con las potencias, eliminando el resto.

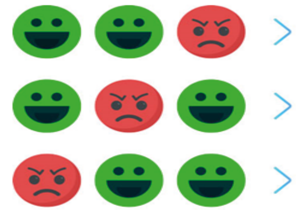


$(0{,}8)^2 \times (0{,}2)^1$

$(0{,}8)^2 \times (0{,}2)^1$

$(0{,}8)^2 \times (0{,}2)^1$

A continuación, calculamos la probabilidad total de que solo a 2 de los 3 clientes les guste mi hamburguesa empleando la regla de la suma de probabilidades para eventos mutuamente excluyentes.



$$
\begin{array}{r}
(0{,}8)^2 \times (0{,}2)^1 \; + \\
(0{,}8)^2 \times (0{,}2)^1 \; + \\
(0{,}8)^2 \times (0{,}2)^1 \; + \\
\hline
3 \times (0{,}8)^2 \times (0{,}2)^1
\end{array}
$$

Si esta multiplicación la metemos a la calculadora, nos quedaría 0,384.

¿Es un poco largo este cálculo? Pues sí, está larguísimo. Imagínate que ahora llegan 10 clientes nuevos a mi cafetería y queremos calcular la probabilidad de que a 4 de ellos les guste la matehamburguesa. Me tardaría horas, por eso es mejor usar la fórmula de la función de probabilidad binomial. Esta fórmula permite encontrar la probabilidad para cada posible valor de x (número de éxitos). Veamos los detalles:

3) Función de probabilidad binomial (fórmula)

Para un experimento binomial, sea p la probabilidad de "éxito" y 1-p la probabilidad de un "fracaso" en un solo ensayo; entonces la probabilidad de obtener x éxitos en n ensayos, está dada por la función de probabilidad f(x):

$$f(x) = P(X = x) = \binom{n}{x}(p)^x(1-p)^{n-x}, para\ x = 0, 1, 2, \ldots, n$$

Siendo el coeficiente binomial o número combinatorio

$$\boxed{\binom{n}{x} = \frac{n!}{x!\,(n-x)!}}$$

Por eso, en algunos libros encontrarás la función de probabilidad binomial f(x) con el coeficiente binomial ya incorporado y presentada de la siguiente manera:

$$f(x) = P(X = x) = \frac{n!}{x!\,(n-x)!}(p)^x \underbrace{(1-p)}_{q}{}^{n-x}, para\ x = 0, 1, 2, \ldots, n$$

¿Cuál de las formas es la mejor? Eso ya depende de lo que diga tu corazón, pero recuerda que el resultado es el mismo.

La función de probabilidad binomial se aplica a cualquier experimento binomial. Si una situación demuestra las propiedades de un experimento binomial y se conocen los valores de n y p, se puede usar la ecuación de arriba para calcular la probabilidad de x éxitos en n ensayos. Recuerda, antes de usar la fórmula de probabilidad binomial, siempre verifica que te encuentres ante un experimento binomial.

4) Media, varianza y desviación estándar de la distribución binomial

La media, la varianza y la desviación estándar se pueden encontrar con estas fórmulas:

Media:

$$\mu = np$$

Varianza:

$$\sigma^2 = np(1 - \mathrm{p})$$

Desviación estándar:

$$\sigma = \sqrt{np(1-p)}$$

Ejemplo 1

La probabilidad de que a un cliente nuevo le guste la matehamburguesa de Jorge es de 0,8. Si llegan 5 clientes nuevos a la cafetería, ¿cuál es la probabilidad de que solo a 3 de ellos les guste la matehamburguesa?

Desarrollo:

Antes de aplicar la fórmula, verificamos que se trate de un experimento binomial. Para ello, tiene que cumplir con las 4 condiciones que mencionamos arriba. Efectivamente, se trata de un experimento binomial.

En este caso, vamos a centrarnos en los clientes a los que les gusta esta hamburguesa, por ello diremos que:

X = número de clientes nuevos de 5 a los que les gusta la matehamburguesa

Entonces consideramos un éxito si al cliente le gusta esta hamburguesa.

Aplicaremos la fórmula binomial:

$$f(x) = P(X = x) = \binom{n}{x}(p)^x(1-p)^{n-x}$$

Ahora colocamos los valores de n, k y p. Recuerda que n es el número de ensayos, k el número de éxitos y p la probabilidad de éxito.

$$n = 5 \quad \wedge \quad x = 3 \quad \wedge \quad p = 0{,}8$$

Reemplazamos estos valores en la fórmula:

$$f(x) = P(X = x) = \binom{n}{x}(p)^x(1-p)^{n-x}$$

$$f(3) = P(X = 3) = \binom{5}{3}(0{,}8)^3(1-0{,}8)^{5-3}$$

$$f(3) = P(X = 3) = \frac{5!}{3!\,(5-3)!}(0{,}8)^3(0{,}2)^2$$

$$f(3) = P(X = 3) = \frac{5 \times 4 \times \cancel{3!}}{\cancel{3!}\,2!}(0{,}8)^3(0{,}2)^2$$

$$f(3) = P(X = 3) = \frac{20}{2 \times 1}(0{,}512)(0{,}04)$$

$$f(3) = P(X = 3) = 10(0{,}512)(0{,}04)$$

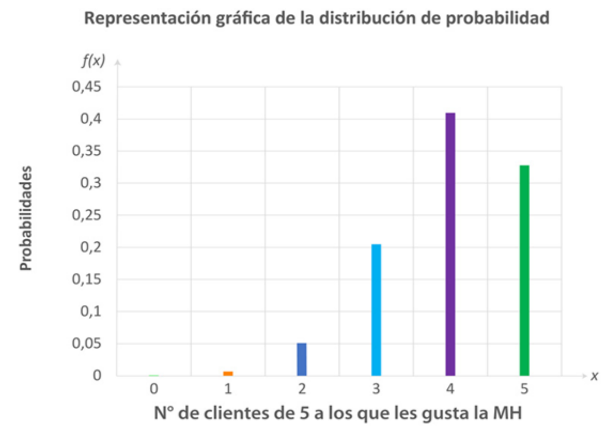
$$f(3) = P(X = 3) = \mathbf{0{,}2048}$$

La respuesta sería 0,2048.

Recuerda que X, nuestra variable aleatoria, es el número de clientes nuevos de 5 a los que les gustan las hamburguesas de Jorge. Aunque el problema no lo pide, vamos a elaborar la tablita de distribución de probabilidad. Para calcular todas las probabilidades, usaré la misma fórmula de arriba.

x	0	1	2	3	4	5
f(x)	0,0003	0,0064	0,0512	0,2048	0,4096	0,3277

Vamos a elaborar una gráfica para representar esta distribución de probabilidad, empleando un diagrama de barras.



Algunos usan los histogramas, también es válido.

Ejemplo 2

De todas las flores plantadas por una empresa de jardinería, el 90% sobrevive. Si se plantan 10 flores ¿cuál es la probabilidad de que 9 o más sobrevivan?

Desarrollo:

Antes de aplicar la fórmula, verificamos que se trate de un experimento binomial. Para ello, tiene que cumplir con las 4 condiciones que mencionamos arriba. Efectivamente, se trata de un experimento binomial.

En este caso, vamos a centrarnos en las flores que sobreviven, por ello diremos que:

X = número de flores de 10 que sobreviven

Entonces consideramos un éxito si la flor sobrevive. A las que flores que se mueren, las consideramos como un fracaso.

Aplicaremos la fórmula binomial:

$$f(x) = P(X = x) = \binom{n}{x}(p)^x(1-p)^{n-x}$$

Nos piden calcular la probabilidad de 9 o más sobrevivan

$$P(X \geq 9)$$

Este problema tiene trampa, porque dado que se plantaron 10 flores, la máxima cantidad de flores que pueden sobrevivir es 10, por lo tanto:

$$P(X \geq 9) = P(X = 9) + P(X = 10)$$

– ¿Y no pueden sobrevivir 11 flores?

– No se puede, porque solo se plantaron 10.

– ¿Y si las flores tienen hijitos bonitos?

– Alumno por favor, tome menos azúcar. Concéntrese y sigamos con la clase.

Ahora colocamos los valores de n, k y p. Recuerda que n es el número de ensayos, k el número de éxitos y p la probabilidad de éxito. En este caso:

$$n = 10 \quad \wedge \quad p = 0{,}9$$

Regresamos con la fórmula de arriba:

$$P(X \geq 9) = P(X = 9) + P(X = 10)$$

Vamos a calcular cada probabilidad por separado, empezando con P(X = 9):

$$P(X = 9) = \binom{10}{9}(0{,}9)^9(1-0{,}9)^{10-9}$$

$$P(X = 9) = \frac{10!}{9!\,(10-9)!}(0{,}9)^9(1-0{,}9)^{10-9}$$

$$P(X = 9) = \frac{10 \times \cancel{9!}}{\cancel{9!}1!}(0{,}9)^9(0{,}1)^1$$

$$P(X = 9) = 10(0{,}9)^9(0{,}1)^1$$

$$P(X = 9) = 0{,}3874$$

Continuamos con P(X

$$P(X = 10) = \binom{10}{10}(0{,}9)^{10}(1-0{,}9)^{10-10}$$

$$P(X = 10) = \frac{10!}{10!\,(10-10)!}(0{,}9)^{10}(0{,}1)^0$$

$$P(X = 10) = \frac{10!}{10!\,0!}(0{,}9)^{10}(1)$$

$$P(X = 10) = \frac{\cancel{10!}}{\cancel{10!}1}(0{,}9)^{10}(1)$$

$$P(X = 10) = 1(0{,}9)^9$$

$$P(X = 10) = 0{,}3487$$

Regresamos con esta fórmula:

$$P(X \geq 9) = P(X = 9) + P(X = 10)$$

Y reemplazamos lo calculado:

$$P(X \geq 9) = 0{,}3874 + 0{,}3487$$
$$P(X \geq 9) = 0{,}7361$$

Ejemplo 3

Considere un experimento binomial con dos ensayos y p=0,4.

a) Calcular la probabilidad de no obtener ningún éxito.

b) Calcular la probabilidad de obtener al menos 1 éxito.

Desarrollo:

Iniciamos definiendo la variable aleatoria de interés en nuestro experimento binomial:

X = número de éxitos en n ensayos.

x = 0; 1; 2.

El enunciado nos dice que: n = 2 y que p = 0,4; con ello podemos definir la función de probabilidad de X.

$$f(x) = P(X = x) = \binom{n}{x} p^x (1-p)^{2-x}$$
$$f(x) = P(X = x) = \binom{2}{x} 0{,}4^x (1-0{,}4)^{2-x}$$
$$f(x) = P(X = x) = \binom{2}{x} 0{,}4^x (0{,}6)^{2-x}$$

a) Calcular la probabilidad de no obtener ningún éxito: P(X = 0).

$$f(x) = P(X = 0) = \binom{2}{0} 0{,}4^0 (0{,}6)^{2-0}$$
$$f(x) = P(X = 0) = (1)(1)(0{,}6)^2$$
$$f(x) = (X = 0) = \mathbf{0{,}36}$$

Calcular la probabilidad de obtener al menos 1 éxito.

$$P(X \geq 1) = P(X = 1) + P(X = 2)$$

Pero:

$$f(1) = P(X = 1) = \binom{2}{1} 0{,}4^1 (0{,}6)^{2-1}$$

Reemplazamos:

$$P(X \geq 1) = P(X = 1) + P(X = 2)$$
$$P(X \geq 1) = 0{,}48 + 0{,}16$$
$$P(X \geq 1) = \mathbf{0{,}64}$$

Esa sería la respuesta: 0,64.

Una forma alternativa de desarrollar el apartado b, sería usando la siguiente

"La suma de las probabilidades para todos los resultados del experimento debe ser igual a 1".

$$\sum P(x) = 1$$

$$P(X=0)+P(X=1)+P(X=2)=1$$

$$P(X=1)+P(X=2)=1-P(X=0)$$

Pero recuerda que P(X=0) ya lo calculamos en el apartado a, y es igual a 0,36.

$$P(X=1)+P(X=2)=1-0{,}36$$

$$P(X=1)+P(X=2)=\mathbf{0{,}64}$$

1.27.Una Distribuccion Binomial tiene estas características

1) El resultado de cada ensayo de un experimento se clasifica en 2 categorías mutuamente excluyentes a saber: éxito o fracaso.

2) La variable aleatoria cuenta el número de éxitos en una cantidad fija de ensayos.

3) La probabilidad de éxito permanece igual a todos los ensayos. Lo mismo sucede con la probabilidad de un fracaso.

4) Los ensayos son independientes lo cual significa que el resultado de un ensayo no afecta el resultado de otro.

¿CÓMO SE CALCULA UNA DISTRIBUCIÓN BINOMIAL?

Para elaborar una distribución de probabilidad binomial se

necesita:

1) El número de ensayos

2) La probabilidad de éxitos en cada ensayo.

Por ejemplo:

Si un examen al término de un seminario de contabilidad y auditoría contiene 20 preguntas de opinión múltiple, el número de ensayos es 20.

Si cada pregunta tiene 5 opciones y solo 1 es correcta la probabilidad de éxito en cada ensayo que tiene una persona que desconoce la materia es de 0.20, de este modo la probabilidad de que una persona sin conocimiento del tema adivine la respuesta correcta a una pregunta tiene un valor de 0.20. Por lo tanto, se satisface las condiciones descritas para una distribución binomial.

Mediante la fórmula que tenemos a continuación podemos encontrar la distribución de probabilidad binomial.

$$P(x)=nCx\pi^{x}(1-\pi)^{n-x}$$

DONDE

P(x)= Probabilidad del evento a realizar

C= Combinación

n= Número de ensayos

π = Probabilidad de éxito de cada ensayo.

x= El número de éxitos.

Nota:

Se utiliza la letra griega π para representar un parámetro de población binomial.

No confundirse con la constante matemática π = 3.1416.

Aplicación:

Entre 2 ciudades hay 5 vuelos diarios si la probabilidad de que un vuelo llegue retrasado es de 0.20 ¿cuál es la probabilidad de que ninguno de los vuelos se retrase el día de hoy? ¿Cuál es la probabilidad de que exactamente llegue tarde hoy?

$$P(x) = nCx\pi^x(1-\pi)^{n-x}$$
$$P(0) = 5C_0(0.20)^0(1-0.20)^{5-0}$$
$$P(0) = 0.33$$

$$nCr = \frac{n!}{r!\,(n-r)!}$$
$$5C0 = \frac{5!}{(5-0)!}$$
$$5C0 = \frac{5!}{(5)!}$$
$$5C0 = \frac{5*4*3*2*1}{5*4*3*2*1}$$
$$5C0 = 1$$

Encontrar la probabilidad de que 1 de los 5 vuelos llegue retrasado.

$$P(x) = nCx\pi^x(1-\pi)^{n-x}$$
$$P(1) = 5C_1(0.20)^1(1-0.20)^{5-1}$$
$$P(1) = 0.410$$

$$nCr = \frac{n!}{r!\,(n-r)!}$$
$$5C1 = \frac{5!}{(5-1)!}$$
$$5C1 = \frac{5!}{(4)!}$$
$$5C1 = \frac{5*4*3*2*1}{4*3*2*1}$$
$$5C1 = 5$$

Conclusión:

No es factible de que 1 de los 5 vuelos llegue retrasado.

CALCULO DE LA TABLA DE DISTRIBUCIÓN BINOMIAL

n= 5

π=0.20

x= 0

$$P(x) = nCx\pi^x(1-\pi)^{n-x}$$
$$P(1) = 5C_0(0.20)^0(1-0.20)^{5-0}$$
$$P(1) = 0.33$$

$$nCr = \frac{n!}{r!\,(n-r)!}$$
$$5C0 = \frac{5!}{0!\,(5-0)!}$$
$$5C0 = \frac{5!}{1(5)!}$$
$$5C0 = \frac{5*4*3*2*1}{5*4*3*2*1}$$
$$5C0 = 1$$

n= 5

π=0.20

x= 1

$$P(x) = nCx\pi^x(1-\pi)^{n-x}$$
$$P(1) = 5C_1(0.20)^1(1-0.20)^{5-1}$$
$$P(1) = 0.410$$

$$nCr = \frac{n!}{r!\,(n-r)!}$$
$$5C1 = \frac{5!}{1!\,(5-1)!}$$
$$5C1 = \frac{5!}{(4)!}$$
$$5C1 = \frac{5*4*3*2*1}{4*3*2*1}$$
$$5C1 = 5$$

n= 5

π= 0.20

x= 2

$$P(x) = nCx\pi^{x}(1-\pi)^{n-x}$$
$$P(1) = 5C_{2}(0.20)^{2}(1-0.20)^{5-2}$$
$$P(1) = 0.2048$$

$$\mathbf{nCr = \frac{n!}{r!\,(n-r)!}}$$
$$5C2 = \frac{5!}{2!\,(5-2)!}$$
$$5C2 = \frac{5!}{(2*1)(3)!}$$
$$5C2 = \frac{5*4*3*2*1}{(2)(3*2*1)}$$
$$5C2 = 20$$

n= 5

π=0.20

x= 3

$$P(x) = nCx\pi^{x}(1-\pi)^{n-x}$$
$$P(3) = 5C_{3}(0.20)^{3}(1-0.20)^{5-3}$$
$$P(3) = 0.051$$

$$\mathbf{nCr = \frac{n!}{r!\,(n-r)!}}$$
$$5C3 = \frac{5!}{3!\,(5-3)!}$$
$$5C3 = \frac{5*4*3*2*1}{(3*2*1)(2)!}$$
$$5C3 = \frac{5*4}{2*1}$$
$$5C3 = 10$$

n= 5

π=0.20

x= 4

$$P(x) = nCx\pi^{x}(1-\pi)^{n-x}$$
$$P(4) = 5C_{4}(0.20)^{4}(1-0.20)^{5-4}$$
$$P(4) = 0.0064$$

$$\mathbf{nCr = \frac{n!}{r!\,(n-r)!}}$$
$$5C4 = \frac{5!}{4!\,(5-4)!}$$
$$5C4 = \frac{5*4*3*2*1}{(4*3*2*1)(1)!}$$
$$5C4 = \frac{5}{1}$$
$$5C4$$

n= 5

π=0.20

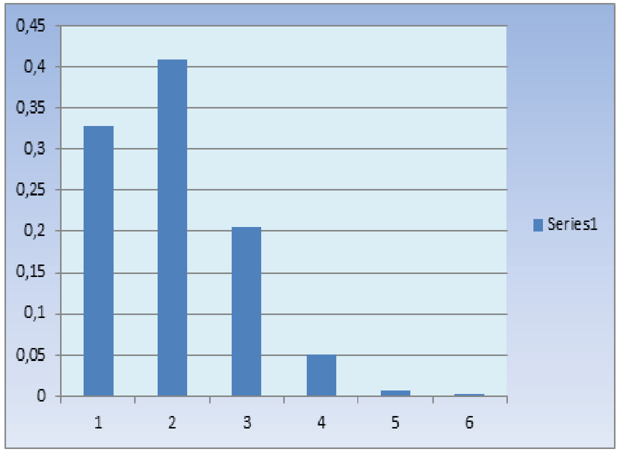
x= 5

$$P(x) = nCx\pi^{x}(1-\pi)^{n-x}$$
$$P(1) = 5C_{5}(0.20)^{5}(1-0.20)^{5-5}$$
$$P(1) = 0.0003$$

$$\mathbf{nCr = \frac{n!}{r!\,(n-r)!}}$$
$$5C5 = \frac{5!}{5!\,(5-5)!}$$
$$5C5 = \frac{5*4*3*2*1}{(5*4*3*2*1)(0)!}$$
$$5C5 = \frac{1}{1}$$
$$5C5 = 1$$

NÚMERO DE VUELOS ATRASADOS	Px
0	0.3277
1	0.4096
2	0.2048
3	0.0512
4	0.0064
5	0.003
TOTAL	1

GRÁFICO DE LA VARIABLE ALEATORIA



Conclusión:

La probabilidad de que un vuelo llegue tarde es mínima.

Aplicaciones

El Departamento de autopista de Florida informa que un puente levantadizo sobre el Rio de Gulp permanece levantado, bloqueando el tránsito de vehículos 20% de las veces usted debe pasar en el auto por esa ruta, una vez al día en los próximos 7 días y se desea predecir el número de veces que el puente estará elevado cuando usted se aproxime.

a) Esta situación satisface la hipótesis de distribución de la probabilidad Binomial

b) Calcule la probabilidad de que el puente se halle levantado cada vez que usted se acerque.

c) Utilizando la formula calcule la probabilidad de que el puente este levantado en 3 de las 7 veces que usted llegue ahí.

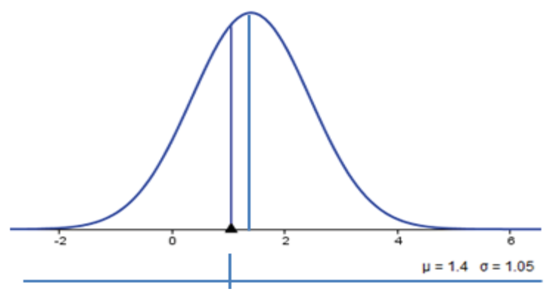
d) Aplique la formula, para determinar la probabilidad de que este levantado una vez.

e) Utilice la tabla de probabilidad Binomial para verificar sus respuestas a los mismos a, b c, d.

Respuestas

Días de semana X	P(X)	XPX	X-U	$(X-U)^2$	$(X-U)^2$ PX
0	0.2097	0.2097	-1.61	2.5921	0.5436
1	0.3670	0.3669	-0.61	0.3721	0.1365
2	0.2753	0.5506	0.39	0.1521	0.0419
3	0.1147	0.3441	1.39	1.9321	0.2216
4	0.02867	0.1148	2.39	5.7121	0.1639
5	0.0043	0.0215	3.39	11.4921	0.0494
6	0.0009	0.0024	4.39	19.2721	0.0077
7	0.0000	0.0000	5.39	29.0521	0.0004
	1.000	1.61			1.165

GRÁFICO



Conclusión:

No hay la probabilidad de que esté levantado una vez.

1.28.Distribucion de Posibilidades de Poisson

“Se usa para determinar la probabilidad de la ocurrencia en un número determinado de eventos cuando estos ocurren en un continuo de espacio o tiempo”

Existen muchas distribuciones de probabilidades discretas, pero para nuestro análisis se centrará solo en dos. La distribución binomial qu $\left(\sigma = 1,05 \right)$ acabamos de concluir y la distribución de Poisson que es el tema a tratarse.

“La distribución de Poisson se debe su nombre a Simeón Denisse Poisson de nacionalidad francesa.”

La distribución de Poisson se utiliza para describir ciertos tipos de procesos, vamos a dar varios ejemplos referentes al tema:

Entre las que se encuentra la distribución de llamadas telefónicas que llegan a un computador; La demanda (necesidades) de los pacientes que requieren servicio en una institución de salud; Las llegadas de camiones y automóviles a una caseta de cobro; el número de accidentes registrados en una intersección de calles; el número de pacientes que llegan a un consultorio médico en un cierto intervalo de tiempo.

1.29.Características de los procesos que producen una Distribuccion de Probabilidad de Poisson

Ejemplo:

El número de vehículos que pasan por una sola caja de una caseta de cobro en una hora pico, sirve como una ilustración de las características de la distribución de probabilidad de Poisson.

1. El promedio (media) del número de vehículos que llegan por hora pico, puede estimarse a partir de datos sobre el tráfico que se tengan disponibles.

2. Si dividimos la hora pico en periodos (intervalos) de un segundo cada uno, encontraremos que las siguientes afirmaciones son verdaderas.

a) La probabilidad de que exactamente un vehículo llegue a una caja por segundo es muy pequeña y es constante para cada intervalo de segundo.

b) La probabilidad de quedos o más vehículos lleguen en un intervalo de un segundo es tan pequeña que le podemos asignar un valor de cero.

c) El número de vehículos que llegan en un intervalo dado de un segundo es independiente del tiempo en que dicho intervalo se presenta en la hora pico.

d) El número de llegadas en cualquier intervalo de un segundo no depende del número de llegadas de cualquier otro intervalo de segundo.

La distribución de probabilidad de Poisson como hemos demostrado tiene que ver con ciertos procesos que pueden ser descritos por una variable aleatoria discreta.

La letra X por lo general representa a esta variable discreta, y puede tomar los valores enteros (0, 1,2); Utilizamos la letra mayúscula X para representar a la variable aleatoria; y la letra minúscula x para señalar un valor especifico de que dicha variable puede tomar.

Es decir, calcularíamos con la siguiente fórmula:

$$P(X) = \frac{\mu^{X} e^{-\mu}}{x!}$$

FÓRMULA DE POISSON

DONDE

Px: Probabilidad de que se va a calcular para un valor dado de x

μ: Es la media del número de ocurrencias (éxitos) en un intervalo específico

e: La constante cuyo valor es 2.71828

x: Número de ocurrencias o éxitos

x!: Factorial

MEDIA DE POISSON

$$\mu = n \,.\pi$$

DONDE

μ: Media

n: Número total de ensayos

π: Probabilidad de éxitos

Nota

$$\sigma^2 = n \,.\pi$$

La varianza de distibución de POISSON es igual a su media.

Ejemplo:

La probabilidad de que sea devuelto un cheque expendido por un banco es 0.0003 y se cambia a efectivo 10000 cheques.

¿Cuál sería el número medio de los documentos con rebote?

$$\mu = n \,.\pi$$

$$\mu = 10000(0{,}0003)$$

$$\mu = 3$$

Conclusión

Solo 3 cheques pueden ser rebotados.

Aplicaciones

Considerese en el aeropuerto de aerolíneas TAME, rara vez se pierde equipaje. En la myor parte de los vuelos no se observa un mal manejo de las maletas; algunos pasajeros reportan una valija perdida, unos cuantos tienen maletas perdidas, rara vez para un vuelo se tiene 3 y así sucesivamente supóngase que una muetra aleatoria de 1000 viajes aéreos revela un total de 300 maletas pérdidas.

¿Calcular con formula de POISSON la probabilidad de no perder ninguna maleta?

DATOS:

n= 1000

x= 300

μ=300/1000

μ=0.30

$$P(0) = \frac{(0.30)^0(2.71828)^{-0.3}}{0_!}$$

$$P(0) = \frac{1(0.7408)}{1}$$

$$P(0) = 0.7408 \quad = \quad 74\%$$

$$P(X) = \frac{\mu^X e^{-\mu}}{x_!}$$

$$P(1) = \frac{(0.30)^1(2.71828)^{-0.3}}{1_!}$$

$$P(1) = \frac{0.30\ (0.7408)}{1}$$

$$P(1) = 0.2222 \quad = 22\%$$

$$P(X) = \frac{\mu^X e^{-\mu}}{x_!}$$

$$P(2) = \frac{(0.30)^2(2.71828)^{-0.3}}{2_!}$$

$$P(2) = \frac{0.09\ (0.7408)}{2 * 1}$$

$$P(2) = 0.0333 \quad = 3.33\%$$

Aplicación

La señora Liliana Bayas está encargada de los préstamos en el Banco Costa Azul, con base a sus años de experiencia estima la probabilidad de que un solicitante no sea capaz de pagar su préstamo es de 0.25. El mes pasado realizo 40 préstamos.

¿Cuál es la probabilidad de que 3 préstamos no sean pagados a tiempo?

¿Cuál es la probabilidad de que por lo menos 3 préstamos no se liquiden a tiempo?

Datos:

n= 40

x= 3

μ= 0.025*40

μ= 10

$$P(X) = \frac{\mu^X e^{-\mu}}{x_!}$$

$$P(3) = \frac{(10)^3(2.71828)^{-10}}{3_!}$$

$$P(3) = \frac{1000\ (0.000045400)}{3 * 2 * 1}$$

$$P(3) = 0.0075667 \quad = \ 0.76\%$$

Conclusión:

No hay la probabilidad de que los préstamos sean pagados a tiempo.

EJERCICIOS DE APLICACIÓN

11. Supóngase una distribución binomial en la que n= y π= 0.60

a) Consulte el apéndice de A y enuncie la probabilidad, para valores de x desde 0 hasta3.

b) Determine la media y la desviación estándar de la distribución a partir de las definiciones generales dadas en las formulas (6.1) y (6.2). (6,4) y (6,5).

NÚMERO DE VECES "X"	PROBABILIDAD P(X)	X P(X)
0	0,064	0
1	0,288	0,288
2	0,432	0,864
3	0,216	0,648
TOTAL	1	μ=1,8

$0.-\ P(x) = C_x \pi^x (1-\pi)^{n-x}$ $\quad C_r = \frac{n!}{r!(n-r)!}$

$P(0) = C_0\ (o,60)^0 (1-0{,}60)^{3-0}$ $\quad C_0 = \frac{3*2*1}{0(3-0)}$

$P(0) = 1\ \ (o,60)^0 (0{,}40)^3$ $\quad C_0 = \frac{3*2*1}{0(3*2*1)}$

$P(0) = 0{,}064$ $\quad C_0 = 1$

$1.-\ P(x) = C_x \pi^x (1-\pi)^{n-x}$ $\quad C_r = \frac{n!}{r!(n-r)!}$

$P(1) = C_1\ (o,60)^1 (1-0{,}60)^{3-1}$ $\quad C_1 = \frac{3*2*1}{1(3-1)}$

$P(1) = 3\ \ (o,60)^1 (0{,}40)^2$ $\quad C_1 = \frac{3*2*1}{1(2*1)}$

$P(1) = 0{,}288$ $\quad C_1 = 3$

$2.-\ P(x) = C_x \pi^x (1-\pi)^{n-x}$ $\quad C_r = \frac{n!}{r!(n-r)!}$

$P(2) = C_1\ (o,60)^2 (1-0{,}60)^{3-2}$ $\quad C_2 = \frac{3*2*1}{2*1(3-2)}$

$P(2) = 3\ \ (o,60)^2 (0{,}40)^1$ $\quad C_2 = \frac{3*2*1}{2*1(1)}$

$P(2) = 0{,}432$ $\quad C_2 = 3$

$3.-\ P(x) = C_x \pi^x (1-\pi)^{n-x}$ $\quad C_r = \frac{n!}{r!(n-r)!}$

$P(3) = C_1\ (o,60)^3 (1-0{,}60)^{3-3}$ $\quad C_3 = \frac{3*2*1}{3*2*1(3-3)}$

$P(3) = \mathbf{1}\ (o,60)^3 (0{,}40)^0$ $\quad C_3 = \frac{3*2*1}{3*2*1(0)}$

$P(3) = 0{,}216$ $\quad C_3 = 1$

μ= ∑ [X P(X)]

μ= 0 + 0,288 + 0,864+ 0,648

μ= 1, 8

N*π (1-π)

3*(0, 60) (1-0, 60)

3*(0, 60) (0, 40)

0, 72

Conclusión:

La desviación estándar es favorable debido a que refleja un 72%.

13. Una encuesta de corretaje financiero (de EUA) reportan que 30% de los inversionistas individuales han empleado a un corredor de descuento; esto es, uno que no cobra las comisiones completas. Es una muestra seleccionada al azar de nueve inversionistas, ¿Cuál es la probabilidad de que:

a) Exactamente dos de los individuos de la muestra hayan empleado a un corredor de descuento?

b) Exactamente cuatro de ellos han recurrido a un corredor de ese tipo?

c) Ninguno haya recurrido a un corredor de descuento?

$$2.-\ P(x) = C_x \pi^x (1-\pi)^{n-x} \qquad C_r = \frac{n!}{r!(n-r)!}$$

$$P(2) = C_2\,(0{,}30)^2(1-0{,}30)^{9-2} \qquad C_2 = \frac{9\ast8\ast7\ast6\ast5\ast4\ast3\ast2\ast1}{2\ast1(9-2)}$$

$$P(2) = 36\,(0{,}30)^3(0{,}70)^7 \qquad C_2 = \frac{9\ast8\ast7\ast6\ast5\ast4\ast3\ast2\ast1}{2\ast1(7\ast6\ast5\ast4\ast3\ast2\ast1)}$$

$$P(2) = 0{,}27 \qquad C_2 = 36$$

Conclusión:

La probabilidad de que exactamente dos individuos hayan empleado a un corredor de descuento es nula con un 27%.

$$4.-\ P(x) = C_x \pi^x (1-\pi)^{n-x} \qquad C_r = \frac{n!}{r!(n-r)!}$$

$$P(4) = C_4\,(0{,}30)^4(1-0{,}30)^{9-4} \qquad C_4 = \frac{9\ast8\ast7\ast6\ast5\ast4\ast3\ast2\ast1}{4\ast3\ast2\ast1(9-4)}$$

$$P(4) = 126\,(0{,}30)^4(0{,}70)^5 \qquad C_4 = \frac{9\ast8\ast7\ast6\ast5\ast4\ast3\ast2\ast1}{4\ast3\ast2\ast1(5\ast4\ast3\ast2\ast1)}$$

$$P(4) = 0{,}17 \qquad C_4 = 126$$

Conclusión:

La probabilidad de que exactamente cuatro individuos hayan empleado un corredor de descuento es nula con un 17%.

$$4.-\ P(x) = C_x \pi^x (1-\pi)^{n-x} \qquad C_r = \frac{n!}{r!(n-r)!}$$

$$P(0) = C_0\,(0{,}30)^0(1-0{,}30)^{9-0} \qquad C_0 = \frac{9\ast8\ast7\ast6\ast5\ast4\ast3\ast2\ast1}{0(9-0)}$$

$$P(0) = 1\,(0{,}30)^0(0{,}70)^9 \qquad C_0 = \frac{9\ast8\ast7\ast6\ast5\ast4\ast3\ast2\ast1}{0(9\ast8\ast7\ast6\ast5\ast4\ast3\ast2\ast1)}$$

$$P(0) = 0{,}040 \qquad C_0 = 1$$

Conclusión:

La probabilidad de que ninguno haya recorrido a un corredor de descuento es nula con un 4%.

15. Los estándares de la industria automovilística de EUA indican que 10% de los autos nuevos requerirán servicio por garantía en el primer año. La agencia de Jones Nissan en Sumter, Carolina del Sur, vendió 12 automóviles en el mes pasado.

a) ¿Cuál es la probabilidad de que ninguno de estos autos necesite servicio de garantía?

b) Determine la probabilidad de que exactamente uno de ellos requiera tal servicio.

c) Determine la probabilidad de que exactamente dos automóviles lo necesiten.

d) Determine la media y la desviación estándar de esta distribución de probabilidad.

$P(0) = C_x \pi^x (1-\pi)^{n-x}$ $\quad C_r = \frac{n!}{r!(n-r)!}$

$P(0) = C_0\,(0{,}10)^0(1-0{,}10)^{12-0}$ $\quad C_0 = \frac{12*11*10*9*8*7*6*5*4*3*2*1}{0(12-0)}$

$P(0) = 1\,(0{,}10)^0(0{,}90)^{12}$ $\quad C_0 = \frac{12*11*10*9*8*7*6*5*4*3*2*1}{0(12*11*10*9*8*7*6*5*4*3*2*1)}$

$P(0) = 0{,}28$ $\quad C_0 = 1$

Conclusión:

La probabilidad de que ninguno de los automóviles necesite el servicio de garantía en el primer año es baja con un 28%.

$P(1) = C_x \pi^x (1-\pi)^{n-x}$ $\quad C_r = \frac{n!}{r!(n-r)!}$

$P(1) = C_1\,(0{,}10)^1(1-0{,}10)^{12-1}$ $\quad C_1 = \frac{12*11*10*9*8*7*6*5*4*3*2*1}{1(12-1)}$

$P(1) = 12\,(0{,}10)^1(0{,}90)^{11}$ $\quad C_1 = \frac{12*11*10*9*8*7*6*5*4*3*2*1}{1(11*10*9*8*7*6*5*4*3*2*1)}$

$P(1) = 0{,}38$ $\quad C_1 = 12$

Conclusión

La probabilidad de que exactamente uno de los automóviles necesite el servicio de garantía en el primer año es baja con un 38%.

$P(2) = C_x \pi^x (1-\pi)^{n-x}$ $\quad C_r = \frac{n!}{r!(n-r)!}$

$P(2) = C_2\,(0{,}10)^2(1-0{,}10)^{12-2}$ $\quad C_2 = \frac{12*11*10*9*8*7*6*5*4*3*2*1}{2*1(12-2)}$

$P(2) = 66\,(0{,}10)^2(0{,}90)^{10}$ $\quad C_2 = \frac{10*9*8*7*6*5*4*3*2*1}{2*1(10*9*8*7*6*5*4*3*2*1)}$

$P(2) = 0{,}23$ $\quad C_2 = 66$

Conclusión:

La probabilidad de que exactamente dos de los automóviles necesite el servicio de garantía en el primer año es casi nula con un 23%, ninguno de los dos vehículos necesitara el servicio.

$$\mu = n\pi$$

$$\mu = 12(0{,}10)$$

$$\mu = 1{,}2$$

$$\sigma = n\pi$$

$$\sigma = 12(0{,}10)$$

$$\sigma = 1{,}2$$

17. Un estudio reciente realizado por una asociación de contadores mostro que el 23% de los estudiantes de contaduría elijen el ramo de contaduría pública. Se selecciona una muestra de 15 estudiantes.

a) ¿Cuál es la probabilidad de que dos hayan seleccionado contaduría pública?

b) ¿Cuál es la probabilidad de que cinco hayan seleccionado contaduría pública?

c) ¿Cuántos estudiantes se espera que hayan seleccionado contaduría pública?

$$2.- P(x) = C_x \pi^x (1-\pi)^{n-x} \qquad C_r = \frac{n!}{r!(n-r)!}$$

$$P(2) = C_2\,(0{,}23)^4(1-0{,}23)^{9-4} \qquad C_2 = \frac{15*14*13*12*11*10*9*8*7*6*5*4*3*2*1}{2*1(15-2)}$$

$$P(2) = 105\,(0{,}23)^4(0{,}77)^5 \qquad C_2 = \frac{15*14*13*12*11*10*9*8*7*6*5*4*3*2*1}{2*1(13*12*11*10*9*8*7*6*5*4*3*2*1)}$$

$$P(2) = 0{,}19 \qquad C_2 = 105$$

Conclusión:

La probabilidad de que dos estudiantes hayan elegido contaduría pública es nula con un 19%.

$$5.- P(x) = C_x \pi^x (1-\pi)^{n-x} \qquad C_r = \frac{n!}{r!(n-r)!}$$

$$P(5) = C_5\,(0{,}23)^5(1-0{,}23)^{15-5} \qquad C_5 = \frac{15*14*13*12*11*10*9*8*7*6*5*4*3*2*1}{5*4*3*2*1(15-5)}$$

$$P(5) = 3003\,(0{,}23)^5(0{,}77)^{10} \qquad C_2 = \frac{15*14*13*12*11*10*9*8*7*6*5*4*3*2*1}{5*4*3*2*1(10*9*8*7*6*5*4*3*2*1)}$$

Conclusión:

La probabilidad de que cinco estudiantes hayan elegido contaduría pública es nula con un 14%.

CAPÍTULO II

DISTRIBUCIÓN DE UNA VARIABLE ALEATORIA CONTINUA

“La variable puede tomar cualquier valor que este en un intervalo de valores dados, y en los cuales la districucion es continua”

Una distribución de probabilidad continua que es muy importante es la distribución normal.

“Varios matemáticos han contribuido a su desarrollo, entre los que podemos contar al Astrónomo-Matemático del siglo XIX Karl Gauss. En honor a su trabajo, la distribución de probabilidad normal a menudo también se le llama distribución gaussiana”

Existen dos razones básicas para los cuales la distribución normal ocupa un lugar tan prominente en la estadística.

Primero. -Tiene algunas propiedades que la hacen aplicable a un gran número de situaciones en las que es necesario hacer inferencias mediante la toma de muestras.

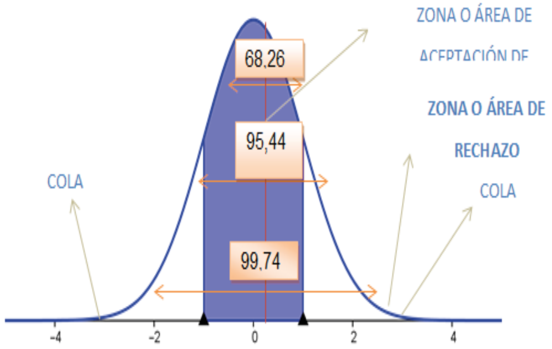
Segundo. - La distribución normal casi ajusta a las distribuciones de frecuencias reales observadas en muchos fenómenos, incluyendo características humanas (pesos, alturas) resultado de procesos físicos (dimensiones y rendimientos) y muchas otras medidas de interés, tanto en el sector público como en el privado.

2.1.Características de la distribucción normal de Probabilidad

Pone de manifiesto varias características importantes de la distribución normal de probabilidad, a continuación, tenemos lo siguiente:

1. La curva tiene un solo pico, por lo tanto, es unimodal, tiene la forma de campana.
1. La media de la población distribuida normalmente cae en el centro de la curva normal y es simétrica.
1. Debido a la simetría de la distribución normal de probabilidad, la media y la moda de distribución se encuentran también en el centro, en consecuencia, para una curva normal, la media, la mediana y la moda tienen el mismo valor.
1. Los dos extremos de la distribución normal de la probabilidad se extienden indefinidamente y nunca tocan el eje horizontal es decir es asintótica.

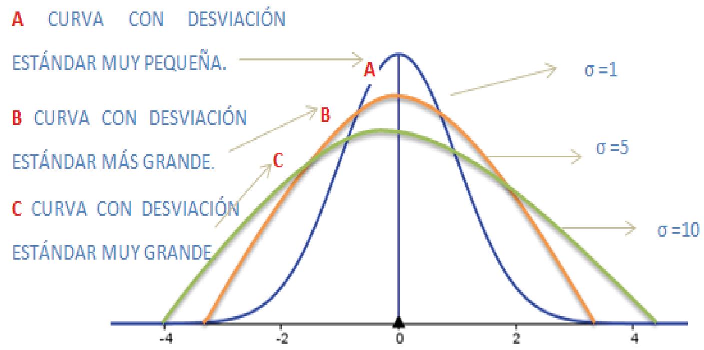
- ✓ Media
- ✓ Mediana
- ✓ Moda



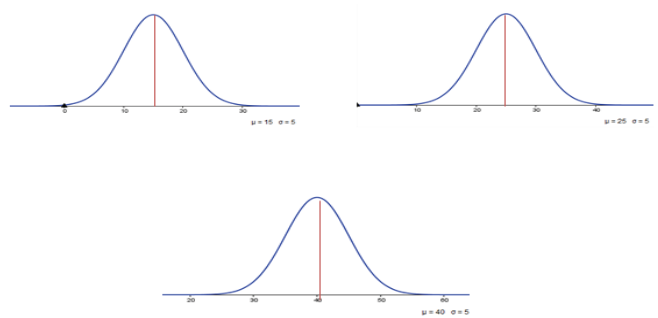
2.2. Áreas Bajo la Curva normal

La mayor parte de las poblaciones reales no se extienden de manera indefinida en ambas direcciones; con estas poblaciones, la distribución normal es una aproximación conveniente. No hay una sola distribución normal, sino una familia de curvas normales.

En el siguiente gráfico se muestran tres distribuciones normales de probabilidad, cada una de las cuales tiene la misma media, pero diferentes desviaciones estándar. Aunque estas curvas difieren una apariencia, las tres son "Curvas Normales".



Otra Distribución: Cuando tiene la misma desviación estándar, pero difiere la muestra.



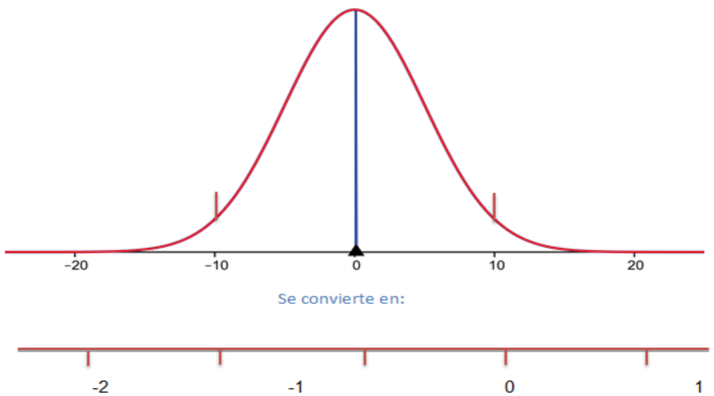
Conclusión:
La media no influye en la distribución estándar.

La mediación del área bajo la curva normal “no importa cuáles son los valores de μ y σ para una distribución de probabilidad normal, el área total bajo la curva es 1”

No es posible ni necesario tener una tabla distinta para cada curva normal posible. En lugar de ella podemos utilizar una distribución de probabilidad normal, estándar para encontrar áreas bajo cualquier curva normal.

Con esta tabla podemos determinar el área o la probabilidad de que la variable aleatoria distribuida normalmente esté dentro de ciertas distancias a partir de la media.

DISTRIBUCIÓN NORMAL ESTÁNDAR



$$Z = \frac{x - \mu}{\sigma}$$

DONDE:

X= Valor de la variable que nos preocupa.

μ= Media de distribución de la variable aleatoria

σ= Desviación estándar.

“Z”= Número de desviaciones, que hay desde x la media de distribución.

Aplicación:

En los siguientes gráficos servirá para reforzar nuestra desviación sobre distribuciones normales de probabilidad.

Con 2 intervalos cada uno con 1 desviación estándar de la media.

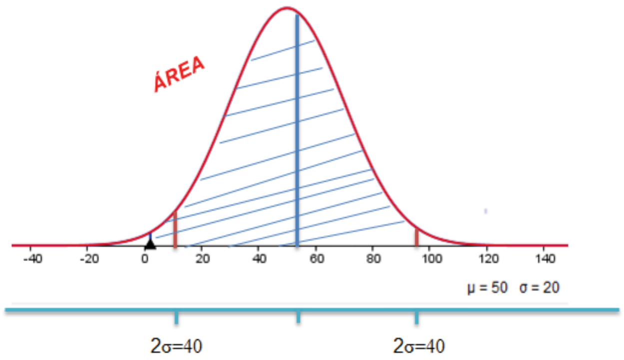
- Distribución A

Datos:

μ=50

σ=20

± 2 Intervalos



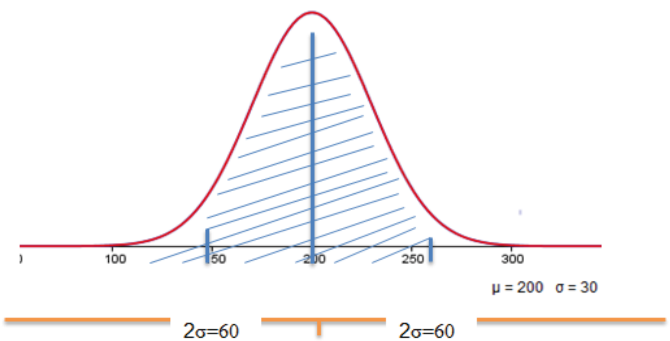
- Distribución B

Datos:

μ=200

σ=30

± 2 Intervalos



Conclusión:

Cada una de ellas tiene una media y una desviación estándar diferente y también podemos ver que las áreas sombreadas en ambas distribuciones en el 95% de área bajo la curva.

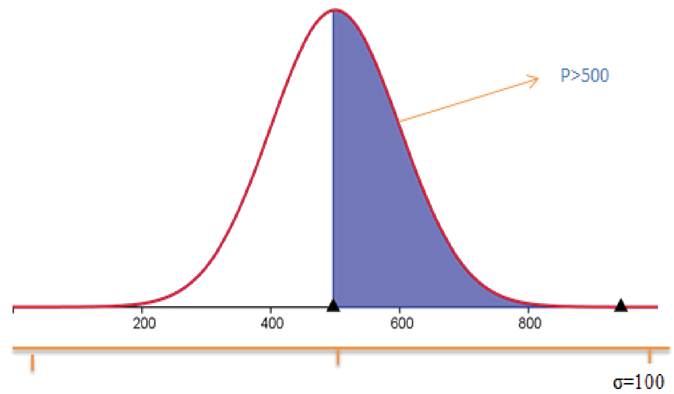
Aplicación 1 :

Cuál es la probabilidad de que un estudio de que ciertos, participantes en un evento en que el tiempo en completar el programa es de 500 horas y que la variable aleatoria normalmente distribuida tiene una desviación estándar de 100 horas.

Datos:

μ=500

σ=100



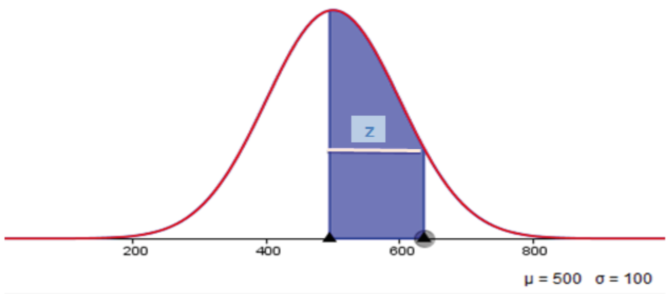
Aplicación 2

Cuál es la probabilidad con los mismos datos de que un candidato elegido al azar se tome entre 500 y 650 horas para completar un programa de entrenamiento.

Datos:

μ=500-650 Horas

σ=10



$$z = \frac{x - \mu}{\sigma}$$

$$z = \frac{650 - 5000}{100}$$

$Z = 1.5$ → Ver Tabla: $1.5 = 0.4332$

Conclusión:

La probabilidad de que un candidato escogido al azar entre 500 y 650 horas para terminar el programa de entrenamiento es ligeramente mayor a 0.4

Aplicación 3

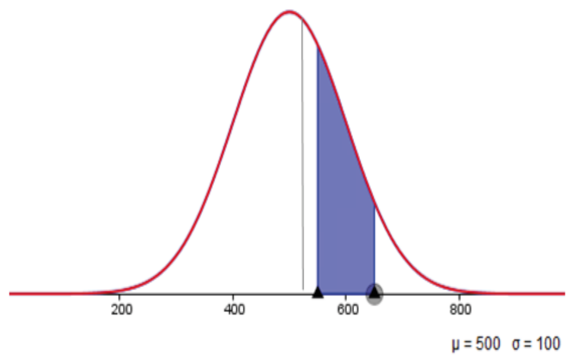
En los programas de entrenamiento el director desea saber la probabilidad de que un participante escogido al azar requiere entre 500 y 650 horas para completar el trabajo utilizaremos los datos anteriores.

Datos:

μ=500-650 Horas

σ=100

INTERVALO= 550-650



$$Z = \frac{X - \mu}{\sigma}$$

$$Z = \frac{650 - 500}{100}$$

$$Z = 1{,}5 = 0{,}4332$$

$$Z = \frac{X - \mu}{\sigma}$$

$$Z = \frac{550 - 500}{100}$$

$$Z = 0{,}5 = 0{,}1915$$

- ENTONCES:

0.4332 ⟶ Probabilidad de que la variable aleatoria este entre la media y 650 horas.
0.1915 ⟶ Probabilidad de que la variable aleatoria este entre la madia y 550 horas.
0.2417 ⟶ Probabilidad de que la variable aleatoria este entre las 550 y 650 horas.

Conclusión:

La probabilidad de que el participante pueda completar el programa de entrenamiento entre las 550 y 650 horas es baja.

Con información con ingresos semanales con los supervisores de turno de una industria de vidrio los ingresos semanales siguen una distribución normal con una media de $1000 y una desviación estándar de $100.

¿Cuál es la probabilidad de elegir a un supervisor de turno en la industria de vidrio?

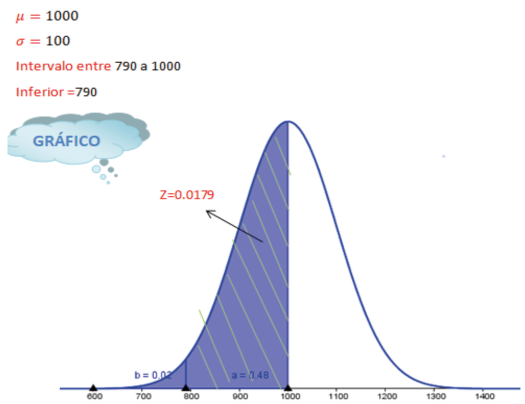
1) Este entre los $700 y $1000
2) Sea inferior a $790

μ=1000

σ=100

Intervalo entre 790 a 1000

Inferior =790



$$Z = \frac{X - \mu}{\sigma}$$

$$Z = \frac{790 - 1000}{100}$$

$$Z = -\frac{210}{100\sigma}$$

$$Z = 2.10$$

$$Z = 0.4821$$

Z= 48%

- INFERIOR A 790

$0.5000 - 0.4821$

0.0179

Z=2%

Conclusión:

La probabilidad de elegir un supervisor de turno con sueldo entre$ 790 y $1000 es de 48%, mientras que el 2% es la probabilidad mínima de elegir a un supervisor de turno con un sueldo inferior a \$790.

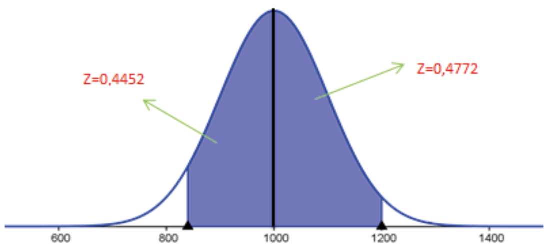
Volviendo a la distribución de los ingresos semanales con los supervisores de turno de una industria de vidrio los ingresos semanales siguen una distribución normal con una media de \$1000 y una desviación estándar de \$100.

¿Cuál es el área bajo la curva de los intervalos entre \$840 y \$1200?

μ=1000

σ=100

Intervalo entre 840 a 1200



$$z = \frac{X - \mu}{\sigma}$$

$$Z = \frac{840 - 1000}{100}$$

$$Z = -\frac{160}{100}$$

$$Z = 1.6$$

$$Z = 0.4452$$

Z=45%

$$Z = \frac{X - \mu}{\sigma}$$

$$Z = \frac{1200 - 1000}{100}$$

$$Z = \frac{200}{100}$$

$$Z = 2$$

$$Z = 04772$$

Z=48%

$$Z = 0.4452 + 0.4772$$

$$Z = 0.9224$$

Z=92%

Conclusión:

Si hay la probabilidad de elegir un supervisor de turno con sueldo entre\$ 790 y \$1000 es de 92%.

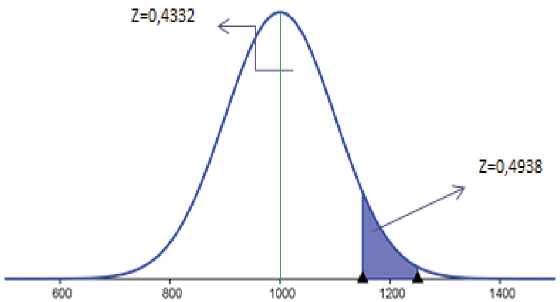
Considerando de nuevo el ejemplo de los ingresos semanales con los supervisores de turno de una industria de vidrio los ingresos semanales siguen una distribución normal con una media de \$1000 y una desviación estándar de \$100.

¿Cuál es el área bajo la curva de los intervalos entre \$1150 y \$1250?

μ=1000

σ=100

Intervalo entre \$1150 y \$1250



$$Z = \frac{X - \mu}{\sigma}$$
$$Z = \frac{1150 - 1000}{100}$$
$$Z = -\frac{150}{100}$$
$$Z = 1.5$$
$$Z = 0.4332$$
Z=43%

$$Z = \frac{X - \mu}{\sigma}$$
$$Z = \frac{1250 - 1000}{100}$$
$$Z = \frac{250}{100}$$
$$Z = 2.5$$
$$Z = 0.4938$$
Z=49%

$$Z = 0.4332 - 0.4938$$
$$Z = 0.0606$$
Z= 6%

Conclusión

No hay la probabilidad de elegir un supervisor de turno con

sueldo entre $1150 y $1250 es mínima del 6%.

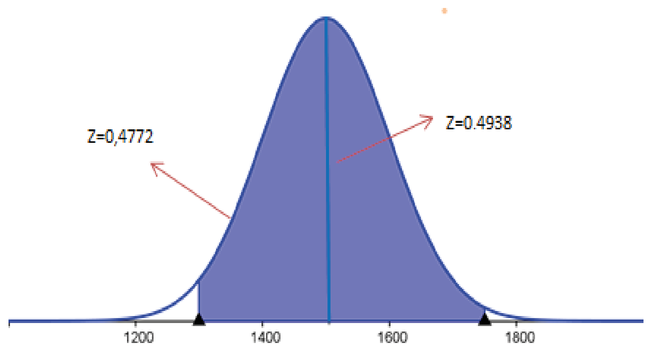
Aplicación :

Supóngase que en una asociación de comerciantes de productos agrícola se escoge al azar a un comerciante que tenga un ingreso entre $1300 y $1750 con una media de $1500 y una desviación estándar de $100.

μ=1500

σ=100

Intervalo entre $1150 y $1250.



$$Z = \frac{X - \mu}{\sigma}$$
$$Z = \frac{1300 - 1500}{100}$$
$$Z = -\frac{200}{100}$$
$$Z = 2$$
$$Z = 0.4772$$
Z=48%

$$Z = \frac{X - \mu}{\sigma}$$
$$Z = \frac{1750 - 1500}{100}$$
$$Z = \frac{250}{100}$$
$$Z = 2.5$$
$$Z = 0.4938$$
Z=49%

$$Z = 0.4772 + 0.4938$$
Z=0.9710

Z=97%

Conclusión:

Si hay la probabilidad de elegir un comerciante con sueldo entre \$1300 y \$1750 es 97%.

Taller

Con los ingresos semanales de los supervisores de turno de una industria de vidrio. Los ingresos siguen una distribución normal con una media de 1000\$.

Considerando el nuevo ejemplo de los ingresos semanales de los supervisores de la industria de vidrio con los datos ya conocidos la media de \$1000, desviación estándar \$100.

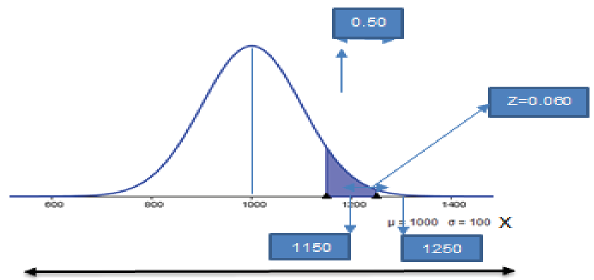
¿Cuál es el área bajo la, curva normal entre \$1150 y \$1250?

Datos

μ=1000

σ =100

Intervalo: 1150 y 1250



$$z = \frac{x - \mu}{\sigma}$$
$$z = \frac{1150 - 1000}{100}$$
$$z = 1.5 = 0.4332$$
$$z = \frac{x - \mu}{\sigma}$$
$$z = \frac{1250 - 1000}{100}$$
$$z = 2.5 = 0.4938$$

ENTONCES RESTAMOS:

0.4938
-0.4332
0.0606

Conclusión:

La probabilidad de elegir un supervisor de turno entre 1150 y 1250 con un 6% con una probabilidad mínima.

Aplicación 4

Con la información de los ingresos mensuales de los agentes de tránsito de la ciudad de Ambato, con una distribución normal, con una media de $900 y una desviación estándar de $100

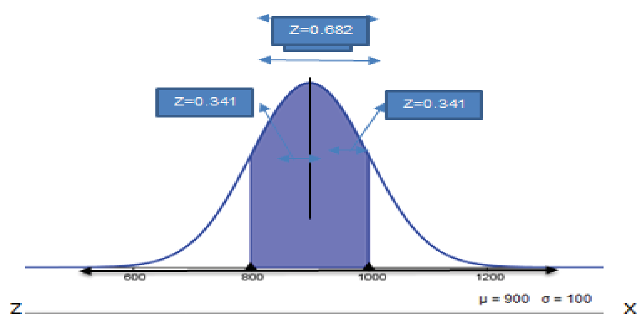
¿Cuál es el área bajo la curva normal entre los sueldos de $800 a $1000?

Datos

μ=900

σ =100

Intervalo: 800 y 1000



$$z = \frac{800 - 900}{100}$$

$$z = -1.0 = 0.3413$$

$$z = \frac{x - \mu}{\sigma}$$

$$z = \frac{1000 - 900}{100}$$

$$z = 1 = 0.3413$$

Conclusión:
La probabilidad de que los ingresos mensuales de los agentes de tránsito de la ciudad de Ambato es 68% teniendo una probabilidad favorable.

2.3. Muestreo

POBLACIÓN

Son todos los conjuntos de personas, objetos, cosas o datos acerca de los cuales se requiere información, los mismos que requieren ser definidos o identificados, sus propiedades particulares, antes de ser observados

Ejemplo:

Estudiantes de la Universidad Técnica de Ambato, Docentes Facultad Contabilidad y Auditoría, Equipos de futbol del Ecuador, Número de autos vendidos en una concesionaria, Discografía de un artista.

PARÁMETRO

Es una media usada para describir alguna característica de una población o universo.

Para simbolizar la población, se utiliza símbolos griegos ejemplo la media aritmética , desviación típica σ = (sigma), número de población N=Población.

ESTADÍGRAFO

Es una medida usada para describir algunas características de una muestra para lo cual se utiliza letras latinas para diferenciar del universo o población ejemplo:

La media aritmética representamos por media de ẋ, desviación típica S, número de elementos N; etc.

Para poder determinar una muestra representativa es necesario escoger adecuadamente cada elemento, de acuerdo a un criterio específico y a condiciones bien controladas. Esto requiere de varios pasos:

1. Definir la naturaleza de la población con lo que se va a trabajar.
2. Tipo de diseño de la muestra.
3. Grado de precisión de la muestra, lo suficientemente representativos y adecuado que permita “Inferir”, características de la población de estudio.

Inferir: Extraer un juicio o conclusión a partir de hechos, proposiciones o principios, sean generales o particulares.

Con respecto a la definición de la población, es decir identificar con claridad la naturaleza de sus elementos que la componen, para considerar la correcta selección de la muestra.

Ejemplo:

Para obtener la información acerca del salario promedio de los profesores universitarios.

Luego de haber determinado bien la población, el investigador procede a seleccionar los elementos que tomaron parte de la muestra, para ello debe utilizar técnicas correctas que permitan disminuir al máximo los errores y/o bien la comodidad del investigador.

2.4. Muestra

Es un subconjunto del conjunto universo o población, cuyos elementos discrepan y se asemejan en algunas características respecto al universo dado.

Para que una muestra permita estimar adecuadamente un parámetro, es necesario que tenga características de representatividad, condición esta que no es fácil garantizar, por lo que se acude comúnmente al muestreo aleatorio (azar) que tiene la particularidad de poder determinar la no representatividad, luego de un largo proceso de muestreos similares.

El muestreo aleatorio simple garantiza a todos los elementos de una población la posibilidad de pertenecer a la muestra o muestras que se tomen de ella. Este sistema de la selección de los elementos de una muestra, con su variante; el muestreo aleatorio estratificado pertenece al grupo de muestras probabilísticas.

"Cuando la población es pequeña (menor a cien) el muestreo no tiene objetivo toda vez que se puede utilizar el total de la población".

2.5. Tipos o clases de muestra

Las clases de muestra representativa son numerosas, dependiendo del tiempo, dinero y habilidad disponible para obtenerlos, por lo que revisamos las siguientes:

1. De acuerdo con el número de muestras tomadas de una población

1.1. Muestreo simple: Cuando una población se toma solamente una muestra para la toma de decisiones.

1.2. Muestreo Doble: Cuando el análisis de la primera muestra no es decisivo, entonces se requiere de una segunda para emitir una decisión.

1.3. Muestreo Múltiple: Cuando se requiere muestras sucesivas con la finalidad de llegar a emitir una decisión.

2. De acuerdo con las formas de seleccionar los elementos de

una muestra

2.1. Probabilidad: Todos los elementos tienen igual posibilidad de ser seleccionados en la muestra.

2.1.1. Aleatorias: Todos los elementos de una población tienen la posibilidad de ser seleccionados en la muestra.

2.1.2. Al Azar Simple: Mediante un listado de todos los elementos se sortean mediante medios mecánicos para obtener la muestra; por ejemplo: escribir el nombre de cada elemento y colocar en un recipiente para extraerlo, uso de un programa de sorteo en computadora.

2.1.3. Al Azar Sistemático: Es igual al anterior, pero aplicando el siguiente procedimiento:

Primero. Determinar la constante K.

DONDE:

K= Constante de referencia

N= Elementos de la población

n= Unidades deseadas de una muestra

- **Segundo.** Se efectúa el sorteo para elegir un número que sea inferior o igual al valor "A".

A + 0 k

A + 1 k

A + 2 k

A + 3 k

.

.

.

A + (n-1) k

- **Tercero.** Todos los números obtenidos serán los elementos de la muestra.

Aplicación:

Datos:

N= 2760

n= 60

$$k = \frac{N}{n}$$

$$k = \frac{2760}{60}$$

$$k = 46$$

Tomando "A" = 40, pues es menor que K

Reemplazando = A + 0 k

40+0(46)=40

40+1(46)=86

40+2(46)=132

40+3(46)=178

40+4(46)=224

40+59(46)=2754

- Cuarto.

Conclusión:

De acuerdo al listado los elementos seleccionados serán 40, 86, 132, 178, 224… y 2745.

2.1.4. **MUESTREO POR CONGLOMERADOS:** Todos los elementos de la muestra son seleccionados en grupos, parte de las mismas características similares, es decir, el universo está dividido en un número infinito de conjuntos, escogiendo alguno de ellos. Por ejemplo, parroquias, barrios, manzanas, empresas, aldeas.

2.1.5. **MUESTRAS ESTRATIFICADAS:** Consiste en subdividir a la población en estratos de acuerdo a sus características, para luego extraer al azar un cierto número de elementos en cada uno de los grupos homogéneos, para determinar en qué manera actuaría la gente de acuerdo a un problema relacionado con la educación; se puede subdividir a la población de acuerdo a la edad, sexo, niel económico, educacional, religioso, etc.

2.1.6. **MUESTRA PROPORCIONAL:** Permite otorgar la mayor representatividad. Esta técnica requiere que, de los diversos extractos se seleccione al azar un número de unidades o elementos de acuerdo con su proporción de población total que represente a cada uno de los grupos. Ejemplo: así, el 20% de la población electoral está compuesta por graduados; entonces el 20% debe obtenerse de este grupo

3. NO PROBABILÍSTICAS

No se puede estimar la probabilidad de que cada elemento estará en la muestra.

3.1. **MUESTRA ACCIDENTAL:** Los elementos son todos tomados en forma

circunstancial. Ejemplo: los 50 primeros curos que representen la muestra de un colegio.

3.2. **MUESTRAS POR CUOTAS:** Cuando los elementos de la muestra se determinan en igual cantidad de cada categoría, es decir, se toma igual porción. Ejemplo: 35 mujeres y 35 hombres.

3.3. **MUESTRA INTENCIONADA POR EL CRITERIO:** Elección de unidades en forma fortuita según características de relevancia. Ejemplo: cuando se identifica una muestra en lugares donde mayor votación ha tenido un candidato.

MUESTRA

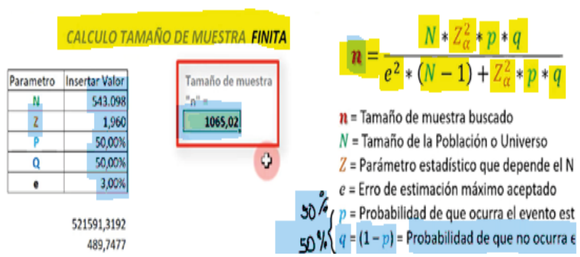
Es un subconjunto del conjunto universo o población, cuyos elementos se asemejan y discrepan en algunas características respecto al universo dado.

Para que una muestra permita estimar adecuadamente un parámetro, es necesario que tenga características de representatividad, condición que no es fácil de garantizar, por lo que se acude comúnmente al muestreo aleatorio (azar), que tiene 1 particularidad de poder determinar la representatividad luego de un largo proceso de muestreos similares.

El muestreo aleatorio simple garantiza a todos los elementos de una población la posibilidad de pertenecer a la muestra o muestras que se tomen de ella. Este sistema de selección de los elementos de una muestra, con su variante: el muestreo aleatorio estratificado pertenece al grupo de muestras probabilísticas.

Cuando la población es pequeña (menor a cien), el muestreo no tiene objeto toda vez que puede utilizarse la población total.

FÓRMULAS PARA EL CÁLCULO DE LA MUESTRA



N

Calcular el tamaño de muestra para una población de 543.098 consumidores de una marca de bebida gaseosa "y", donde el investigador asigna un nivel de confianza de 95% y un margen de error de 3%. Donde se desconoce la probabilidad "p" del evento

Nivel de confianza	Z_{alfa}
99,7%	3
99%	2,58
98%	2,33
96%	2,05
95%	1,96
90%	1,645
80%	1,28
50%	0,674

CALCULO TAMAÑO DE MUESTRA **INFINITA**

Parametro	Insertar Valor
Z	1,960
P	50,00%
Q	50,00%
e	3,00%

0,9604
0,09%

Tamaño de muestra "n" =
1067,11

$$n = \frac{Z_{\alpha}^{2} * p * q}{e^{2}}$$

n = Tamaño de muestra buscado
Z = Parámetro estadístico que depende el N
e = Erro de estimación máximo aceptado
50% p = Probabilidad de que ocurra el evento est
50% $q = (1 - p)$ = Probabilidad de que no ocurra

Nivel de confianza	Z_{alfa}
99,7%	3
99%	2,58
98%	2,33
96%	2,05
95%	1,96
90%	1,645
80%	1,28
50%	0,674

Calcular el tamaño de muestra para una **población desconocida**, donde el investigador asigna un nivel de confianza de **95%**, un margen de error de **3%** y se desconoce la probabilidad "p" del evento que se está estudiando.

FORMULA PARA EL CÁLCULO DEL TAMAÑO DE MUESTRA POBLACION FINITA:

$$n = \frac{z^2 pqN}{E^2(N-1) + z^2 pq}$$

FORMULA PARA EL CÁLCULO DEL TAMAÑO DE MUESTRA POBLACION INFINITA:

$$n = \frac{z^2 pq}{E^2}$$

n = Tamaño de la muestra
N = Tamaño de la Población o Universo
Z = Parámetro estadístico que depende de el Nivel de confianza
e = Error de estimación máximo aceptado
p = Probabilidad de que ocurra el evento
q = (1 - p) Probabilidad de que no ocurra el evento

Tabla normal de variación de nivel de confianza

90%	1.62
91%	1.71
92%	1.79
93%	1.85
94%	1.90
95%	1.96
96%	2.12
97%	2.25
98%	2.38
99%	2.58

Taller de aplicación. -

Ejemplo 1. Se requiere realizar una ivestigación acerca de la Percepción de los estudiantes sobre las clases virtuales; en la carrera de ingeniería comercial en Utepsa, para ello se recabaron los siguientes datos:

N. de estudiantes de la carrera Ing.Comercial:850 alumnos

Nivel de confianza: 90%

Error muestra: 5%

R. Alumnos.

Tareas:

Ejemplo 2.

Se realizan elecciones para elegir al nuevo director de cierta institución educativa, que consta de 5 universidades, el total de alumnos es de 10100 se requiere tomar una muestra para conocer la tendencia del voto entre los alumnos.Para este estudio se requiere un nivel de confianza del 95%, un porcentaje de error del 3%,si además se conoce que la tendencia de aceptación de cierto candidato es del 35%.

Ejemplo 3.

Se desea determinar cuantas personas en el mercado de Cuitláhuac. Vera Cruz han comprado sandia en el último mes con el fin de verificar si la idea de venderlas sería buena.Para esto se requiere obtener una muestra y aplicar un estudio estadístico, muestra requerida debe contar con un 90% de nivel de confianza, un error máximo del 10%, si se observa que un dia determinado 6 de cada 10 personas compran sandía en el mercado.

Ejemplo 4.

Para un municipio se repartiran 100 paquetes electorales, cada paquete va a constar de 750 boletas, se desea corroborar que en ninguno de los paquetes falte algo, para esto se realizara una muestra representativa de los paquetes que cuenten con un nivel de confianza del 95%, un error máximo del 10%, si se conoce que años anteriores sobre el 5% de los paquetes han tenido boletas faltantes.

Fórmulas de otros Autores

1. $n = \frac{z^2 NPq}{(n-1)E^2 Z^2 Pq}$

2. $n = \frac{N}{E^2(N-1)+1}$

3. $n = \frac{N\,\sigma^2 Z^2}{(N-1)\,E^2 \sigma^2 Z^2}$

4. $n = \frac{NPq}{(N-1)\left(\frac{E^2}{K^2}\right)+(Pq)}$

5. $n = \frac{(Z^2)PqN}{(NE^2)+PqZ^2}$

DONDE:

σ = Desviación estándar

N = Población

σ_X = Desviación media

X = Media

E = Error estándar

P = Población de ocurrencia (0.5)

q = Probabilidad de no ocurrencia (0.5)

Nc = Nivel de confianza (7) 1.96 95%

K = Coeficiente de error

σ^2 = Varianza

Z = Nivel de confianza

Aplicaciones

Un investigador quiere conocer la proporción de estudiantes que ven un determinado programa de televisión. Si la población es de 2762 y se suma un valor de error del 10% y un 95% de nivel de confianza. ¿Qué tamaño de la muestra debe tomar?

DATOS:

N= 2762

E= 10% / 100 = 0.1

Z= 95% en tabla = 1.96

$\sigma^2 = Pq = (0.5)^2$

$$n = \frac{N\,\sigma^{2z^2}}{(N-1)E^2\sigma^2Z^2}$$

$$n = \frac{(2762)(0.5^2)(1.96^2)}{(2762-1)0.1^2 0.5^2 1.96^2}$$

$$n = \frac{2652.6248}{28.57}$$

$$n = 92.85 = 93$$

Conclusión :
Se debe tomar la muestra de 93 personas de una población de 2762, con un margen de confiabilidad del 95%.

Para una investigación de campo, mediante estimación, se

propone un margen de error del 6% y un nivel de confianza del 99%.

Datos:

E= 6% / 100 = 0.6

Z= 99% en tabla = 2.58

$\sigma^2 = (0.5)^2$

$$n = \frac{\sigma^2\, Z^2}{E^2}$$

$$n = \frac{(0.5)^2\,(2.58)^2}{(0.06)^2}$$

$$n = \frac{1.6641}{0.0036}$$

$$n = 462.25$$

Conclusión:

Por el tamaño de la muestra, es de imaginarse que la población es demasiado extensa, tan ves ascienda a los 4000 elementos.

2.6. Usos frecuentes de desviaciones estándar

SESGO

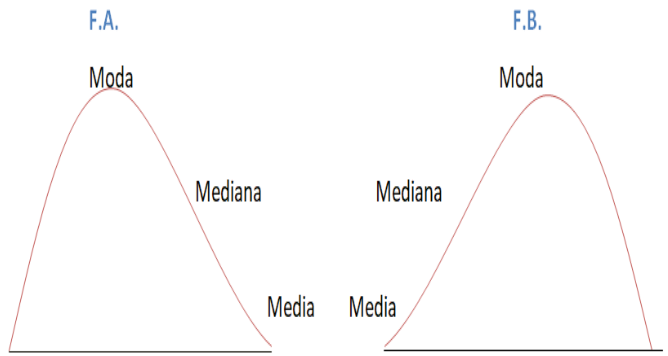
No todas las distribuciones son normales, algunas están sesgadas a la izquierda o a la derecha en la gráfica que comparemos se encuentran las causas de distribución. Ejemplo peso de las personas.

FIGURA A

Se dice que la distribución está a la derecha. Parecería que pocas personas las más pesadas que están en el extremo superior de la escala de peso, si algunas quizá algunos hombres más altos dan la tabla derecha. En la segunda distribución de peso.

FIGURA B

Unas cuantas mujeres diminutas harán la distribución hacia el extremo inferior, derecho que se desvíe a la izquierda



Conclusión:

En ambos casos, la moda es por definición, la observación que más ocurre con más frecuencia por lo tanto está en el pico de la distribución. Sin embargo como dijo anteriormente, por du naturaleza la media se ve más afectada por las observaciones

extremas. Por lo tanto, es halada en la dirección de sesgo más de lo que esta la mediana la cual está en algún sitio la media y la moda.

PARA MEDIR EL SESGO

El sesgo puede medirse mediante el coeficiente de Pearson.

$$P = \frac{3(\bar{X} - \text{MEDIANA})}{S}$$

CON LAS SIGUIENTES CONDICIONES:

Si $P < 0$ → Los datos están sesgados hacia la izquierda

Si $P > 0$ → Los datos están sesgados hacia la derecha

Si $P = 0$ → Están distribuidos normalmente

Ejemplo:

Utilizando la lista de estudiantes de la UTA, se calcula la media 78,7 , la desviación estándar 12,14 Ϭ, mediana 78,33 Me, dado la información, el rector puede ver claramente que los datos están sesgados a la derecha, debido que la media excede a la mediana además se desea una media de grado de sesgo.

Datos:

$\overline{X} = 78{,}7$

$S = 12{,}14$

$Me = 78{,}33$

$$P = \frac{3(\overline{X} - Me)}{s}$$

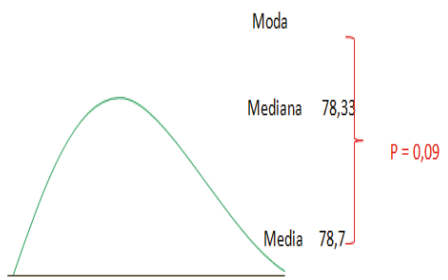
$$P = \frac{3(78{,}7 - 78{,}3)}{12{,}14}$$

$$P = \frac{1{,}11}{12{,}14}$$

$$P = 0{,}09$$

Conclusión:

Sesgo es mayor que cero eso significa que está sesgado a la derecha



2.7. Intervalo de confianza

"Conjunto de valores obtenidos a partir de los datos muéstrales en la que hay una determinada probabilidad que se encuentra en parámetro"

Ejemplo:

Una empresa de jeans el ingreso anual medio de los trabajadores es de $25000 dólares el intervalo de esta estimación puede ser de $61000- $6900 dólares.

Este seguro que este parámetro poblacional se encuentra en el intervalo dando una probabilidad.

Se puede indicar por ejemplo que se tiene una seguridad del 90% de que el salario anual medio de los trabajadores de jeans de esta región $61000- $69000

La información que se obtenga acerca de la forma de distribución muestra, de la media muestra, es decir la distribuciones muestra permite localizar un intervalo que tenga determinado probabilidad de contener a la media poblacional.

El teorema de límite establece lo siguiente.

Calculando escogemos **95%** $\xrightarrow[\text{oo}]{\text{VER TABLA}}$ 1.96 ⇨ Se observa que se encuentra la desviación estándar a la media poblacional

Cuando establecemos **99%** $\xrightarrow[\text{oo}]{\text{VER TABLA}}$ 2.58 ⇨ Se observa que se encuentra la desviación estándar a la media poblacional

Aquí las desviación estándar de la que se trata, es la desviación estándar de la desviación muestra.

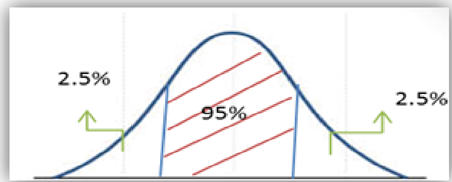
A esta desviación se llama error estándar.
A los intervalos calculamos de esta manera se llama intervalo

de confianza.

$95\% \xrightarrow[\text{oo}]{\text{VER TABLA}} 1.96$ ⇒ 5% QUEDA EN DOS COLAS.

$5\% \div 100 = 0.05$ ⇒ NIVEL DE SIGNIFICANCIA

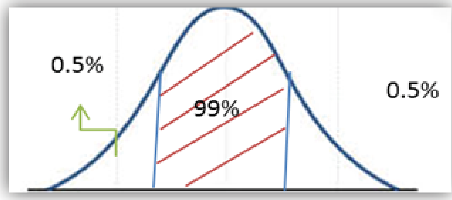
$Z0.05 \div 2 = 0.025$



$99\% \xrightarrow[\text{oo}]{\text{VER TABLA}} 2.58$ ⇒ 5% QUEDA EN DOS COLAS.

$1\% \div 100 = 0.01$ ⇒ NIVEL DE SIGNIFICANCIA

$0.01 \div 2 = 0.005$



2.8.Introducción a la Hipótesis Estadísticas

Conceptos Generales

-El mejor ejemplo para entender la prueba de hipótesis es el veredicto de un jurado.

-La hipótesis nula y alternativa

H0: el acusado es inocente.

H1: el acusado es culpable

.

-La acusación proviene de una sospecha de culpabilidad. -La hipótesis H0 (status quo) se establece en oposición a H1 y se mantiene a menos que se apoye H1 con evidencia "más allá de una duda razonable". -El no rechazo de H0 no implica inocencia, sino que la evidencia fue insuficiente para lograr una condena. De esta manera el jurado no necesariamente acepta H0, sino que no la rechaza.

Una prueba o contraste de una hipótesis estadística es una regla o procedimiento que conduce a una decisión de aceptar o rechazar cierta hipótesis con base en los resultados de una muestra.

Los procedimientos de prueba de hipótesis dependen del empleo de la información contenida en una muestra aleatoria de la población de interés.

Si esta información es consistente con la hipótesis se concluye que es verdadera.

En caso contrario, la información es inconsistente con la hipótesis, se concluye que es falsa.

Estadístico de prueba. Valor obtenido a partir de la información muestral, se utiliza para determinar si se rechaza o no la hipótesis.

Toda prueba de hipótesis lleva implícita un estadístico para probarla. Este estadístico depende del problema planteado y de la codificación de las variables.

Región crítica. Región de rechazo de la hipótesis nula H0.

Se denomina C (subconjunto del espacio muestral) a la región crítica de una prueba dada si nos lleva a rechazar la hipótesis nula H0 cuando la muestra cae en la región C.

Valor crítico. El punto que divide la región de aceptación y la región de rechazo de la hipótesis nula H0 .

Nivel de significancia. Probabilidad de rechazar la hipótesis nula cuando es verdadera. Se denota con α.

Error Tipo I. Rechazar la hipótesis nula cuando en realidad es verdadera. La probabilidad de cometer un error tipo I se le llama nivel de significancia.

Error Tipo II. Aceptar la hipótesis nula cuando en realidad es falsa.

La probabilidad de cometer el error tipo II es imposible de calcular a menos que tengamos una hipótesis alternativa específica. Se denota con β.

Al probar cualquier hipótesis estadística, hay cuatro situaciones posibles que determinan si nuestra decisión es correcta o errónea (ver la siguiente tabla).

		Situaciones Posibles	
		H_0 verdadera	**H_0 falsa**
Decisiones Posibles	**No rechazar H_0**	Decisión correcta	Error tipo II
	Rechazar H_0	Error tipo I	Decisión correcta

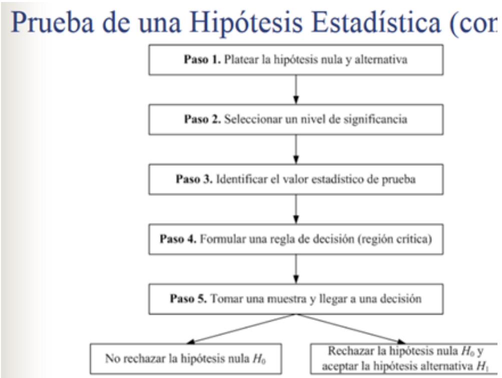
$$\alpha = P(\text{error tipo I}) = P(\text{Rechazar } H_0 | H_0 \text{ verdadera}) = P(\text{Muestra dentro de } C | H_0)$$

$$\beta = P(\text{error tipo II}) = P(\text{No rechazar } H_0 | H_0 \text{ falsa}) = P(\text{Muestra fuera de } C | H_1)$$

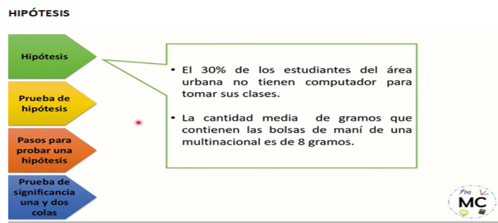
Los valores para la significancia de una prueba de uso más común son 0.01, 0.05 y 0.10; o sea, el investigador está dispuesto a permitir 1%, 5% o 10% de cometer un error tipo I.

Idealmente se desearía que la probabilidad de error tipo I fuera igual a cero. Sin embargo, si se desea α = 0, nunca se podría tomar la decisión de rechazar la hipótesis nula.

La decisión de rechazar la hipótesis nula es importante, ya que la decisión se basa en una muestra y no en la población; por lo cual existe la posibilidad de cometer un error tipo I



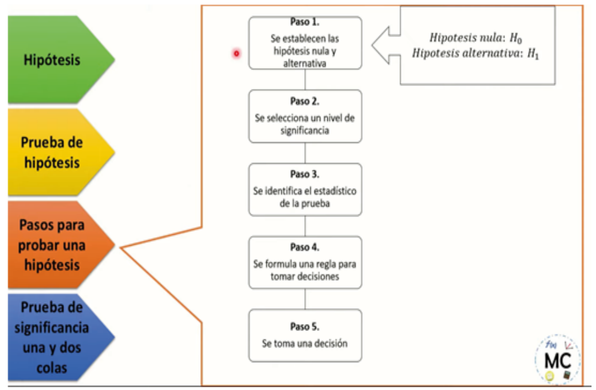
Una hipótesis es un enunciado acerca de un parámetro poblacional que va a estar sujeta a verificación



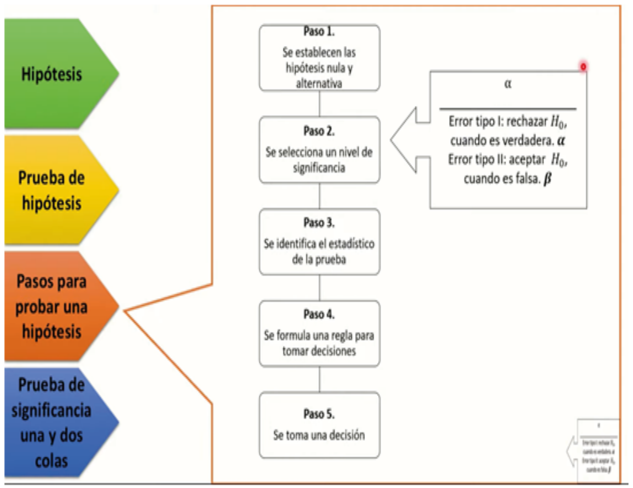
La Prueba de Hipótesis va a ser entonces un procedimiento en el cual se muestra evidencias para verificar dicha hipótesis.

Pasos para probar una hipótesis

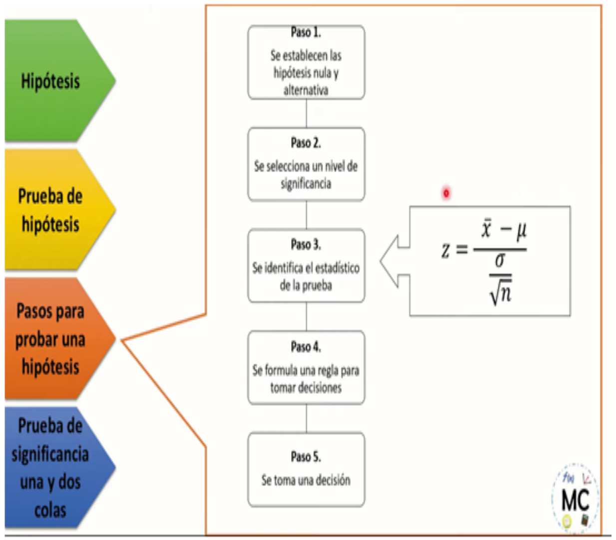
Paso1.-



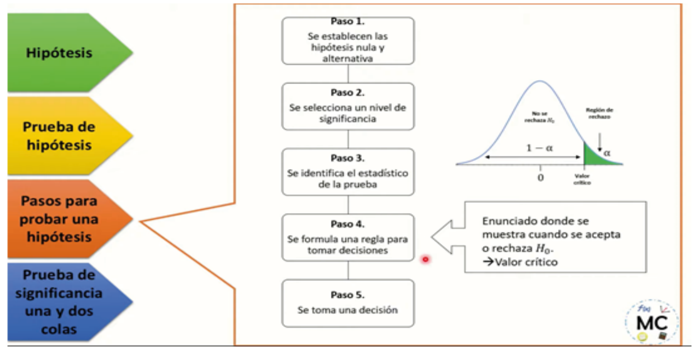
PASO 2.-



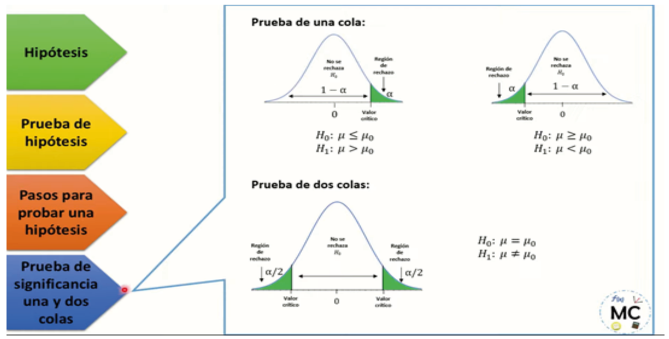
PASO 3.-

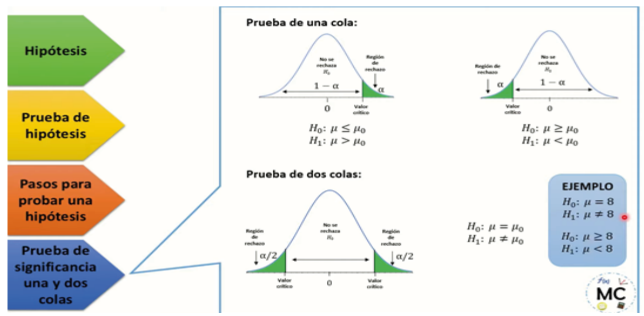


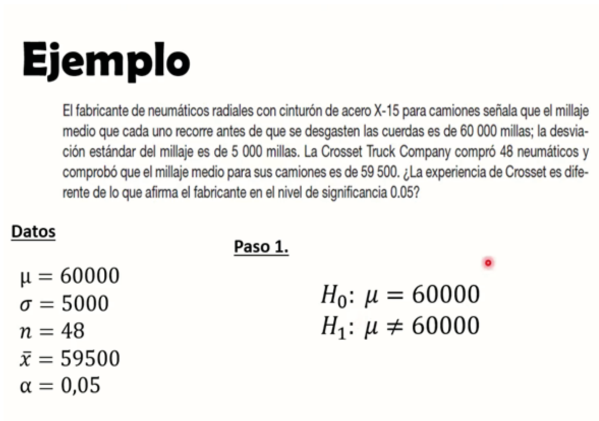
PASO 4 Y 5 .-



Prueba de significancia







M = media poblacional

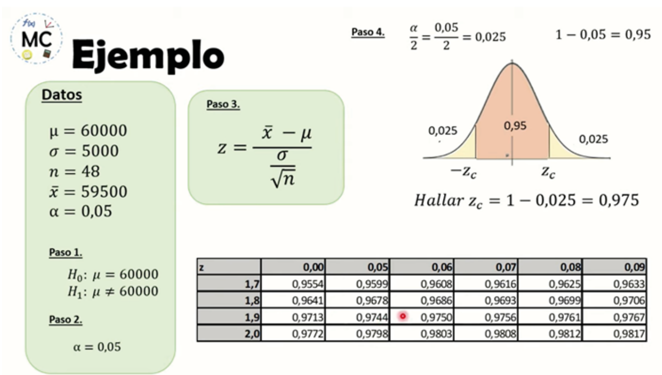
Σ = desviación estandar

N = datos

Ẋ = media muestral

A = nivel de significacia.

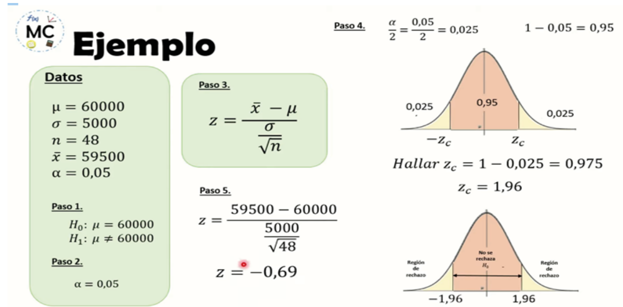
Se trata de una prueba a dos colas



z	0,00	0,05	0,06	0,07	0,08	0,09
1,7	0,9554	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9798	0,9803	0,9808	0,9812	0,9817

El Z crítico será -1.96 y 1.96, obtenido de la Tabla de Distribución en la parte de las Probabilidades.

se va a aceptar la hipótesis nula si el valor de z calculado se encuentra entre- 1.96 y 1.96 si queda por fuera se va a la región de rechazo.



Como el valor de z =-0.69 cae dentro de la región entre -1.96 y 1.96 de no rechazo, se puede concluir que la medida de la población no es distinta de 60.000.

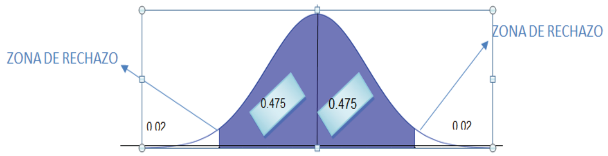
FUENTE: Prueba de hipótesis de una muestra ejercicios resueltos - YouTube

ERROR ESTÁNDAR DE LA MEDIA CUANDO NO SE CONOCE LA DESVIACIÓN ESTÁNDAR POBLACIONAL

“Cuando no se conoce la deviación estándar poblacional se calcula por:

vereficación del 95%

$$S_{\bar{x}} = \frac{S}{\sqrt{n}}$$



- SUMANDO

0.4750+0.5000=0.9500 95%

CÁLCULO DEL INTERVALO DE CONFIANZA

Partamos de un ejemplo donde necesitamos hacer una investigación para determinar el salario de los egresados de economía. Se calculó la media muestral que es $27000.00 y la desviación estándar es de $200. Suponemos que la muestra tiene 80 observaciones. El intervalo de confianza está en un 95% entre 26608 y 27392.

Para su cálculo utilizaremos las siguientes fórmulas:

$$\bar{x} \pm 1.96\frac{s}{\sqrt{n}} \longrightarrow 95\%$$

$$\bar{x} \pm 2.58\frac{s}{\sqrt{n}} \longrightarrow 99\%$$

$$IC\% = \bar{X} \pm Z\,(S_{\bar{X}})$$

DONDE:

IC= Intervalo de Confianza

$\bar{X}$ = Media Muestral

Z = Coeficiente de Confianza

$S_{\bar{X}}$ = Error típico de la muestra

n = Muestra

NOTA: Cuando el tamaño de la muestra es al menos igual a 30 generalmente se acepta la teoría del límite central asegura una distribución normal de las medias muéstrales.

En los cálculos se puede utilizar la distribución estándar normal "Z"

$\bar{X} \pm Z\frac{s}{\sqrt{n}}$ $\longrightarrow$ Intervalo confianza para la media

Aplicaciones:

Un departamento de flora y fauna silvestre ha proporcionado un alimento especial para crías de truchas arcoíris en un estanque, una muestra de los pesos de 40 pececillos revelo que su media muestral es de 402.7 gramos, la desviación estándar de la muestra 8.3 gramos.

a) ¿Cuál es el precio medio estimado de la población? ¿Cómo se denomina el valor estimado?

b) ¿Cuál es el intervalo de confianza del 99%?

c) ¿Cuáles son los límites de confianza del 99%?

d) ¿Qué grado de confianza se utiliza?

e) Interprete los resultados.

$\bar{X}$ = 402.7 gr.

S = 8.8 gr.

n = 40

402.7. gr. en 40 pececillos. Estimación Puntual.

$$\bar{X} \pm Z \frac{S}{\sqrt{n}}$$

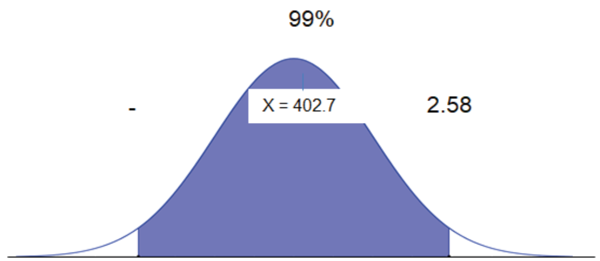
$$402.7 \pm 2.58 \frac{8.8}{\sqrt{40}} 402.7 \pm 2.58 \frac{8.8}{6.32}$$

$$402.7 \pm 2.58\,(1.39)$$

$$402.7 \pm 3.59$$

402.7 + 3.59 = 406.29 LIMITE SUPERIOR

402.7 – 3.59 = 399.11 LIMITE INFERIOR



Conclusión:

Para una muestra de 40 pececillos 402.7 y en la cual tenemos un intervalo de confianza entre 399.11 y 406.29 el cual tiene un nivel de confianza del 99%.

-Una muestra de 10 observadores se selecciona a partir de una población normal para la cual se conoce que la desviación

estándar es 5 y la media muestral vale 20.

a) Determine el error estándar de la muestra

b) Explique porque se puede utilizar la formula para obtener el intervalo de confianza de 95% aunque el tamaño de la muestra es menor que 30.

c) Determine el intervalo de confianza de 95% para la media de la población.

Cuando la población es menor que 30 no es necesario tomar una muestra

Datos:

$\bar{X} = 20$

$S = 5$

$n = 10$

$$\bar{X} \pm Z \frac{S}{\sqrt{n}}$$

$$20 \pm 1.96 \frac{5}{\sqrt{10}}$$

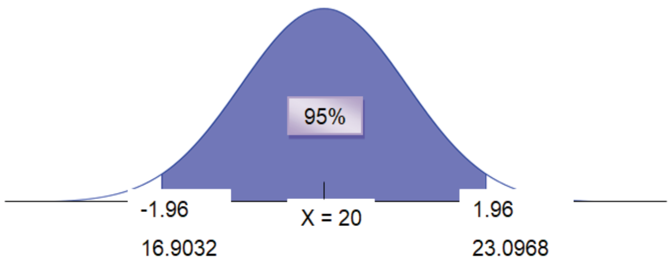
$$\bar{2}0 \pm 1.96 \frac{5}{3.16}$$

$$20 \pm 1.96\ (1.58)$$

$$20 \pm 3.0968$$

$20 + 3.0968 = 23.0968$ LIMITE SUPERIOR

$20 - 3.0968 = 16.9032$ LIMITE INFERIOR



Conclusión:

De un total de 20 observaciones tenemos que el límite superior es 23.0968 y el límite inferior es 16.9032.

INTERVALO DE CONFIANZA PARA UN INTERVALO DE CONFIANZA

Proporción: "P"

"Es una fracción, razón o porcentaje que indica la parte de la muestra o población que tiene características determinadas"

Ejemplo

Noventa y dos de Cien encuestados están de acuerdo con el horario de verano para el ahorro de energía eléctrica.

LA PROPORCIÓN MUESTRAL ES: $\frac{92}{100} = 0.92 \cong 92\%$

Si presenta una proporción muestral, "X" es el número de éxitos "n" número de objetos muestreados por lo tanto podemos calcular mediante la siguiente fórmula:

PROPORCIÓN MUESTRAL: $P = \frac{X}{n}$

CALCULOS:

- INTERVALO DE CONFIANZA PARA PROPORCION

$P \pm Z \sigma p$

- INTERVALO DE CONFIANZA PARA UNA PROPORCION POBLACIONAL

- ERROR ESTANDAR DE LA PROPORCION DE LA MUESTRA

$$\sigma = \sqrt{\frac{P\,(1-P)}{n}}$$

- Remplanzando

$$P \pm Z \sqrt{\frac{P\,(1-P)}{n}}$$

DONDE:

P = Proporción

Z = Intervalo de confianza seleccionado

σ p = Error estándar de la proporción

n = Tamaño de la muestra

Aplicaciones:

De 1600 a 2000 trabajadores sindicalizados que se muestrean dijeron que plantean poner a votación una propuesta para unirse a una federación. Si se utiliza un nivel de confianza de 0.95.

a) ¿Cuál es la estimación del intervalo para la proporción poblacional?

b) ¿A qué conclusión se llegaría con base en intervalo de confianza?

UTILIZANDO FÓRMULA:

$$P \pm Z \sqrt{\frac{P\,(1-P)}{n}}$$

2000 -------- 100%

1600 -------- X

$$X = \frac{(1600)(100)}{2000}$$

X = 80%

P= 80 / 100

P = 0.80

$$P \pm Z \sqrt{\frac{P\,(1-P)}{n}}$$

$$0.80 \pm 1.96 \sqrt{\frac{0.80\,(1-0.80)}{2000}}$$

$$0.80 \pm 1.96 \sqrt{\frac{0.80\,(0.20)}{2000}}$$

$$0.80 \pm 1.96 \sqrt{\frac{0.16}{2000}}$$

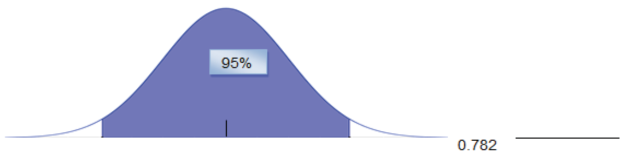
$$0.80 \pm 1.96\,(0.0089)$$

$$0.80 \pm 0.0174$$

DONDE:

0.80 + 0.0174 = 0.8174 LÍMITE SUPERIOR

0.80 – 0.0174 = 0.7826 LÍMITE INFERIOR



Conclusión:

En base a la proporción con un nivel de confianza del 95% y con un límite superior del 0.8174 y un límite inferior del 0.7826.

Aplicaciones

Se realizo una investigación de mercado para estimar la proporción de amas de casa que pueden reconocer la marca de clase de un limpiador con base de la forma y el color del recipiente de las 1400 personas, 420 pudieron identificar la marca del producto. Por lo que se necesita saber:

a) Si se utiliza el grado de confianza en que el intervalo se encuentra la proporción de la población.

b) Cuáles son los límites de Confianza

c) Conclusiones e interpretación de Resultados

a. 1400-100%

420 - x

$= \frac{(420)(100)}{1400}$

=30/100

= 0.30

$P \pm Z\sqrt{\frac{P(1-P)}{n}}$

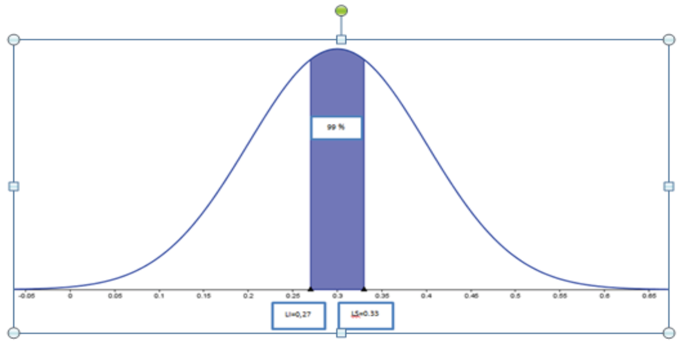
$0.30 \pm 2.58\sqrt{\frac{0.30(1-0.30)}{1400}}$

$0.30 \pm 2.58\sqrt{0.012247}$

0.30 ± 0.03

0.30 + 0.03= 0.33

0.30- 0.03= 0.27



En % De27% - 33%

420 - 100%

X - 27%

$X = \frac{(420)(27)}{100}$

420 - 100%

X - 33%

$X = \frac{(420)(33)}{100}$

Rango

X = 113.40 ⟷ X= 138.60

Conclusión :

Se determina que el grado de confianza es de 0.30% de proporción sobre la población.

Esto refleja un rango de 113.40 – 138.60.

HIPÓTESIS ESTADÍSTICA

"Enunciado acerca de una población con el propósito de Poner a Prueba"

Ejemplo

SISTEMA JURÍDICO: Una persona es inocente hasta que se pruebe que es culpable.

Un jurado tiene la hipótesis de que una persona es inocente y se somete esta hipótesis a una verificación de pruebas antes de llegar a un veredicto.

PASOS PARA PROCEDIMIENTOS PARA COMPROBACION DE UNA HIPOTESIS

1. Plantear la Hipótesis
   - Nula "Ho"
   - Alternativa "H"
2. Seleccionar el Nivel de Significación
3. Identificar el Valor Estadístico de la Prueba.
4. Formular una Regla de Decisión
5. Tomar una Muestra y llegar una Decisión.
   - Acepta Ho
   - Rechaza Ho

CHI CUADRADO

Un sociólogo desea saber si es posible concluir que hay un grado de liberalismo y la posición de la Universidad en una población de estudiantes Universitarios. El sociólogo selecciono una muestra de 500 estudiantes.

	GRADO DE LIBERALISMO			Sub.
	Ligero	Moderado	Alto	Total
Primer Año	30	63	37	130
Segundo Año	19	56	50	125
Tercer Año	16	46	63	125
Cuarto Año	10	38	72	120
Subtotal	75	205	222	500

1. PASO Formulación de la Hipótesis:

Ho: El grado de liberalismo y la clase "No" son independientes en la Población de estudiantes.

Hi: El grado de liberalismo y la clase "Si" son independientes en la Población de los estudiantes.

2. PASO Modelo Matemático

$$X^2 = \sum \frac{(O-E)}{E}$$

3. PASO Elección Prueba Estadística

• Chip Cuadrado

4. PASO Nivel de Significación

5% / 100= 0.05

5. PASO Distribución Muestral

GL= K-1 → "K" → FILAS / COLUMNAS

GL= (F-1) (C-1)

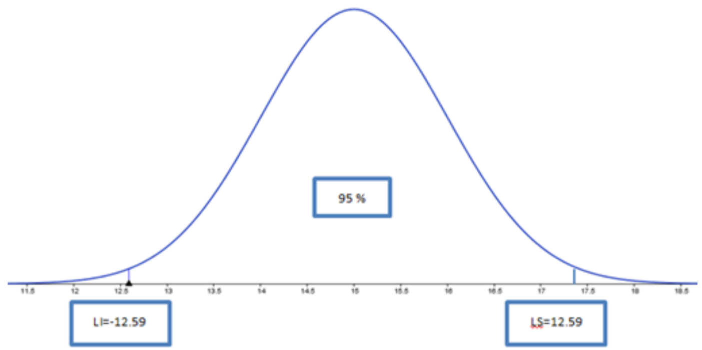
= (4-1) (3-1)

= (3) (2)

GL= 6 → Ver tabla

Gl= **12.592**

6. PASO Definición Zona de Rechazo



7. PASO

O	E	(O-E)	$(O-E)^2$	$\frac{(O-E)^2}{E}$
30	19,50	10,50	110,25	5,65
63	52,78	10,22	104,45	1,98
37	57,72	-20,72	429,32	7,44
19	18,75	0,25	0,06	0,003
56	50,75	5,25	27,56	0,54
50	55,00	-5,50	30,25	0,55
16	18,75	-2,75	7,56	-0,40
46	50,75	-4,75	22,56	0,45
63	55,50	7,50	56,25	1,01
10	18,00	-8,00	64,00	8,00
38	48,72	-10,72	114,92	2,36
72	53,28	18,72	350,42	6,58
500	**500**	**0**	**1317,62**	**30,51**

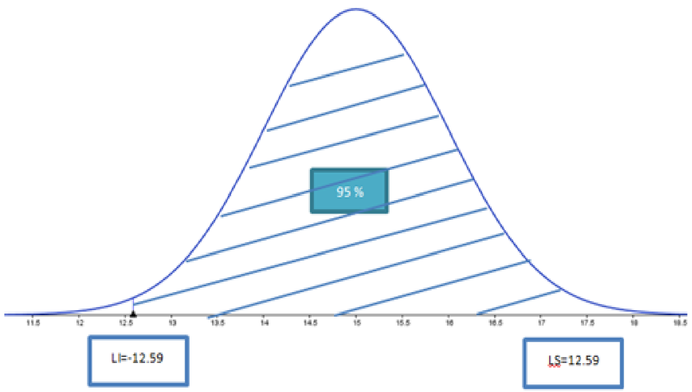
CALCULOS DE ESPERADOS "E"

$\frac{75*130}{500}= 19.5$ $\frac{75*125}{500}= 1$

$\frac{200*130}{500}= 52.58$ $\frac{203*125}{500}=$

$\frac{222*130}{500}= 57.72$

$\frac{75*120}{500}= 18$



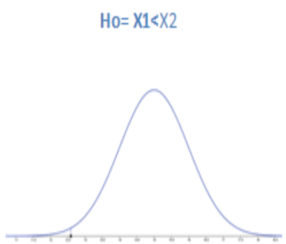
Decisión

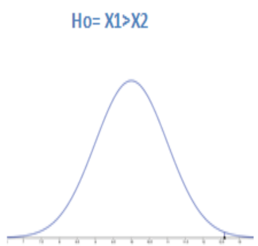
Como XF = 12.59 <Xc= 30.51 Entonces

Aceptamos Hipótesis Alternativa “H” que dice

“El grado de liberalismo y la clase “S” son independientes en la población de los Estudiantes.

Ho= X1-X2





Taller en clases

Ejemplo:

Se lanza una moneda al aire, ¿Cuál es la probabilidad que se obtenga cara?

Total, de evento (cara, sello)

$$p = \frac{1}{2}$$

$$p = 0{,}5 \text{ ó } 50\%$$

q= 1- P

q= 1 – 0.5

q= 0.5

P + q = 1

0.5 + 0.5= 1

1 = 1

Se lanza un dado al aire. ¿Cuál es la probabilidad de que se obtenga un numero 5?

Total de eventos

$$P = \frac{1}{6} = 0.17$$

Sea E el suceso de lanzar un dado una vez salga un 3 o un 4.

¿Cuál es dicha probabilidad? y ¿Cuál es la probabilidad de que no salga 3 o 4?

$$P = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$$

$$Q = 1 - \frac{1}{3}$$

$$Q = \frac{2}{3}$$

Se lanza 2 monedas al aire, ¿Cuál es la probabilidad de que ambas salgan cara?

Total eventos

$$P = \frac{1}{4} \qquad \begin{bmatrix} S & C \\ C & S \\ C & C \\ S & C \end{bmatrix}$$

Actividades de aprendizajes 1.-

1.-Oreganizar un concepto de variable aleatoria:

-Es una probabilidad, es una función que asigna valores numéricos, valores establecidos a los puntos muestrales, a los puntos de una competencia, de un experimento, de un diagrama

de Venn, de un número de veces.

- Es una función que asigna valores numéricos a los puntos muestrales de un experimento.

2.-Probabilidad Clásica

a) Son puntos de objetivo, numero de resultados favorables, número de eventos, numero de ocurrencia, numero de implicación entre: ninguno de los eventos, numero de resultados posibles, numero de posibles empleados, probabilidad específica, número de veces en un conjunto de eventos.

$$\frac{\text{Número de resultados favorables}}{\text{Número de resultados posibles}}$$

3.-Probabilidad Empírica

-Número de veces repetitiva, número de veces que ocurrió un evento pasado, número de veces de evento específico, número de veces en total de unidades sumadas entre:

Número total de funciones, número total de lanzamientos, número total de observaciones.

$$\frac{\text{Número de veces que ocurrio el evento en el pasado}}{\text{Número total de observaciones}}$$

3.-Construya con:

$$P_{(AyB)}; P_A; P_{(AoB)}; P_{(B)}$$
$$-P_{(\sim A)}; P_B^A; P_{(A)}; P_A y P_B y P_C; (P)$$

-Regla de adición

$$P_{(AoB)} = P_{(AoB)} = P_{(A)} + P_{(B)}$$

-Regla de complemento

$$P_{(A)=} 1 - P_{(\sim A)}$$

-Regla especial de la suma

$$P_{(AoB)} = P_{(A)} + P_{(B)} - P_A^B$$

4.-Una maquina llena sacos de papas, la mayor parte de los sacos contienen peso correcto, pero debido a ligeras variaciones en el tamaño de las papas, un saco puede tener ligeramente mayor o menor. Una verificación de 8000 sacos que se llenaron en el mes de abril revelo lo siguiente:

PESO	EVENTO	N° DE PAQUETES	PROBABILIDAD DE OCURRENCIA
Mayor	A	600	0.075
Satisfactorio	B	7200	0.90
Menor	C	200	0.025
TOTAL		8000	1

a) Probabilidad de que el peso del saco no sea el correcto es decir sea mayor o menor.
b) Probabilidad de que el saco sea satisfactorio (ni menor ni mayor).
c) Utilice el diagrama de Venn.

$$P(A) = \frac{600}{8000}$$
$$P(A) = 0{,}075$$

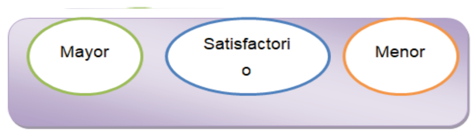
Mayor

$$P(A) = \frac{200}{8000}$$
$$P(A) = 0{,}025$$

Menor

$$P(A) = \frac{7200}{8000}$$
$$P(A) = 0{,}9$$

Satisfactorio



Actividades de aprendizaje 1.-

15. Los eventos A Y B son mutuamente excluyentes. Supóngase que $P(A)=0.30$ y $P(B)=0.20$. ¿Cuál es la probabilidad de que ocurra A o B? ¿Cuál es la probabilidad de que no suceda ni A ni B?

$$P(A \text{ o } B)=PA+PB$$

P (Ao B)=0.30+0.20

P (A o B)=0.50

Conclusión :

Puede hacerse A o puede hacerse B

P (A o B)=1-(PA+P B)

P (Ao B)=1-(0.05+0.02)

P (A o B)=1-0.50

P (A o B)=0.50

Conclusión:

Puede ocurrir B o puede ocurrir B

18. Un estudio de opiniones de diseñadores en lo referente al color primario más conveniente para aplicar en oficinas ejecutivas indico:

a) ¿Cuál es el experimento?

b) ¿Cuál es el evento posible?

c) ¿Cuál es la probabilidad de seleccionar una respuesta específica y descubrir que el diseñador prefiera rojo o verde?

d) ¿Cuál es la probabilidad de que un diseñador no prefiera el amarillo?

a) Estudio de opiniones de diseñadores para la aplicación de colores primarios

b) Cualquier color primario

n= 5

π=0.20

x= 1

$$P(x) = nCx\pi^{x}(1-\pi)^{n-x}$$

c)

$$P\ (\ A\ o\ B\ ) = PB + PC$$

$$P\ (A\ o\ B) = \frac{92}{100} + \frac{91}{400}$$

$$P\ (A\ o\ B) = 0.23 + 0.23$$

$$P\ (A\ o\ B) = 0.46$$

a) No hay probabilidad de que el diseñador prefiera rojo o verde

$$P\ (\ C\ ) = PC$$

$$P\ (A) = \frac{46}{400}$$

$$P\ (A) = 0.12$$

b) La probabilidad que no escoja el amarillo

28. Un estudio realizado por el servicio de Parque Nacionales (de Estados Unidos) reveló que 50% de los vacacionistas que viajan a la región de las Montañas Rocosas van al Parque Yellowstone, 40% visitan Tetons, y 35% van a ambos sitios.

a) ¿Cuál es la probabilidad de que un vacacionista visite al menos una de estas atracciones?

b) ¿Cómo se denomina a la probabilidad 0.35?

c) ¿Los eventos son mutuamente excluyentes? Explique su respuesta

P (X o Y)=PX+PY-P(X o Y)

P (X o Y)=0.50+0.40-0.35

P (X o Y)=0.55

Si hay probabilidad de que un vacacionista visite al menos de una de estas atracciones

a) Probabilidad conjunta

Aun departamento de mercadotecnia se le ha solicitado que enseñe códigos de colores para 42 líneas de discos compactos "CD" que comercializa la empresa GOOD STUDENT. Se van a utilizar 3 colores en cada línea de "CD", pero una combinación de 3 colores que se utiliza en una línea no puede reordenarse y utilizarse para identificar a otra línea diferente. Esto significa que si se usara los colores verde, amarillo y violeta para señalar una línea, entonces amarillo, verde y violeta (o cualquier otra combinación de estos 3 colores) no se podría emplear para identificar otra línea.

¿Serán adecuados 7 colores tomados 3 la vez para codificar

adecuadamente las 42 líneas?

APLICACANDO FÓRMULA DE CONBINACIÓN

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

7C3= 7!/3!(7-3)!

7C3= 7.6.5.4.3.2.1/3.2.1(4)!

7C3= 7.6.5.4.3.2.1/3.2.1(4.3.2.1)

7C3= 7.6.5/3.2.1

7C3= 35

Conclusión:

Utilizando 7 colores con 3 a la vez no satisfacen la necesidad de codificar.

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

8C3= 8!/3!(8-3)!

8C3= 8.7.6.5.4.3.2.1/3.2.1!(5)!

8C3= 8.7.6.5.4.3.2.1/3.2.1(5.4.3.2.1)

8C3= 8.7.6/3.2.1

8C3=56

Conclusión:

Utilizando 8 colores tomados de 3 en 3 darían 56 combinaciones distintas. Esto sería más que suficiente para codificar cromáticamente las 42 líneas.

Actividades de aprendizaje 3.-

Una encuesta de corretaje financiero "Crédito Fácil" reportan que 30% de los inversionistas individuales han empleado a un corredor de descuento; esto es, uno que no cobra las comisiones completas. Es una muestra seleccionada al azar de nueve inversionistas, ¿Cuál es la probabilidad de que:

a) Exactamente dos de los individuos de la muestra hayan empleado a un corredor de descuento?

b) Exactamente cuatro de ellos han recurrido a un corredor de ese tipo?

c) Ninguno haya recurrido a un corredor de descuento?

a)

$$2.-\ P(x) = {}_{n}C_{x}\ \pi^{x}(1-\pi)^{n-x} \qquad {}_{n}C_{r} = \frac{n!}{r!(n-r)!}$$

$$P(2) = {}_{9}C_{2}\ (0{,}30)^{2}(1-0{,}30)^{9-2} \qquad {}_{9}C_{2} = \frac{9*8*7*6*5*4*3*2*1}{2*1(9-2)}$$

$$P(2) = 36\ (0{,}30)^{3}(0{,}70)^{7} \qquad {}_{9}C_{2} = \frac{9*8*7*6*5*4*3*2*1}{2*1(7*6*5*4*3*2*1)}$$

$$P(2) = 0{,}27 \qquad {}_{9}C_{2} = 36$$

Conclusión:

La probabilidad de que exactamente dos individuos hayan empleado a un corredor de descuento es nula con un 27%.

b)

$$4.-\ P(x) = {}_{n}C_{x}\ \pi^{x}(1-\pi)^{n-x} \qquad {}_{n}C_{r} = \frac{n!}{r!(n-r)!}$$

$$P(4) = {}_{9}C_{4}\ (0{,}30)^{4}(\ {}_{9}C_{4} = \frac{9*8*7*6*5*4*3*2*1}{4*3*2*1(9-4)}$$

$$P(4) = 126\ (0{,}30)^{4}(0{,}70)^{5} \qquad {}_{9}C_{4} = \frac{9*8*7*6*5*4*3*2*1}{4*3*2*1(5*4*3*2*1)}$$

$$P(4) = 0{,}17 \qquad {}_{9}C_{4} = 1261-0{,}30)^{9-4}$$

Conclusión:

La probabilidad de que exactamente cuatro individuos hayan empleado un corredor de descuento es nula con un 17%

c)

$$4.-\ P(x) = {}_{n}C_{x}\ \pi^{x}(1-\pi)^{n-x} \qquad {}_{n}C_{r} = \frac{n!}{r!(n-r)!}$$

$$P(0) = {}_{9}C_{0}\ (0{,}30)^{0}(1-0{,}30)^{9-0} \qquad {}_{9}C_{0} = \frac{9*8*7*6*5*4*3*2*1}{0(9-0)}$$

$$P(0) = 1\ (0{,}30)^{0}(0{,}70)^{9} \qquad {}_{9}C_{0} = \frac{9*8*7*6*5*4*3*2*1}{0(9*8*7*6*5*4*3*2*1)}$$

$$P(0) = 0{,}040 \qquad {}_{9}C_{0} = 1$$

Conclusión:

La probabilidad de que ninguno haya recorrido a un corredor de descuento es nula con un 4%.

Actividades de aprendizaje 4.-

2.9. Distribucción de probabilidad binomial

a.- Es un ejemplo de datos excluyentes

b.- Es una reunión de Datos

c.- Es una Distribución de probabilidad discreta

Completar: Es una distribución de probabilidades discretas, los resultados son mutuamente excluyentes, lo cual implica que una respuesta a una pregunta puede ser verdadero o falso.

2.10. La fórmula Binomial esta compuesta por:

P(X): Probabilidad evento a realizarse

C: Combinación

N: Número de ensayos

Π: Probabilidad de éxito en cada ensayo

X: El número de éxitos

2.- Entre Quito y Guayaquil hay 5 vuelos diarios hay la probabilidad de que un vuelo llegue retrasado es de 0,20 ¿Cuál es la probabilidad de que ningún vuelo se retrase el día de hoy?

N° DE AUTOS VENDIDOS	PROBABILIDAD Px	xP(x)	x-μ	$(x-\mu)^2$	$(x-\mu)^2P(x)$
0	0,3277	0,3277	0,6723	0,45198729	0,148116235
1	0,4096	0,4096	0,5904	0,34857216	0,142775157
2	0,2048	0,4096	0,7952	0,63234304	0,129503855
3	0,0512	0,1536	0,9488	0,90022144	0,046091338
4	0,0064	0,0256	0,9936	0,98724096	0,006318342
5	0,003	0,015	0,997	0,994009	0,002982027
TOTAL	1,0	1,3			0,5

3.- Jorge está encargada de los préstamos en el Banco Pichincha, estima la probabilidad de que un solicitante no sea capaz de pagar su préstamo es de 0.25. El mes pasado realizo 40 préstamos.

a. ¿Cuál es la probabilidad de que 3 préstamos no sean pagados a tiempo?

b. ¿Cuál es la probabilidad de que por lo menos 3 préstamos no se liquiden a tiempo?

DATOS:

n=40

x=3

μ=0.025*40

μ=10

a: $$P(X) = \frac{\mu^X e^{-\mu}}{X!}$$

$$P(3) = \frac{(10)^3(2.71828)^{-10}}{3!}$$

$$P(3) = \frac{1000\ (0.000045400)}{3*2*1}$$

$$P(3) = 0.0075667 \quad = 0.76\%$$

Conclusión:

No hay la probabilidad de que los préstamos sean pagados a tiempo.

b: $$P(X) = \frac{\mu^X e^{-\mu}}{X!}$$

$$P(3) = \frac{(10)^3(2.71828)^{-10}}{3!}$$

$$P(3) = \frac{1000\ (0.000045400)}{3*2*1}$$

$$P(3) = 0.0075667 \quad = 0.76\%$$

Conclusión:
No hay la probabilidad de que los préstamos sean pagados a tiempo.

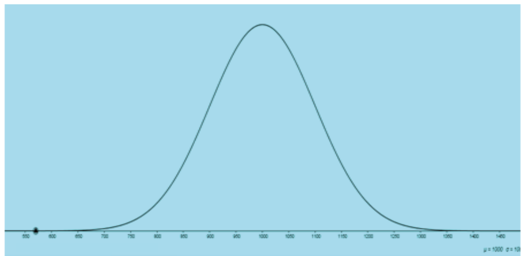
4.- Cual es la Probabilidad de que un estudiante de la Carrera de contabilidad y Auditoría del quinto semestre, es escogido al azar entre 420-570 hora para completar el programa de preparación de tesis de grado con una media 1000 y una distribución estándar de 100

Datos:

U=100

= 100

Intervalo = 420 -570



0,0942

$$z = \frac{x - \mu}{\sigma}$$

$$z = \frac{420 - 1000}{100}$$

$$z = \frac{-580}{100}$$

$$z = -5{,}8 + 3.0 = 0. = -2{,}8$$

$$z = 0{,}4974$$

$$z = \frac{x - \mu}{\sigma}$$

$$z = \frac{570 - 1000}{100}$$

$$z = \frac{-430}{100}$$

$$z = -4{,}30 + 3.0 = 0. = -1{,}3$$

$$z = 0{,}4032$$

Entonces Restamos:

0.4974

-0.4032

0.0942

9.42%

Conclusión:

La probabilidad de escoger, a un estudiante de la carrera de contabilidad y Auditoría es del 9,42%

COMPROBACIÓN DE HIPÓTESIS CON EXCEL

ESTADÍSTICO Z

Introducción. -

El estadístico Z es una medida utilizada en estadística para estandarizar y comparar valores en una distribución normal, ya que permite estandarizar y comparar valores en una distribución normal, representando cuántas desviaciones estándar un valor observado se encuentra por encima o por debajo de la media de la población. (Barbara illowsky, 2022)

Pasos para seguir para calcular el Estadístico

Ejemplo

El fabricante de neumáticos radiales con cinturón de acero X-15 para camiones señala que el millaje medio que cada uno recorre antes de que se desgasten las cuerdas es de 60 000 millas; la desviación estándar del millaje es de 5 000 millas. La Crosset Truck Company compró 48 neumáticos y comprobó que el millaje medio para sus camiones es de 59 500. ¿La experiencia de Crosset es diferente de lo que afirma el fabricante en el nivel de significancia 0.057

Se establecen las hipótesis nula y alternativa

Hipótesis nula: Ho

Hipótesis alternativa: H1

Se selecciona un nivel de significancia.

Error tipo I: Rechazar Ho cuando es verdadera α

Error tipo II: Aceptar Ho cuando es falsa

Se identifica el estadístico de la prueba, en el siguiente ejercicio

se utilizará el estadístico de prueba z, con la siguiente fórmula:

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Donde:

x̄: media muestral

μ: media poblacional

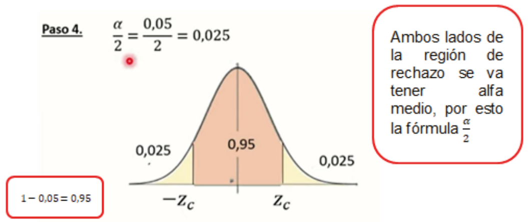
σ: desviación estándar de la población

n: tamaño de la muestra

De acuerdo con cada ejercicio, si se conoce sigma σ y μ como parámetro, no importa si el tamaño de muestra es pequeño, se utiliza el estadístico de prueba z

Se reemplaza los datos del ejercicio

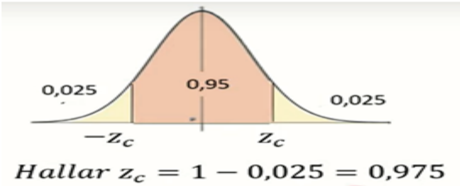
Se fórmula una regla para tomar decisiones Enunciado donde se muestra cuando se acepta o rechaza Ho:



LINK

https://www.youtube.com/watch?v=RS5F_bhNugw&feature=youtu.be

Hallar Zc, para ello se utiliza la tabla de distribución normal acumulada, la que muestra la probabilidad de estar por debajo de un valor



LINK

https://www.youtube.com/watch?v=RS5F_bhNugw&feature=youtu.be

Continuamente al tener el valor de Zc, se busca en la tabla de distribución

z	0,00	0,05	0,06	0,07	0,08	0,09
1,7	0,9554	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9798	0,9803	0,9808	0,9812	0,9817

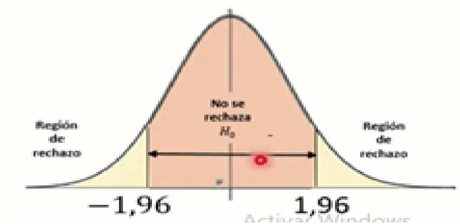
Para conocer Zc se suma el valor de la fila y la columna del valor donde está ubicado:

Zc=1,9+0.06

Zc=1,96

Regla de decisión:

Se determina si se acepta o rechaza la hipótesis nula, si el valor de z calculado se encuentra dentro de Zc (es decir, si está dentro de 1,96 y -1,96)



LINK

[https://www.youtube.com/watch?v=RS5F_bhNugw&feature=-youtu.be](https://www.youtube.com/watch?v=RS5F_bhNugw&feature=youtu.be)

Se toma una decisión

Se calcula el valor Z del paso 3, reemplazando los valores

Paso 5.

$$z = \frac{59500 - 60000}{\frac{5000}{\sqrt{48}}}$$

$$z = -0{,}69$$

LINK

[https://www.youtube.com/watch?v=RS5F_bhNugw&feature=-youtu.be](https://www.youtube.com/watch?v=RS5F_bhNugw&feature=youtu.be)

Se realiza la conclusión

Como el valor de z se encuentra dentro de la región de no rechazo, se concluye mencionado que la media de la población no es distinta de 60000.

Coclusiones. -

El estadístico Z permite comparar un valor observado con una distribución normal estándar, lo que proporciona información sobre la posición relativa del valor en relación con la media y la desviación estándar de la población. Si el valor Z es positivo, significa que el valor observado está por encima de la media, mientras que un valor Z negativo indica que está por debajo de la media.

Utilizando el estadístico Z, se pueden calcular intervalos de confianza para estimar un parámetro poblacional. Estos intervalos proporcionan información sobre la variabilidad esperada del valor de interés y permiten establecer límites en los que se encuentra con un determinado nivel de confianza.

T STUDENT

Introducción. -

La prueba t de Student, también conocido como prueba t, es una herramienta estadística que se utiliza para comparar las medias de dos grupos y determinar si existen diferencias significativas entre ellas. (López, 2019)

Aquí te presento una explicación paso a paso sobre cómo

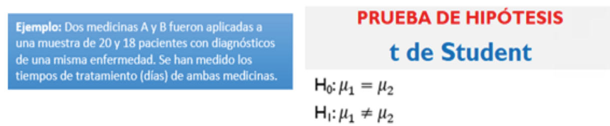
realizar la prueba t de Student para dos muestras independientes. Supongamos que tienes dos grupos de datos medicinas A y B fueron aplicadas a una muestra de 20 y 18 pacientes con diagnósticos de una misma enfermedad. Se han medido los tiempos de tratamiento (días) de ambas medicinas.

Pasos por seguir para calcular T Student

Formulación de hipótesis nula y alternativa

- Hipótesis nula (H0): Es la suposición de que no hay diferencias significativas entre las medias de los dos grupos. En términos estadísticos, H0 plantea que la diferencia entre las medias es igual a cero.

- Hipótesis alternativa (H1): Es la afirmación opuesta a la hipótesis nula. Plantea que hay una diferencia significativa entre las medias de los dos grupos.



Reúne los datos

Obtén las muestras de ambos grupos (Grupo A y Grupo B) y registra los valores de cada muestra. Asegúrate de que ambas muestras sean independientes entre sí y que cumplan con los supuestos necesarios para realizar la prueba t de Student.

Ejemplo: Dos medicinas A y B fueron aplicadas a una muestra de 20 y 18 pacientes con diagnósticos de una misma enfermedad. Se han medido los tiempos de tratamiento (días) de ambas medicinas.

t de Student

$H_0: \mu_1 = \mu_2$
$H_1: \mu_1 \neq \mu_2$

Nro	Medicina A	Medicina B
1	6	7
2	5	6
3	6	7
4	7	9
5	5	5
6	7	8
7	6	7
8	4	6
9	3	7
10	6	9
11	6	8
12	5	7
13	7	8
14	8	7
15	6	6
16	5	8
17	8	7
18	4	6
19	6	
20	6	

Calcula la media y la desviación estándar de cada muestra - Calcula la media (promedio) de cada muestra sumando todos los valores y dividiendo entre el número total de elementos en cada muestra, con la siguiente fórmula

= SUMA (NUMERO DE DATOS A) /TOTAL DE DATOS A Y B

= SUMA (NUMERO DE DATOS B) /TOTAL DE DATOS A Y B

Ejemplo: Dos medicinas A y B fueron aplicadas a una muestra de 20 y 18 pacientes con diagnósticos de una misma enfermedad. Se han medido los tiempos de tratamiento (días) de ambas medicinas.

PRUEBA DE HIPÓTESIS
t de Student

$H_0: \mu_1 = \mu_2$
$H_1: \mu_1 \neq \mu_2$

Nro	Medicina A	Medicina B
1	6	7
2	5	6
3	6	7
4	7	9
5	5	5
6	7	8
7	6	7
8	4	6
9	3	7
10	6	9
11	6	8
12	5	7
13	7	8
14	8	7
15	6	6
16	5	8
17	8	7
18	4	6
19	6	
20	6	
	5,8	7,11

- Calcula la desviación estándar de cada muestra, que mide la dispersión de los datos alrededor de la media, con la siguiente fórmula.

= PROMEDIO (NUMERO DE DATOS EN A)

= PROMEDIO (NUMERO DE DATOS EN B)

Ejemplo: Dos medicinas A y B fueron aplicadas a una muestra de 20 y 18 pacientes con diagnósticos de una misma enfermedad. Se han medido los tiempos de tratamiento (días) de ambas medicinas.

PRUEBA DE HIPÓTESIS
t de Student

$H_0: \mu_1 = \mu_2$
$H_1: \mu_1 \neq \mu_2$

Estadístico de prueba

$\bar{X}_1$ = 5,80
$\bar{X}_2$ = 7,11

Nro	Medicina A	Medicina B
1	6	7
2	5	6
3	6	7
4	7	9
5	5	5
6	7	8
7	6	7
8	4	6
9	3	7
10	6	9
11	6	8
12	5	7
13	7	8
14	8	7
15	6	6
16	5	8
17	8	7
18	4	6
19	6	
20	6	
	5,8	7,11

Calcula la diferencia de medias y la desviación estándar combinada

Donde:

n1 y n2 son los tamaños de las muestras de los grupos A y B, lo realizamos con la siguiente fórmula.

= CONTAR (NUMERO DE DATOS A)

= CONTAR (NUMERO DE DATOS B)

Ejemplo: Dos medicinas A y B fueron aplicadas a una muestra de 20 y 18 pacientes con diagnósticos de una misma enfermedad. Se han medido los tiempos de tratamiento (días) de ambas medicinas.

PRUEBA DE HIPÓTESIS
t de Student < 30

$H_0: \mu_1 = \mu_2$
$H_1: \mu_1 \neq \mu_2$

Estadístico de prueba

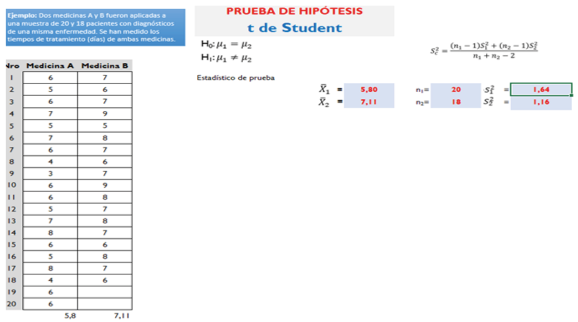
$\bar{X}_1$ = 5,80 n_1= 20
$\bar{X}_2$ = 7,11 n_2= 18

Nro	Medicina A	Medicina B
1	6	7
2	5	6
3	6	7
4	7	9
5	5	5
6	7	8
7	6	7
8	4	6
9	3	7
10	6	9
11	6	8
12	5	7
13	7	8
14	8	7
15	6	6
16	5	8
17	8	7
18	4	6
19	6	
20	6	
	5,8	7,11

s1 y s2 son las desviaciones estándar de las muestras de los grupos A y B, Lo encontramos con la siguiente fórmula.

= VAR.S(NUMERO DE DATOS DE A)

= VAR.S(NUMERO DE DATOS DE B)



Ejemplo: Dos medicinas A y B fueron aplicadas a una muestra de 20 y 18 pacientes con diagnósticos de una misma enfermedad. Se han medido los tiempos de tratamiento (días) de ambas medicinas.

PRUEBA DE HIPÓTESIS
t de Student

$H_0: \mu_1 = \mu_2$
$H_1: \mu_1 \neq \mu_2$

Estadístico de prueba

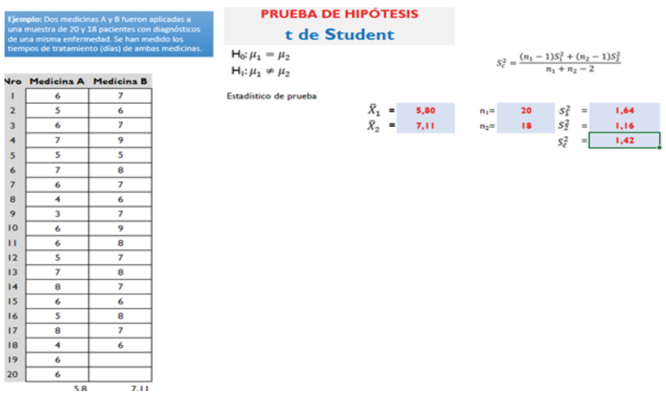
$S_c^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$

$\bar{X}_1$ = 5,80 n_1= 20 S_1^2 = 1,64
$\bar{X}_2$ = 7,11 n_2= 18 S_2^2 = 1,16

Nro	Medicina A	Medicina B
1	6	7
2	5	6
3	6	7
4	7	9
5	5	5
6	7	8
7	6	7
8	4	6
9	3	7
10	6	9
11	6	8
12	5	7
13	7	8
14	8	7
15	6	6
16	5	8
17	8	7
18	4	6
19	6	
20	6	
	5,8	7,11

- Calcula la desviación estándar combinada (Sc) utilizando la siguiente fórmula:

$$s_c^2 = \frac{(n_1-1)s_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$$



Ejemplo: Dos medicinas A y B fueron aplicadas a una muestra de 20 y 18 pacientes con diagnósticos de una misma enfermedad. Se han medido los tiempos de tratamiento (días) de ambas medicinas.

PRUEBA DE HIPÓTESIS
t de Student

$H_0: \mu_1 = \mu_2$
$H_1: \mu_1 \neq \mu_2$

Estadístico de prueba

$S_c^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$

$\bar{X}_1$ = 5,80 n_1= 20 S_1^2 = 1,64
$\bar{X}_2$ = 7,11 n_2= 18 S_2^2 = 1,16
S_c^2 = 1,42

Nro	Medicina A	Medicina B
1	6	7
2	5	6
3	6	7
4	7	9
5	5	5
6	7	8
7	6	7
8	4	6
9	3	7
10	6	9
11	6	8
12	5	7
13	7	8
14	8	7
15	6	6
16	5	8
17	8	7
18	4	6
19	6	
20	6	
	5,8	7,11

Cálculo del estadístico t

- Utiliza la siguiente fórmula para calcular el estadístico t:

$$t = \frac{\overline{x_1} - \bar{x}_2}{\sqrt{\frac{s_c^2}{n_1} + \frac{s_c^2}{n_2}}}$$

Ejemplo: Dos medicinas A y B fueron aplicadas a una muestra de 20 y 18 pacientes con diagnósticos de una misma enfermedad. Se han medido los tiempos de tratamiento (días) de ambas medicinas.

PRUEBA DE HIPÓTESIS
t de Student

$H_0: \mu_1 = \mu_2$
$H_1: \mu_1 \neq \mu_2$

Estadístico de prueba

$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_c^2}{n_1} + \frac{S_c^2}{n_2}}}$

$S_c^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$

$\bar{X}_1$ = 5,80 n_1= 20 S_1^2 = 1,64
$\bar{X}_2$ = 7,11 n_2= 18 S_2^2 = 1,16
S_c^2 = 1,42

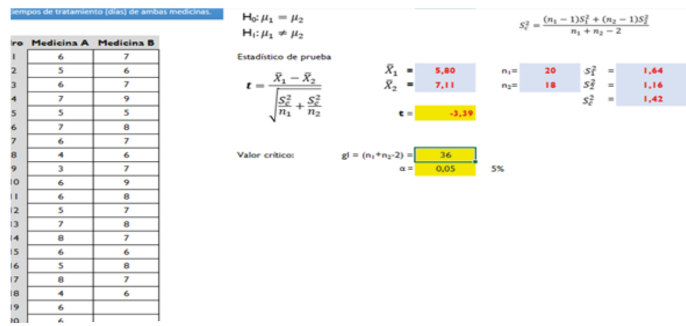
t = -3,39

Nro	Medicina A	Medicina B
1	6	7
2	5	6
3	6	7
4	7	9
5	5	5
6	7	8
7	6	7
8	4	6
9	3	7
10	6	9
11	6	8
12	5	7
13	7	8
14	8	7
15	6	6
16	5	8
17	8	7
18	4	6
19	6	
20	6	
	5,8	7,11

Determina los grados de libertad

- Los grados de libertad (gl) para este tipo de prueba se calculan utilizando la siguiente fórmula:

$$gl = (n_1 + n_2 - 2)$$



tiempos de tratamiento (días) de ambas medicinas.

$H_0: \mu_1 = \mu_2$
$H_1: \mu_1 \neq \mu_2$

Estadístico de prueba

$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_c^2}{n_1} + \frac{S_c^2}{n_2}}}$

$S_c^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$

$\bar{X}_1$ = 5,80 n_1= 20 S_1^2 = 1,64
$\bar{X}_2$ = 7,11 n_2= 18 S_2^2 = 1,16
S_c^2 = 1,42

t = -3,39

Valor crítico: gl = (n_1+n_2-2) = 36
α = 0,05 5%

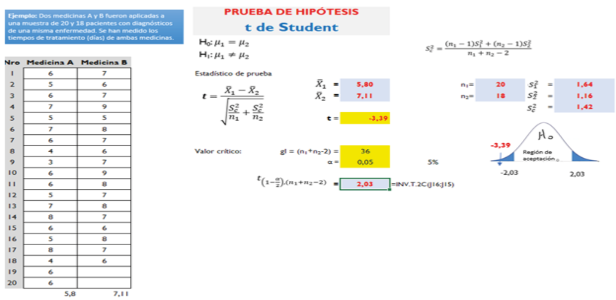
Nro	Medicina A	Medicina B
1	6	7
2	5	6
3	6	7
4	7	9
5	5	5
6	7	8
7	6	7
8	4	6
9	3	7
10	6	9
11	6	8
12	5	7
13	7	8
14	8	7
15	6	6
16	5	8
17	8	7
18	4	6
19	6	
20	6	

Busca el valor crítico o puntuación t crítica

- Utiliza el nivel de significancia deseado y los grados de libertad para encontrar el valor crítico de t en una tabla de distribución t de Student o utilizando software estadístico, mediante la fórmula.

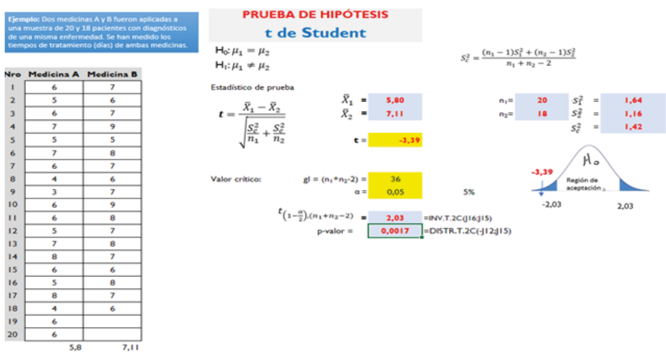
$$t_{\left(1-\frac{\alpha}{2}\right),(n_1+n_2-2)}$$

= INV.T.2C (∝; VALOR CRÍTICO)



Calculamos el p–valor

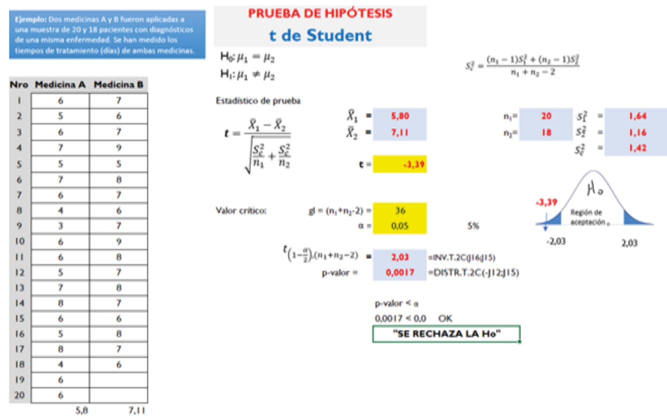
= =DISTR.T.2C(-t ; gl)



Compara el valor calculado de t con el valor crítico

- Si el valor calculado de t es mayor que el valor crítico de t, entonces rechazamos la hipótesis nula (H0) y aceptamos la hipótesis alternativa (H1), lo que indica que hay una diferencia significativa entre las medias de los grupos A y B.

- Si el valor calculado de t es menor o igual que el valor crítico de t, no tenemos evidencia suficiente para rechazar la hipótesis nula y concluimos que no hay una diferencia significativa entre las medias de los grupos A y B.



Conclusiones. -

Las conclusiones obtenidas a partir de un t-test (t de Student) dependerán de los resultados y del análisis realizado.

En general, las conclusiones del t-test deben ser presentadas y comunicadas de manera clara y objetiva, junto con cualquier consideración sobre las limitaciones y el contexto específico del estudio. El t-test es una herramienta poderosa para el análisis de datos y la toma de decisiones, pero su correcta interpretación y aplicación son fundamentales para obtener resultados válidos y significativos, también podemos decir que el t- student es una técnica estadística funcional y de gran utilidad para comparar las medias de dos grupos y detectar diferencias significativas entre ellos. Su aplicabilidad en muestras pequeñas y su amplio uso en diversos campos lo hacen una herramienta confiable y efectiva en el análisis de datos. Sin embargo, es esencial utilizarlo correctamente y comprender sus limitaciones para obtener conclusiones válidas y relevantes

CHI CUADRADO

Introducción. -

La prueba de chi cuadrado es un método utilizado en estadística para determinar si existe una asociación significativa entre dos variables categóricas en una muestra de datos. El chi cuadrado se basa en la comparación de las frecuencias observadas en diferentes categorías con las frecuencias esperadas que se obtendrían si no hubiera asociación entre las variables. Cuanto mayor sea la diferencia entre las frecuencias observadas y esperadas, mayor será el valor del estadístico de chi cuadrado y más probable será que se rechace la hipótesis nula de no asociación. (Narvaez, 2018)

Pasos a seguir para calcular Chi cuadrado

El Chi-cuadrado (χ^2) es una prueba estadística utilizada para evaluar si existe una relación significativa entre dos variables categóricas. A continuación, se detallan los pasos para realizar el Chi-cuadrado. Ejemplo:

Se tienen los datos de 891 pasajeros del Titanic, las clases al que pertenecía y si sobrevivió o no al naufragio.

Pasajero	Sobrevivió	Clase
1	No	3era
2	Si	1era
3	Si	3era
4	Si	1era
5	No	3era
6	No	3era
7	No	1era
8	No	3era
9	Si	3era
10	Si	2da
11	Si	3era
12	Si	1era
13	No	3era
14	No	3era
15	No	3era
16	Si	2da
17	No	3era
18	Si	2da
19	No	3era

20	Si	3era
21	No	2da
22	Si	2da
23	Si	3era
24	Si	1era
25	No	3era
26	Si	3era
27	No	3era
28	No	1era
29	Si	3era
30	No	3era
31	No	1era
32	Si	1era
33	Si	3era
34	No	2da
35	No	1era
36	No	1era
37	Si	3era
38	No	3era
39	No	3era
40	Si	3era
41	No	3era
42	No	2da
43	No	3era
44	Si	2da
45	Si	3era
46	No	3era

47	No	3era
48	Si	3era
49	No	3era
50	No	3era
51	No	3era
52	No	3era
53	Si	1era
54	Si	2da
55	No	1era
56	Si	1era
57	Si	2da
58	No	3era
59	Si	2da
60	No	3era
61	No	3era
62	Si	1era
63	No	1era
64	No	3era

ETC…….

Formulación de hipótesis

- Hipótesis nula (H0): No hay asociación significativa entre las dos variables categóricas.

- Hipótesis alternativa (H1): Existe una asociación significativa entre las dos variables categóricas.

X	H_0: Independencia de las variables
	H_1: Variables relacionadas ✓

LINK:

https://www.kaggle.com/c/titanic/data

Creación de la tabla de contingencia

Crea una tabla de contingencia (o tabla de frecuencias) que muestre la distribución conjunta de las dos variables categóricas que se desean analizar. Esta tabla tendrá filas

Tabla de contingencia

		Clase			
		1era	2da	3era	Total
Sobrevivió?	Si	136	87	119	342
	No	80	97	372	549
	Total	216	184	491	**891**

		Clase			
		1era	2da	3era	Total
Sobrevivió?	Si	**63%**	**47%**	**24%**	**38%**
	No	**37%**	**53%**	**76%**	**62%**
	Total	100%	100%	100%	100%

y columnas que representan las categorías de las dos variables, y en las celdas se indicará la frecuencia observada de ocurrencias conjuntas.

LINK:

https://www.kaggle.com/c/titanic/data

Cálculo de frecuencias esperadas:

Calcula las frecuencias esperadas para cada celda de la tabla

bajo la suposición de que no hay asociación entre las variables. Para esto, se utiliza la fórmula: frecuencia esperada = (total de la fila) x (total de la columna) / (tamaño total de la muestra).

		Clase			
		1era	2da	3era	Total
Sobrevivió?	Si	82,91	70,63	188,46	342,00
	No	133,09	113,37	302,54	549,00
	Total	216,00	184,00	491,00	891,00

		Clase			
		1era	2da	3era	Total
Sobrevivió?	Si	34,00	3,80	25,60	0,00
	No	21,18	2,36	15,95	0,00
	Total	0,00	0,00	0,00	0,00
		55	6	42	102,89

LINK:

https://www.kaggle.com/c/titanic/data

Cálculo del estadístico de Chi-cuadrado

Calcula el estadístico de Chi-cuadrado utilizando la fórmula: χ^2 = Σ [(frecuencia observada - frecuencia esperada) ²/ frecuencia esperada] donde Σ representa la

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} > \text{Valor crítico:}$$

$$\underbrace{102{,}89 > 5{,}99}$$

Estadístico de prueba **SE RECHAZA LA HIPOTESIS NULA**

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \qquad \chi^2 = 102{,}89$$

Valor crítico:	gl = (r-1)(c-1) =	2	
	α =	0,05	5%
	$\chi^2_{(1-\alpha).(r-1)(c-1)}$ =	5,99	=INV.CHICUAD.CD(W18;W17)

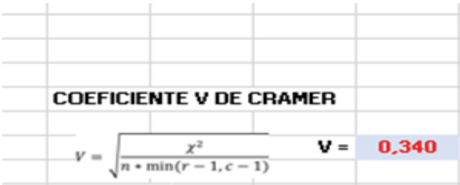
sumatoria sobre todas las celdas de la tabla.

LINK:

https://www.kaggle.com/c/titanic/data

Valor crítico y nivel de significancia

Consulta la tabla de distribución de Chi-cuadrado para obtener el valor crítico correspondiente al nivel de significancia deseado y los grados de libertad obtenidos. El nivel de significancia típico suele ser 0.05 o 0.01.



LINK:

https://www.kaggle.com/c/titanic/data

Comparación y decisión. -

Compara el valor calculado de Chi-cuadrado con el valor crítico. Si el valor calculado es mayor que el valor crítico, se rechaza la hipótesis nula y se concluye que existe una asociación significativa entre las variables. En caso contrario, no se rechaza la hipótesis nula, y no se encuentra evidencia suficiente para afirmar una asociación significativa.

LINK:

https://www.kaggle.com/c/titanic/data

Conclusiones

Se pueden calcular las frecuencias esperadas y observadas para cada categoría. Si las frecuencias observadas difieren significativamente de las esperadas, esto indica que ciertas categorías tienen una mayor influencia en la asociación entre las variables.

Es importante tener en cuenta que la prueba chi cuadrado solo evalúa la asociación entre variables categóricas y no establece relaciones causales 1. Además, la prueba chi cuadrado asume que las frecuencias de las categorías son independientes y que se cumple el tamaño de muestra requerido.

ANÁLISIS DE LA VARIANZA "ANOVA"

Introducción. -

Es una técnica estadística utilizada para comparar las medias de

tres o más grupos en una variable
continua. El objetivo del ANOVA es determinar si existen diferencias significativas entre las medias de los grupos que se están comparando.

Para ello, se utiliza la variabilidad de las observaciones entre y dentro de los grupos. (Andrade, 2019)

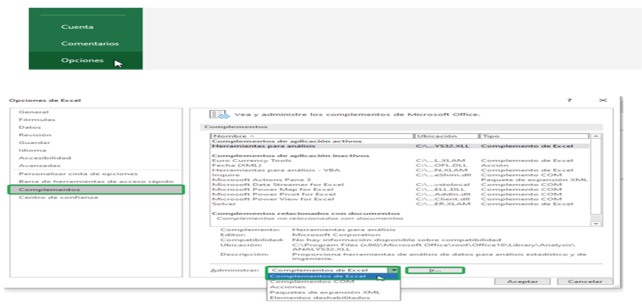
Pasos a seguir para calcular el análisis de la varianza “Anova”

Instalar el paquete de análisis de datos

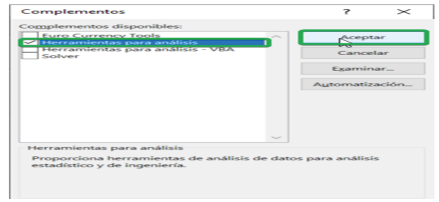
Antes de poder empezar con el análisis de una tabla ANOVA en Excel debemos instalar el paquete de análisis de datos para poder crear una tabla ANOVA.

En la barra de herramientas, hacemos clic en “Archivo” y luego en “Opciones” (en la esquina inferior izquierda).

En la nueva ventana de opciones, debemos seleccionar “Complementos”. En la sección “Administrar”, seleccionamos “Complementos de Excel” y luego hacemos click en “Ir”



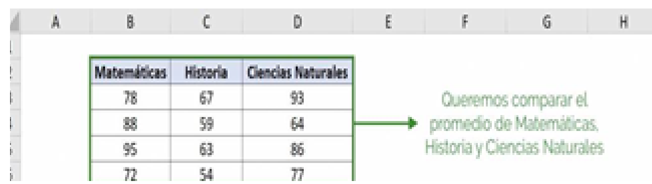
Aparecerá una nueva ventana, en la que debemos verificar “Herramientas para análisis” y luego hacer clic en “Aceptar”.



Si vamos a la pestaña “Datos” de la barra de herramientas, veremos que se ha agregado una nueva sección llamada “Análisis”, la que incluye la herramienta de “Análisis de datos”.

Anova de un Factor

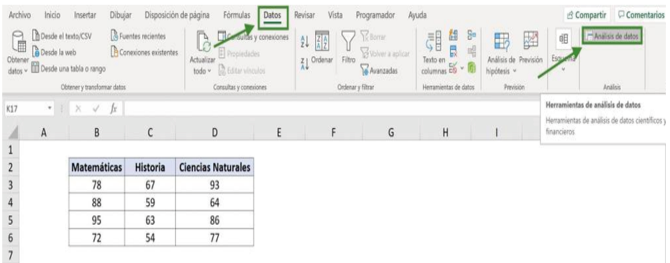
Para entender qué es una tabla ANOVA de un factor veremos un ejemplo. En este caso, que tenemos las notas de un alumno para tres distintas asignaturas: Matemáticas, Historia y Ciencias Naturales.



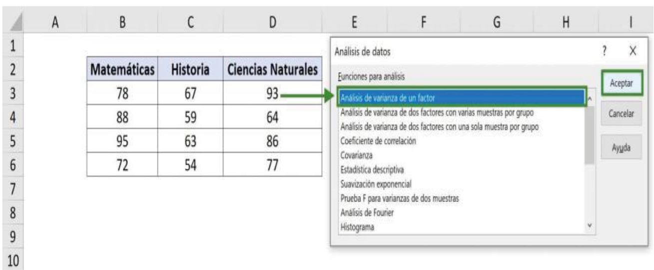
Nuestro propósito es saber si es que el promedio que tiene en cada materia es estadísticamente igual a las otras materias. Esto es, ver si el promedio en Matemáticas es igual el promedio en Historia y el Ciencias Naturales.

Para esto debemos:

Vamos a la pestaña "Datos" y apretamos en "Análisis de datos"

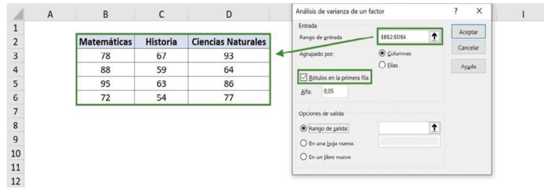


En la ventana que se abre debemos seleccionar "Análisis de Varianza de un factor"

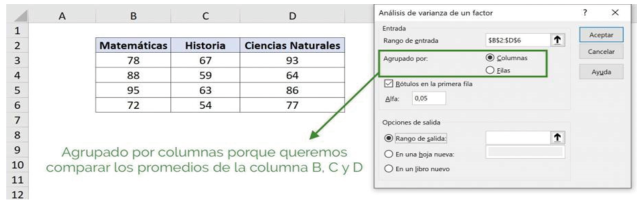


En "Rango de entrada" seleccionamos nuestra tabla con datos.

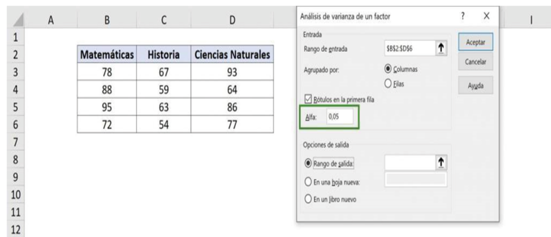
En este caso es \$B\$2: \$D\$6 y como incluimos los títulos debemos seleccionar la opción "Rótulos en la primera fila".



En la sección de "Agrupado por", debemos elegir el sentido de cada grupo: como cada grupo es una columna distinta seleccionamos "Columnas". Esto es porque queremos comparar una columna con otra.

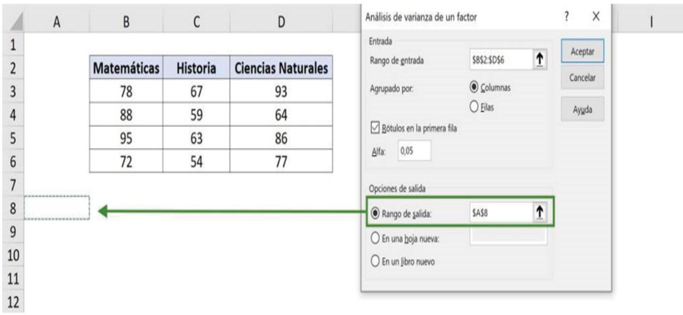


Para el valor de "Alfa" en general se usa 5%, lo que significa que se estima con una confianza del 95%. Si quisiéramos ser más estrictos podemos elegir un valor menor, como 1% (para tener un 99% de confianza). Puedes profundizar en este concepto en este artículo.

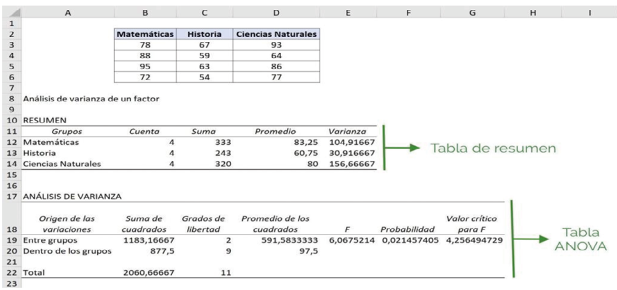


En "Rango de salida" debemos elegir la celda en donde comenzará la información del análisis de varianza, es decir, la esquina superior izquierda.

En este caso seleccionaremos A8. Finalmente hacemos clic en “Aceptar”.



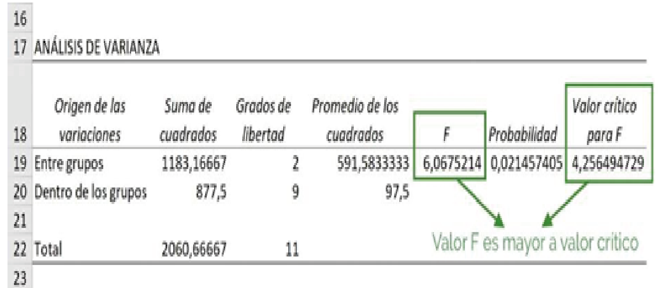
Así, obtenemos como resultado una tabla de resumen de observaciones, promedio y varianza de cada grupo y una tabla ANOVA.



Luego de obtener la tabla ANOVA nuestro objetivo es ver si se cumple o no nuestra hipótesis inicial. si el valor F es mayor al valor crítico de F, vamos a rechazar la hipótesis de que los promedios de los grupos son iguales. Es decir, si F es mayor al valor crítico, vamos a decir al menos uno de los promedios de

estos grupos es estadísticamente diferente.

Para ver esto en la tabla ANOVA misma vamos a la fila que dice “Entre grupos” y vemos que el valor F es 6.06, mientras que el valor crítico de F es 4.25. Esto significa que los promedios de las materias de Matemáticas, Historia y Ciencias Naturales no son todos iguales entre sí, es decir, al menos un promedio es distinto, por lo que se rechaza la hipótesis.



Una vez ya realizado este proceso si se desea saber cual es distinto podemos aplicar el test T para ver cada par de grupos si son iguales o distintos entre sí

Conclusiones. -

Existen diferencias significativas entre al menos dos de las medias de los grupos analizados. Esto implica que al menos dos grupos difieren en términos de la variable en estudio.

La variable independiente (factor) utilizada en el análisis tiene un impacto significativo en la variable dependiente. Esto significa que el factor utilizado para categorizar los grupos tiene una influencia en la variable que se está analizando.

BIBLIOGRAFÍA

- Andrade, J. (2019). Staditical discovery. Obtenido de Staditical discovery: https://www.jmp.com/es_co/statistics-knowledge-portal/one-way-anova.html
- BARBARA ILLOWSKY, D. C. (2022). Estadistica. Obtenido de Estadistica: https://assets.openstax.org/oscms-prodcms/media/documents/Introduccion_al_la_estadistica_-_WEB.pdf
- López, J. F. (2019). Economipedia . Obtenido de Economipedia : https://economipedia.com/definiciones/distribucion-t-de-student.html
- MateClassroom. (6 de Junio de 2021). Prueba de hipótesis de una muestra de ejercicios resueltos. Obtenido de YouTube: https://www.youtube.com/watch?v=RS5F_bhNugw&t=21s
- Narvaez, M. (2018). questionpro . Obtenido de questionpro : https://www.questionpro.com/blog/es/prueba-de-chi-cuadrado-de-pearson/
- Hogg, R. V., & Tanis, E. A. (2020). “Probability and Statistical Inference” (9th ed.). Pearson.
- Casella, G., & Berger, R. L. (2008). “Statistical Inference” (2nd ed.). Duxbury Press.
- Devore, J. L., & Berk, K. N. (2018). “Modern Mathematical Statistics with Applications” (2nd ed.). Springer.
- Agresti, A., & Finlay, B. (2018). “Statistical Methods for the Social Sciences” (5th ed.). Pearson.
- Freedman, D., Pisani, R., & Purves, R. (2007). “Statistics” (4th ed.). W. W. Norton & Company.
- Johnson, R. A., & Kuby, P. (2020). “Elementary

Statistics" (12th ed.). Cengage Learning.

- Rice, J. A. (2006). "Mathematical Statistics and Data Analysis" (3rd ed.). Cengage Learning.
- Wasserman, L. (2004). "All of Statistics: A Concise Course in Statistical Inference." Springer.
- Montgomery, D. C., & Runger, G. C. (2017). "Applied Statistics and Probability for Engineers" (7th ed.). Wiley.
- Lehmann, E. L., & Casella, G. (1998). "Theory of Point Estimation" (2nd ed.). Springer.
- Matemóvil.
- Valencia, Roberto. (2018). Libro de Estadística.
-

CASA EDITORA DEL POLO

ISBN: 978-9942-621-61-0

9 789942 621610