



Universidad
Continental

PROBABILIDAD Y ESTADÍSTICA

Guía de Trabajo



Visión

Ser una de las 10 mejores universidades privadas del Perú al año 2020, reconocidos por nuestra excelencia académica y vocación de servicio, líderes en formación integral, con perspectiva global; promoviendo la competitividad del país.

Misión

Somos una universidad privada, innovadora y comprometida con el desarrollo del Perú, que se dedica a formar personas competentes, íntegras y emprendedoras, con visión internacional; para que se conviertan en ciudadanos responsables e impulsen el desarrollo de sus comunidades, impartiendo experiencias de aprendizaje vivificantes e inspiradoras; y generando una alta valoración mutua entre todos los grupos de interés.

Universidad Continental

Material publicado con fines de estudio

Código: UC0677

2017



Presentación

El material, está diseñado para orientar al estudiante, el desarrollo de aplicaciones prácticas relacionadas al avance teórico de la asignatura de Probabilidades y Estadística. Esta asignatura está diseñada para proporcionar al estudiante, los conocimientos y habilidades necesarias en la recolección, procesamiento, análisis e interpretación de datos numéricos, así como también en el desarrollo de la teoría de la probabilidad, la cual ha aumentado el alcance de las aplicaciones de la estadística.

En general, contiene un compendio de guías prácticas para ser desarrolladas de manera (secuencial), está estructurada en cinco unidades: en la primera unidad se desarrollara la introducción a la asignatura; en la segunda unidad, un resumen y la aplicación de graficas de datos; en la tercera unidad, los estadísticos necesarios para describir, explorar y comparar datos; en la cuarta unidad, se desarrollara el tema de probabilidades; finalmente, en la quinta unidad se busca ampliar el tema de probabilidad hasta las distribuciones de probabilidad y su aplicación.

La elaboración de la presente guía es fruto de la investigación que ha sido enriquecido a partir de la revisión de manuales y libros de la asignatura, utilizando organizadores y demás cuadros para que sea de fácil entendimiento para el estudiante.

Es recomendable que el estudiante antes de desarrollar la guía el material de trabajo desarrolle una permanente lectura de estudio para entender el procedimiento y trabaje con seriedad.

Estimado estudiante, este material llega a sus manos con la intención de guiarlos en su aprendizaje en la asignatura pero así mismo debe ser complementada con la biografía propuesta en el silabo del curso.

El equipo de docentes



ÍNDICE

	Pág.
ÍNDICE.....	4
PRIMERA UNIDAD	5
Semana N°1: INTRODUCCIÓN A LA ESTADÍSTICA	6
Semana N°2: DISTRIBUCIÓN DE FRECUENCIAS.....	10
Semana N°3: DISTRIBUCIÓN DE FRECUENCIAS BIDIMENSIONALES, RECONSTRUCCIÓN DE TABLAS	21
Semana N°4: GRÁFICA ESTADÍSTICOS.....	25
SEGUNDA UNIDAD	32
Semana N°5: MEDIDAS DE TENDENCIA CENTRAL.....	33
Semana N°6: MEDIDAS DE DISPERSIÓN	41
Semana N°7: MEDIDAS DE POSICIÓN RELATIVA	47
Semana N°8: MEDIDAS DE FORMA	52
TERCERA UNIDAD	56
Semana N°10: PROBABILIDADES.....	57
Semana N°11: REGLA DE LA SUMA Y DE LA MULTIPLICACIÓN	64
CUARTA UNIDAD	76
Semana N°13: VARIABLES ALEATORIAS	77
Semana N°14: DISTRIBUCIONES DE PROBABILIDAD DISCRETAS.....	82
Semana N°15: DISTRIBUCIONES DE PROBABILIDAD CONTINUAS.....	88
Semana N°16: APLICACIONES DE LA DISTRIBUCIÓN NORMAL	96
ANEXOS	103



GUÍA DE PRÁCTICA DE PROBABILIDADES Y ESTADÍSTICA

PRIMERA UNIDAD INTRODUCCIÓN A LA ESTADÍSTICA, RESUMEN Y GRÁFICA DE DATOS

- GUÍA DE PRÁCTICA N° 1: Introducción a la Estadística.
- GUÍA DE PRÁCTICA N° 2: Distribución de frecuencias unidimensionales.
- GUÍA DE PRÁCTICA N° 3: Distribución de frecuencias bidimensionales, reconstrucción de tablas.
- GUÍA DE PRÁCTICA N° 4: Gráfica de datos

**Semana N°1: INTRODUCCIÓN A LA ESTADÍSTICA****1. TEMA: Estadística, división, Objetivos, definiciones, tipos de datos. Métodos y fuentes de recolección de datos.****2. PROPÓSITO:**

- Define la estadística e identifica los tipos de datos en situaciones cotidianas.

Estadística.- Es la ciencia por medio de la cual se recogen, organizan, presentan, analizan e interpretan datos con el fin de proporcionar una toma de decisiones más eficaz. (Lind, Marshal y Wathen. 2015).

División.- Los métodos estadísticos se dividen en dos grandes ramas.

Estadística descriptiva.- Conjunto de métodos cuya misión es ordenar, describir y resumir los datos recogidos. Con estos métodos se elaboran tablas, gráficos y medidas descriptivas de tendencia central, posición, dispersión y deformación. La finalidad de la Estadística Descriptiva no es generalizar resultados sobre el fenómeno que ha producido los datos bajo estudio, sino solamente su descripción.

Estadística inferencial.- Conjunto de métodos con los que se generalizan o infieren los resultados de una muestra a la población de donde fue extraída. Para realizar esta inferencia es necesario que sean dadas con una medida de confiabilidad. Comprende la teoría de estimación y prueba de hipótesis.

Objetivo.- La estadística tiene por objetivos:

1. Describir grandes colecciones de datos empíricos reduciéndolos a un pequeño número de características que concentra la parte más importante y significativa de la información proporcionada por los datos se expresan por un conjunto de indicadores, medidas de resumen o estadígrafos.

2. Realizar el análisis estadístico de datos experimentales y de los fenómenos observados. La preocupación del análisis estadístico es inferir propiedades para una población sobre la base de resultados muestrales conocidos. Aquí se presenta varios problemas que presenta la Estadística, la estimación estadística, el cálculo de probabilidades, las pruebas estadísticas, etc.

3. Predecir el comportamiento de los fenómenos en el futuro, lo cual constituye la máxima aspiración práctica de toda ciencia. Este objetivo de predicción y previsión está implícito tanto en la descripción como en el análisis estadístico, puesto que en general interesa orientar la toma de decisiones con vigencia y afecto en el futuro.

Naturalmente que las estimaciones y proyecciones dependen del grado de conocimiento del comportamiento del pasado y presente de las variables en estudio.

Definiciones.-

Datos.- Son las observaciones recolectadas (como mediciones, géneros, respuestas de encuestas).

Población.- Es el conjunto completo o total de los elementos (puntuaciones, personas, medidas, etc.) que se va estudiar y que posee al menos una característica común observable, cuyo estudio nos interesa o acerca de los cuales se desea información. La población debe estar perfectamente definida en el tiempo y en el espacio, de modo que ante la presencia de un potencial integrante de la misma, se pueda decidir si forma parte o no de la población bajo estudio. La población puede ser según su tamaño de dos tipos:

Población finita: cuando se tiene un número determinado de elementos.

Población infinita: cuando el número de elementos es indeterminado, o tan grande que pudiesen considerarse infinitos.

Censo.- Es un método de recolección de datos mediante el cual la información se obtiene analizando a la totalidad de los elementos que componen la población o universo bajo estudio.

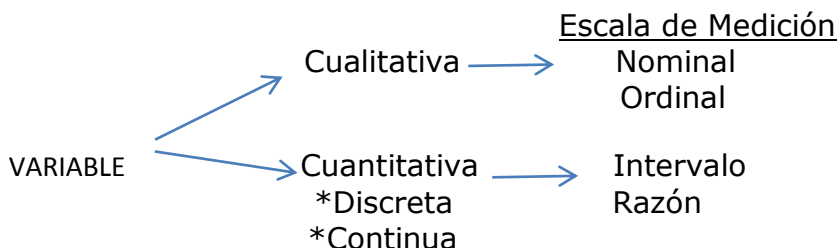
Parámetro.- Es un número que describe alguna característica de la población o medida de resumen de una población. Se considera como un valor verdadero de la característica estudiada y para determinar su valor es necesario utilizar la información poblacional completa, y por lo tanto la decisión se toma con certidumbre total.



Estadístico.- Es un número que describe alguna característica de la muestra o medida de resumen de una muestra y la toma de decisión contiene un grado de incertidumbre.

Unidad o Elemento estadístico.- Es cada uno de los elementos que componen la población, es llamado también individuo.

Tipo de variables.- Las variables por su naturaleza y escala de medición se clasifican en:



Variables cualitativas.- Se dividen en diferentes categorías que se distinguen por algunas características **no numéricas**. Como por ejemplo: el género, deporte preferido, etc.

Variables cuantitativas.- Cuando el valor de la variable se expresa por una cantidad, es de carácter numérico, el dato o valor puede resultar de la operación de contar o de medir y se divide en: discretas y continuas.

*Variable discreta.- Son aquellas que surgen del procedimiento del conteo y suelen tomar valores enteros. Como el número de hijos por familia, número de llamadas por día, etc.

* Variable continua.- Cuando los datos son susceptibles de ser medidos. Pueden asumir valores decimales. Ejemplos; área de parcelas, ingresos, producción de maíz, peso, estatura, tiempo de servicios, horas trabajadas, etc.

Escalas de Medición.-

NOMINAL	Son aquellas variables cualitativas que establecen la distinción de los elementos en diversas categorías (nombres, etiquetas), sin implicar algún orden entre ellas	Color de ojos, género, etc.
ORDINAL	Aquellas variables cualitativas que implican orden entre sus categorías, pero no grados de distancia igual entre ellas, están referidas a un orden de jerarquía, donde las categorías expresan una posición de orden.	Grado de Instrucción, grado de satisfacción de los clientes, etc.
INTERVALO	Son aquellas variables cuantitativas que suponen a la vez orden y grados de distancia iguales entre las diversas categorías, pero no tienen origen natural, sino convencional, tiene un cero relativo.	Temperatura Coeficiente de inteligencia
RAZÓN	Estas variables comprenden a la vez todos los casos anteriores, distinción, orden, distancia y origen único natural; el valor se expresa con un número real tiene un cero absoluto.	Edad, peso, ingresos, producción, etc.



EJERCICIOS RESUELTOS:

En los ejercicios 1 al 4 determine si el valor dado es un estadístico o un parámetro.

1.-Tamaño de la familia. Se selecciona una muestra de hogares y el número promedio (media) de personas por familia es de 2,58 (según datos de la Oficina censal Peruana).

2.-Política. En la actualidad, el 42% de los gobernadores regionales del Perú son de izquierda.

3.-Titanic. En un estudio de los 2223 pasajeros del "Titanic", se encontró que 706 sobrevivieron cuando se hundió.

4.- Audiencia televisiva. Se selecciona una muestra de ciudadanos peruanos y se descubre que la cantidad de tiempo promedio (media) que ven la televisión es de 4,6 horas al día

Respuesta: 1.Estadístico 2.Parámetro 3.-Parámetro 4.-Estadístico

En los ejercicios 5 y 6 determine si la variable es cualitativa o cuantitativa y señale si son discretas o continuas.

5.- a) Lugar de residencia

b) Número de vecinos de un edificio.

c) Profesiones de empleados.

d) Número de llamadas telefónicas.

e) Consumo de gasolina cada 200 km.

Respuesta: a) Cualitativa, b) Cuantitativa discreta c) Cualitativa d) Cuantitativa discreta e) Cuantitativa continua.

6.- Con el fin de conocer la mejor forma de viajar de una población han preparado una encuesta. Algunas de las preguntas trataron sobre: N° de días de viaje, dinero empleado, número de equipajes, zonas geográficas, medios de transporte, naturaleza del viaje (negocios, turismo, salud etc.) y Numero de personas.

Respuesta:

V. Cualitativa: Zonas geográficas, medios de transporte y naturaleza del viaje.

V. Cuantitativa discreta: N° De días, N° De equipajes y N° De personas.

V. Cuantitativa continua: Dinero empleado.

Ejercicios Propuestos:

En los ejercicios 1 a 8 determine cuál de los cuatro niveles de medición (nominal, ordinal, de intervalo o de razón) es el más apropiado.

1.-Maratón en Huancayo. Los números en las camisetas de los corredores de maratones.

2.-Producto de consumo. Las calificaciones que da la revista *Somos* de "la mejor compra": recomendado, no recomendado".

3.-NSS. Los números de seguridad social.

4.-Encuesta de bebidas. El número de respuestas "si" recibidas cuando se les preguntó a 500 estudiantes si alguna vez se habían embriagado en la universidad.

5.-Cigarras. Los años de aparición de cigarras: 1936; 1953; 1970; 1987 y 2004.

6.-Mujeres ejecutivas. Los salarios de mujeres que son directoras generales de corporaciones.

7.- Calificaciones. Calificaciones de las películas de una estrella, dos estrellas, tres estrellas y cuatro estrellas.

8.-Temperaturas. Las temperaturas actuales en las ciudades del Perú.

9.-El Directorio y la Gerencia de la Universidad Continental han realizado un estudio para conocer la opinión de los padres de familia de los estudiantes en general, respecto a las nuevas carreras que se vienen ofertando. Para ello, durante la semana de matrículas se aplicó una encuesta a 860 padres de familia elegidos aleatoriamente, dentro de las instalaciones del campus universitario, donde se obtuvo como resultado que el 87% de los encuestados se manifestaron en total acuerdo por la innovación en carreras profesionales que se viene impulsando (las respuestas iban de "Totalmente en desacuerdo" a "Totalmente de acuerdo"). De estudios anteriores se sabía que sólo el 55% de los padres de familia de los estudiantes de esta casa superior de estudios apoyaban la iniciativa de fomentar nuevas carreras en años pasados.

Del enunciado anterior, indique:

a) Población y parámetro: _____



- b) Muestra y estadístico: _____
c) Variable: _____
d) Tipo de variable: _____
e) Nivel de medición: _____
f) Unidad estadística: _____

10.- Responda las siguientes preguntas:

I. Cuando Alberto Fujimori fue elegido por tercera vez para la presidencia, recibió el 49,9% de los 12 066 229 votos emitidos. El conjunto de todos esos votos es la población a considerar. ¿El 49,9% es un parámetro o un estadístico?

II. En el inciso (I) indica el total de votos emitidos en la elección presidencial del 2000. ¿Estos valores son cualitativos, cuantitativos discretos o cuantitativos continuos?

III. ¿Qué nivel de medida tiene la variable en función del Inciso (II)?

- A) Estadístico, cualitativos, nominal
B) Parámetro, cuantitativo discreto, intervalo
C) Estadístico, cuantitativo discreto, escala
D) Parámetro, cuantitativo continuo, escala
E) Parámetro, cuantitativo discreto, razón

Referencias bibliográficas consultadas y/o enlaces recomendados

- Bejarano, M. (1995) Estadística descriptiva, probabilidades y lineamientos para la elaboración del protocolo de investigación en ciencias de la Salud. Universidad Peruana Cayetano Heredia.
- Triola F. (2009) Estadística Elemental. Editorial Mexicana. Décima Edición.

**Semana N°2: DISTRIBUCIÓN DE FRECUENCIAS****1. TEMA: Distribuciones de Frecuencias Unidimensionales: variables cualitativas y cuantitativas****2. PROPÓSITO:**

- Elabora distribuciones de frecuencia y gráficos, utilizando el software estadístico Excel.
- Organiza los datos cualitativos en una tabla de frecuencia y gráficos, para interpretar los resultados.

DISTRIBUCIÓN DE FRECUENCIAS

La siguiente fase a la recolección de datos es la crítica de campo, lo que significa realizar la validación y consistencia de los datos obtenidos durante el trabajo de campo, con la finalidad que en la fase de organización y clasificación de los datos sea consistente y veraz los datos a procesarse, analizarse y con la cual se va tomar decisiones.

Para iniciar la organización de datos definiremos algunos conceptos:

1. Clase: Es una división de la variable. Se denota como subíndice con la letra "i" y el número total de clases con "m"

2. Frecuencia: Es las veces que se repite una clase de la variable. Estas son: Simples y Acumuladas

2. 1 Frecuencias Simples: Son aquellas frecuencias que sólo corresponden a una clase de la variable entre ellas tenemos:

2.1.1. Frecuencia Absoluta Simple

La frecuencia absoluta simple de la clase ci es el número fi , de observaciones que presentan una modalidad perteneciente a esa clase.

Además se cumple que:
$$\sum_{i=1}^m f_i = n$$

2.1.2 Frecuencia Relativa Simple

Frecuencia relativa simple de la clase ci es el cociente hi , entre las frecuencias absolutas de dicha clase y el número total de observaciones, es decir: $h_i = \frac{f_i}{n}$

Obsérvese que fi es el *tanto por uno* de observaciones que están en la clase ci . Multiplicado por $\frac{100}{n}$ representa el porcentaje de la población que comprende esa clase.

También cumple:
$$\sum_{i=1}^m h_i = 1$$

2.1.3 Frecuencia Porcentual Simple

Frecuencia Porcentual Simple de la clase ci es el producto de pi , entre las frecuencias relativas de dicha clase por 100, es decir $p_i = h_i * 100$

Cumple lo siguiente:
$$\sum_{i=1}^m p_i = 100$$

2.2 Frecuencias Acumuladas: Aquellas frecuencias que se obtienen por la suma de dos o más clases de la variable

2.2.1 Frecuencia Absoluta Acumulada

Fi , se calcula sobre variables cuantitativas, y es el número de elementos de la población cuya modalidad es inferior o equivalente a la modalidad ci :

2.2.2 Frecuencia Relativa Acumulada

$$F_i = f_1 + f_2 + \dots + f_i = \sum_{k=1}^i f_k$$

Hi , se calcula sobre variables cuantitativas, siendo el tanto por uno de los elementos de la población que están en alguna de las clases y que presentan una modalidad inferior o igual a la ci , es decir:



$$H_i = \frac{F_i}{n} = h_1 + h_2 + \dots + h_i = \sum_{k=1}^i h_k$$

2.2.3 Frecuencia Porcentual Acumulada

Pi, se calcula sobre variables cuantitativas, siendo el tanto por ciento de los elementos de la población que están en alguna de las clases y que presentan una modalidad inferior o igual a la ci, es decir :

$$P_i = H_i * 100 = p_1 + p_2 + \dots + p_i = \sum_{k=1}^i p_k$$

DISTRIBUCIÓN DE FRECUENCIAS

Llamaremos **distribución de frecuencias o tabla de frecuencias** al arreglo de filas y columnas que contiene al conjunto de clases junto a las frecuencias correspondientes a cada una de ellas. Una **tabla estadística** sirve para clasificar y ordenar los datos estadísticos.

Distribución de Frecuencias Unidimensionales: Se puede aplicar a variables cualitativas y cuantitativas y se le denomina así cuando se está observando UNA SOLA VARIABLE de cada elemento de una muestra o población.

Organización de datos cualitativos

Determinar si los datos por procesar corresponden a variables cualitativas nominales u ordinales.

Un punto que es necesario resaltar es la elección de un título apropiado para cada cuadro, tabla o figura.

EJEMPLO:

El restaurante "Pizzería Italia" busca evaluar el grado de satisfacción de sus clientes: donde B es Bueno, R es Regular, D es Deficiente. Se tienen los resultados en el siguiente cuadro:

B	B	R	R	R	R	B	B	D	D	D	B	B	B	R	R	R	B	B	B	B	B	B
R	R	R	R	R	D	D	R	R	B	R	D	D	D	D	R	R	R	B	B	B	R	R
R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	R	D	D	D

Para organizar los datos elaboremos la tabla de frecuencias, grafico respectivo e interpretación.

Después del conteo de los datos por categorías tenemos:

Bueno (B): 17 Regular (R): 40 Malo (M): 12

SATISFACCION DEL CLIENTE	Frecuencia Absoluta fi	Frecuencia Absoluta Acumulada Fi	Frecuencia Relativa hi	Frecuencia Relativa Acumulada Hi	Frecuencia Porcentual pi%	Frecuencia Porcentual Acumulada Pi%
BUENO						
REGULAR						
DEFICIENTE						
TOTAL						

TABLA DE FECUENCIA RELLENADA



Tabla No.1
Nivel de satisfacción de los clientes

SATISFACCION DEL CLIENTE	Frecuencia Absoluta f_i	Frecuencia Absoluta Acumulada F_i	Frecuencia Relativa h_i	Frecuencia Relativa Acumulada H_i	Frecuencia Porcentual $p_i\%$	Frecuencia Porcentual Acumulada $P_i\%$
BUENO	17	17	0,25	0,25	25%	25%
REGULAR	40	57	0,58	0,83	58%	83%
DEFICIENTE	12	69	0,17	1	17%	100%
TOTAL	69		1		100%	

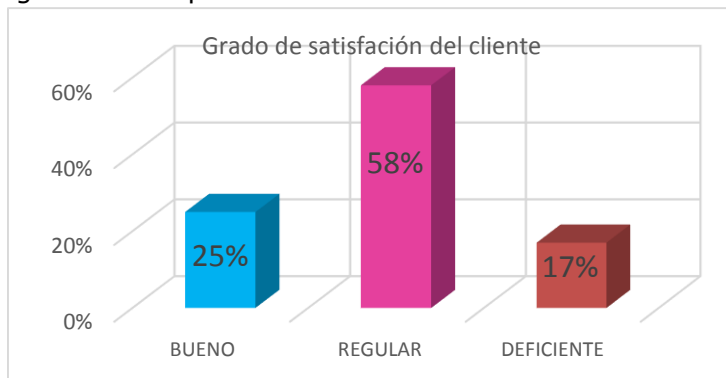
Aquí va el conteo de cada categoría

Divide la frecuencia absoluta (f_i) entre el total de datos: $=17/69=0,25$

$p_i\%$ se obtiene multiplicando la frecuencia relativa (h_i) por 100

$57 = 17 + 40$

- F_i , H_i , $P_i\%$ son las frecuencias acumuladas donde la primera fila siempre será igual a su respectiva frecuencia anterior.



INTERPRETACIÓN:

De los 69 comensales encuestados de la “Pizzería Italia” más de la mitad están satisfechos regularmente con un 58% y otro porcentaje menor del 25% manifiesta que es bueno.

Organización de datos cuantitativos discretos.

Según Toma, J (2011), nos dice que cuando se tienen datos cuantitativos discretos cuyo número de resultados posibles no es grande (no es mayor de 15), la información puede ser clasificada y presentada directamente sin pérdida de la identidad de la misma. Para ello se ordenan los valores de la variable según magnitud y se obtienen las frecuencias absolutas mediante un proceso de conteo.

Tabla No.2
Número de trabajadores eventuales de 68 empresas comerciales
Huancayo -2015

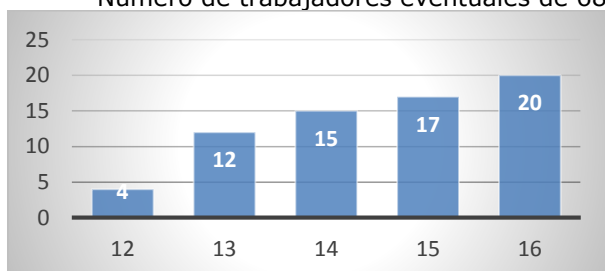
Número de trabajadores eventuales	Frecuencia Absoluta f_i	Frecuencia Absoluta Acumulada F_i	Frecuencia Relativa h_i	Frecuencia Relativa Acumulada H_i	Frecuencia Porcentual $p_i\%$	Frecuencia Porcentual Acumulada $P_i\%$
-----------------------------------	---------------------------	-------------------------------------	---------------------------	-------------------------------------	-------------------------------	---



12	4	4	0.06	0.06	6%	6%
13	12	16	0.18	0.24	18%	24%
14	15	31	0.22	0.46	22%	46%
15	17	48	0.25	0.71	25%	71%
16	20	68	0.29	1	29%	100%
Total	68		1		100%	

Figura No.2

Número de trabajadores eventuales de 68 empresas comerciales. Huancayo-2016



INTERPRETACION:

De las 68 empresas comerciales encuestadas, el 29% de las empresas tiene 16 trabajadores eventuales mientras que el 6% tiene a 12 trabajadores eventuales dentro de su empresa.

Organización de datos cuantitativos continuos

Cuando se tiene información de una variable cuantitativa continua, los datos son clasificados de acuerdo con ciertos intervalos de clase, por lo cual es necesario tener en cuenta la amplitud y los números de intervalos a generar.

NÚMERO DE INTERVALOS DE CLASE:

El símbolo a usar es "m" no existen normas definidas respecto al número ya que si los intervalos de clase son muy pocos se pierden detalles y si son muchos aparte de lo tedioso del trabajo se manifiestan irregularidades. La mayoría de los analistas recomiendan entre 5 y 20 intervalos de clase. Por regla general es preferible hallar utilizando la **regla de Sturges**; donde: $m=1+3,32 \log (n)$ donde n =Número de observaciones.

REGLAS GENERALES PARA FORMAR TABLAS DE DISTRIBUCION DE FRECUENCIAS CONTINUAS.

Intervalos cerrados y abiertos >

Para **datos cuantitativos continuos** se sigue el procedimiento siguiente:

1. Hallar el rango: $R= X_{Max} - X_{min}$
2. Hallar el número de intervalos de clase, utilizando la regla de Sturges: $m=1+3,32 \log (n)$.
3. Hallar la amplitud (a) $a=R/m$, cuando el cociente R/m no es exacto, el valor de "a" debe ser redondeando al valor inmediato superior, según las cifras decimales de los datos observados.



4. Generar los límites de los intervalos. Para el primer intervalo se considera como límite inferior (LI) al valor de la observación de menor magnitud, es decir $LI_1=X_{min}$.

Los límites inferiores (LI) y superiores (LS) de los otros intervalos se obtienen hallando:

$$LI_i = LI_{(i-1)} + a \quad \text{para } i=2, 3, \dots, m$$

$$LS_i = LI_{(i+1)} \text{ para } i=1, 2, \dots, m-1; \text{ o también}$$

$$LS_i = LI_{(i-1)} + a \quad \text{para } i=2, 3, \dots, m$$

5.-Cada uno de los intervalos (LI,LS) se considera cerrado a la izquierda y abierto a la derecha, es decir desde LI, a menos de LS. Esta regla no se aplica necesariamente para el último intervalo. Si el último intervalo superior tiene el mismo valor que la observación de mayor magnitud, deberá considerarse cerrado en ambos extremos, es decir, se considera desde LI_m hasta LS_m .(Toma, J. 2011)

Marcas de clase: Cuando se desea calcular una medida estadística a partir de datos clasificados en intervalos de clase, se trabaja con los valores representativos de cada intervalo. A estos valores que representan a la información de las observaciones de cada intervalo se les llama marca de clase y numéricamente se obtiene de la siguiente manera:

$$X_i = \frac{LI_i + LS_i}{2}$$

El procedimiento descrito anteriormente puede ser aplicado también cuando se tienen datos cuantitativos discretos cuyo número de resultados posibles es grande (es mayor de 15), y su representación gráfica ya no serán bastones. (Toma, J. 2011)

- **Variable Continua**

Ejemplo:

Suponga que los datos que se presentan a continuación corresponden a los valores de la inflación anual durante el año 2015, de un total de 50 países del mundo.

8,2	10,2	8,5	13,1	10,2	10,3	11,2	9,7	11,7	13,8
11,4	13,3	12,8	13,9	10,5	10,7	8,4	10,1	9,1	9,3
11,6	9,8	11,8	11,2	11,4	10,6	9,4	9,0	10,5	12,1
15,1	11,0	12,8	12,2	11,8	11,7	8,2	14,4	12,5	10,1
10,3	9,5	14,8	9,5	12,7	12,8	9,6	15,5	9,7	13,6

Elaborando la tabla de distribución de frecuencias:

1. $R = X_{max} - X_{min} = 15,5 - 8,2 = 7,3$
2. $m = 1 + 3.32 \log(50) = 6,640580 = 7$
3. $a = 7,3/7 = 1,04286 = 1,1$ (redondeo por exceso, a la décima superior, considerando que los datos tienen un decimal significativo)
4. Los límites de los intervalos de clase se obtienen de la siguiente manera:

$$LI_1 = 8,2$$

$$LI_2 = 8,2 + a = 8,2 + 1,1 = 9,3$$

$$LI_3 = 9,3 + 1,1 = 10,4$$

$$LS_1 = LI_2 = 9,3$$

$$LS_2 = LI_3 = 10,4.$$

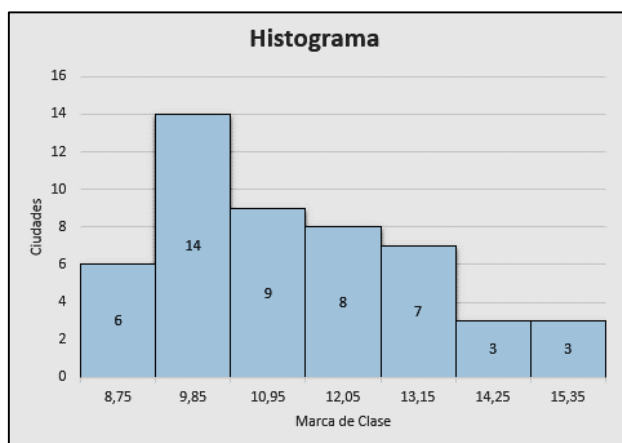


De manera similar se obtienen los otros límites de clase.

Tabla No.3
Distribución de la inflación anual en 50 ciudades.
Perú-2016

Inflación anual (intervalos de clase)	Marca de clase X_i	Ciudades Frec.abs f_i	Frec. Acum. abs. F_i	Frec.rel. h_i	Frec .acum.rel. H_i	Frec. porcentual $p_i\%$	Frec.acum Porcentual $P_i\%$
[8,2 -9,3 >	8,75	6	6	0,12	0,12	12%	12%
[9,3 – 10,4 >	9,85	14	20	0,28	0,40	28%	40%
[10,4 –11,5 >	10,95	9	29	0,18	0,58	18%	58%
[11,5 – 12,6>	12,05	8	37	0,16	0,74	16%	74%
[12,6 –13,7>	13,15	7	44	0,14	0,88	14%	88%
[13,7 – 14,8>	14,25	3	47	0,06	0,94	6%	94%
[14,8 –15,9>	15,35	3	50	0,06	1	6%	100%
Total		50		1		100%	

Con la información de esta tabla se pueden obtener las representaciones gráficas llamadas **histograma de frecuencias, polígono de frecuencias y ojivas.**



INTERPRETACION:
De las 50 países analizadas del mundo, 14 de ellas que tienen una inflación anual de 9,3 a menos de 10,4 que representa el 28%. Y el 88% de los países del mundo tienen una inflación de 8,2 hasta menos de 13,7.



Variable Discreta

Ejemplo:

El Departamento de Personal de la empresa SEDAM JUNÍN ha entregado a la Gerencia el récord de tardanzas de los empleados del personal técnico (medido en minutos) correspondiente al mes de febrero. Los resultados se muestran en la tabla adjunta.

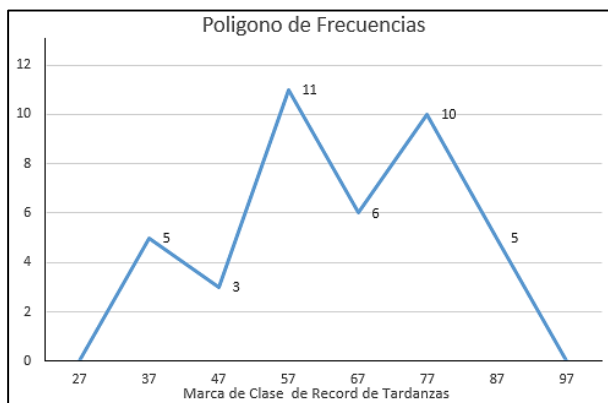
- Elabore la tabla de frecuencias correspondiente, de acuerdo al tipo de variable elabore el polígono de frecuencias.

60	32	85	52	65	77	84	65	57	74
71	81	35	50	35	64	74	47	68	54
80	41	61	91	55	73	59	53	45	77
60	32	85	52	52	65	77	84	77	77

Elaborando la tabla de distribución de frecuencias:

1. $R = X_{\max} - X_{\min} = 91 - 32 = 59$
2. $m = 1 + 3.32 \log(40) = 6,31883917 = 6$
3. $a = 59/6 = 9,833333 = 10$ (redondeo por exceso, a la décima superior, considerando que los datos tienen un decimal significativo)

Record de Tardanzas (minutos)	Marca de clase X_i	Frec.abs f_i	Frec. Acum. abs. F_i	Frec.rel. h_i	Frec. acum.rel. H_i	Frec. porcentual $p_i\%$	Frec.acum Porcentual $P_i\%$
[32 - 42 >	37	5	5	0,125	0,125	12,5%	12,5%
[42 - 52 >	47	3	8	0,075	0,2	7,5%	20%
[52 - 62 >	57	11	19	0,275	0,475	27,5%	47,5%
[62 - 72 >	67	6	25	0,15	0,625	15%	62,5%
[72 - 82 >	77	10	35	0,25	0,875	25%	87,5%
[82 - 92 >	87	5	40	0,125	1	12,5%	100%
Total		40		1		100%	



INTERPRETACION:

De las 40 trabajadores de la empresa SEDAM JUNIN, 11 trabajadores llegan tarde de 52 a menos de 62 minutos tarde, que representa el 27,5%. Por ende, se les descontara a todos los que exceden los 30 minutos acumulados en el mes de Febrero.



TEN PRESENTE: Reglas internacionales de redondeo de números:

Con el objeto de hacer más operativo el manejo de estos datos se redondean. Para "cortar" o redondear, para ello es necesario el uso de las siguientes reglas, aceptadas internacionalmente:

1ª REGLA: Si la cifra que sigue es menor que 5 se deja el dígito precedente intacto.
Ejemplo: Redondear a 2 decimales.

4.123 4.12
2.141 2.14

2º REGLA: Si la cifra que sigue es mayor a 5, se aumenta una unidad el dígito precedente.
Ejemplo: Redondear a 2 decimales

6,176 6,18
4,128 4,13

3º Regla: Si el dígito a la derecha del último requerido es un 5 **seguido de cualquier dígito diferente de cero**, se aumenta una unidad el dígito precedente.

Ejemplo: Redondear a 2 decimales
9,4252 9,43
2,1754 2,18

4º REGLA: Si la cifra que sigue es 5 **no seguido de dígitos** y el número anterior es **par no se modifica**. Pero si el número es impar se aumenta una unidad el dígito precedente.

Ejemplo: Redondeo a 2 decimales
6,545 6,54
1,975 1,98

(Jorge, M. y Edgardo, R. 1998)

EJERCICIOS PARA RESOLVER EN CLASE:

1.- Un estudiante de la Universidad Continental ha elegido como tema de investigación la marca de vehículos que prefieren los trabajadores de esta universidad. Para ello aplicó una encuesta a un grupo de ellos obteniendo los siguientes resultados:

Elabore la tabla de frecuencias correspondiente, de acuerdo al tipo de variable, elabore el gráfico e interprete: f_3 y el $P_4\%$.

HYUNDAI	SUSUKI	NISSAN	TOYOTA	KIA	SUSUKI	KIA	KIA
SUSUKI	TOYOTA	NISSAN	KIA	SUSUKI	SUSUKI	KIA	KIA
KIA	HYUNDAI	SUSUKI	NISSAN	HYUNDAI	HYUNDAI	NISSAN	SUSUKI
SUSUKI	TOYOTA	TOYOTA	NISSAN	HYUNDAI	HYUNDAI	NISSAN	HYUNDAI
NISSAN	TOYOTA	HYUNDAI	SUSUKI	NISSAN	HYUNDAI	SUSUKI	TOYOTA
KIA	KIA	KIA	NISSAN	TOYOTA	SUSUKI	HYUNDAI	KIA
SUSUKI	NISSAN	NISSAN	TOYOTA	KIA	NISSAN	TOYOTA	KIA
TOYOTA	KIA	KIA	NISSAN	NISSAN	KIA	SUSUKI	SUSUKI
NISSAN	KIA	TOYOTA	KIA	TOYOTA	KIA	SUSUKI	NISSAN
SUSUKI	KIA	HYUNDAI	HYUNDAI	TOYOTA	TOYOTA	HYUNDAI	SUSUKI
NISSAN	KIA	TOYOTA	SUSUKI	SUSUKI	NISSAN	TOYOTA	NISSAN

Solución:



Tabla No.4
Marca de vehículos que prefieren los trabajadores de la Universidad
Continental. Huancayo-2013

Marca de vehículos	Frecuencia Absoluta f_i	Frecuencia Absoluta Acumulada F_i	Frecuencia Relativa H_i	Frecuencia Relativa Acumulada H_i	Frecuencia Porcentual $pi\%$	Frecuencia Porcentual Acumulada $Pi\%$
Hyundai						
Susuki						
Kia						
Nissan						
Toyota						

Fuente: Investigación UC

2.-El Departamento de Personal de la empresa PLAZA VEA ha entregado a la Gerencia el récord de tardanzas de los empleados del personal técnico (medido en minutos) correspondiente al mes de febrero. Los resultados se muestran en la tabla adjunta. Elabore la tabla de frecuencias correspondiente, de acuerdo al tipo de variable. Elabora su grafico e interpreta: f_3 y el $P_5\%$

60	32	85	52	65	77	84	65	57	74
71	81	35	50	35	64	74	47	68	54
80	41	61	91	55	73	59	53	45	77

Solución:

1.- $R = \underline{\hspace{2cm}}$

2.- $m = 1 + 3.322 \log(30) = \underline{\hspace{2cm}}$

3.- $A = \underline{\hspace{2cm}}$

TARDANZAS (en minutos)	X_i	f_i	F_i	H_i	H_i	$pi\%$	$Pi\%$



3.- El número de hijos de los trabajadores de la Universidad Continental al 2015 son los siguientes:

3	0	4	2	1
3	1	4	5	3
6	2	1	0	2
1	1	3	3	1

Elabore la tabla de frecuencias de acuerdo a la variable.

TABLA No.
Número de hijos de los trabajadores de la Universidad Continental
Huancayo-2015

NUMERO DE HIJOS	FRECUENCIA ABSOLUTA f_i	FRECUENCIA ACUMULADA F_i	Frec. Relativa h_i	Frec. Relativa acumulada H_i	Frec. Porcentual p_i	Frec. Porcentual acumulada P_i

Fuente: Departamento de personal de la Universidad Continental.

EJERCICIOS PROPUESTOS:

1.-En la Universidad Continental se ha realizado una encuesta a 200 alumnos sobre el tipo de atención de esta institución. El 32% afirma que está muy contento, el 40% está contento, el 23% no está contento, y el resto muy descontento. Elabore la tabla de frecuencias e interprete.

2.-Se ha llevado a cabo una encuesta a 27 empresas sobre el número de microcomputadoras que tienen, encontrando los siguientes resultados:

5	7	9	7	8	5	2	4	3
6	8	7	6	9	8	4	6	4
8	5	9	6	7	9	4	7	5

Elabore la Tabla de frecuencias correspondiente y analice la información.

3.-Se ha llevado a cabo un estudio para evaluar los volúmenes de venta (miles de soles por día) de 24 establecimientos comerciales de Huancayo y se encontraron los siguientes resultados. Elabore la tabla de frecuencias e interprete.

11,7	5,7	10,1	8,5	6,4	2,1
9,1	3,7	5,3	7,8	4,4	9,8
7,4	12,1	5,4	7,4	3,2	1,5
8,4	6,1	5,7	4,7	5,2	4,6



4.-En la fábrica de SAZON LOPESA se hizo un estudio sobre el peso (kg) de los trabajadores con el fin de establecer una orientación sobre nutrición y buena salud. Los resultados fueron los siguientes: Elabore la tabla de frecuencias e interprete.

60	84	112	120	72	61
70	74	68	90,5	81	75
84	64,2	118	84	96,4	65
84	65	97,5	82	98	62

5.- Los directivos de "Real Plaza" realizan una prueba de mercado respecto a la facilidad de navegación en su nuevo sitio web. Selecciona al azar 18 usuarios frecuentes y les solicita que califique la relativa facilidad para navegar como mala (M), buena (B), excelente (E) o sobresaliente (S)". Los resultados son los siguientes, elabore la tabla de frecuencias e interprete.

B	E	S	B	M	S
M	S	E	S	B	E
M	M	B	M	E	S

6.-Para un estudio de accesibilidad, durante 30 días anotamos el número de plazas libres de aparcamiento a las 5 de la tarde. Elabore la tabla de frecuencias e interprete.

1	1	5	0	5	3	0	3	3	2
2	3	1	1	2	1	2	0	1	3
2	1	5	0	2	2	1	3	3	2

7.-Las calificaciones finales del curso de ESTADISTICA donde las notas están sobre 10 son los siguientes. Elabore la tabla de frecuencias e interprete.

4,5	8,0	8,5	7,5	6,5	3,5	6,0
4,5	4,5	8,5	8,5	10,0	7,0	6,5
9,5	7,0	6,0	8,5	6,5	6,5	8,5

8.- Se realizó una encuesta a los trabajadores de la casa de préstamos "Perú Cash", sobre el número de hijos. Elabore la tabla de frecuencias e interprete.

2	1	2	4	1	3	2
3	2	3	2	0	3	4
3	2	1	3	2	1	2

9.-Redondear:

En los siguientes valores, redondea a los decimales indicados.

- 1,0175 Redondea a los 3 decimales _____
- 0,133456 Redondea a los 5 decimales _____
- 2,012501 Redondea a los 3 decimales _____
- 6,0244 Redondea a los 3 decimales _____

Referencias bibliográficas consultadas y/o enlaces recomendados

- Toma, J. y Rubio, J. Estadística Aplicada primera parte. Universidad del Pacífico. Centro de Investigación, 2011.ISBN:978-9972-57-109-1
- Triola F. (2009) Estadística Elemental. Editorial Mexicana. Décima Edición.



Semana N°3: DISTRIBUCIÓN DE FRECUENCIAS BIDIMENSIONALES, RECONSTRUCCIÓN DE TABLAS

1. TEMA: Distribuciones de Frecuencias Bidimensionales. Reconstrucción de tablas

2. PROPÓSITO:

- Organiza y compara dos variables cualitativas y dos variables cuantitativas, elabora su respectiva grafica e interpreta los resultados.

DISTRIBUCIONES BIDIMENSIONALES

Es la elaboración de tablas de frecuencia en función de dos variables pareadas.

1. Datos bivariados:

Valores de dos diferentes variables que se obtienen a partir del mismo elemento de población.

Cada una de las dos variables pueden ser o cualitativas o cuantitativas. Como resultado, los datos bivariados pueden formar tres combinaciones de tipos de variable:

- Ambas variables son cualitativas (atributos): Es frecuente que los datos se ordenen en una tabulación cruzada o tabla de contingencia. Los resultados se presentan en un gráfico de barras agrupadas.
- Una variable es cualitativa (atributo) y la otra es cuantitativa (numérico): Los valores cuantitativos se ven como muestras separadas, con cada conjunto identificado por niveles de la variable cualitativa. Sus resultados pueden mostrarse en un diagrama de puntos o en un diagrama de cajas y bigotes con una escala común.
- Ambas variables son cuantitativas (ambas numéricas): Se acostumbra expresar matemáticamente los datos como pares ordenados (x,y) , donde "x" es la variable de entrada (variable independiente) y "y" es la variable de salida (variable dependiente). Se llaman "emparejados" o "apareados" porque para cada valor de "x" siempre hay un valor correspondiente de "y" de la misma fuente. Sus resultados se presentan en un diagrama de dispersión.

2. Distribuciones absolutas marginales:

Dada una distribución de frecuencias bidimensionales, podemos obtener dos distribuciones de frecuencias absolutas marginales, una con respecto a la variable "x" y la otra respecto a la variable "y".

EJERCICIOS RESUELTOS:

- Un grupo de estudiantes de la facultad de Ingeniería de la Universidad Continental están a punto de iniciar sus prácticas pre profesionales en diferentes regiones de nuestro país (costa, sierra, selva). Se ha encuestado a un grupo de ellos para conocer su género y la región elegida para llevar a cabo dichas prácticas. Los datos se muestran a continuación:

GÉNERO	REGIÓN	GÉNERO	REGIÓN	GÉNERO	REGIÓN
M	Sierra	F	Selva	F	Selva
F	Selva	M	Sierra	M	Selva
M	Sierra	M	Sierra	F	Selva
M	Costa	M	Costa	F	Selva
F	Selva	F	Costa	M	Costa
M	Selva	M	Selva	M	Sierra
F	Costa	M	Selva	F	Selva
F	Selva	M	Sierra	M	Sierra

Organiza los datos en una tabla de contingencia. Luego elabore la distribución de frecuencias marginales y construya su gráfico de barras agrupadas.

Solución:

Primero.- Construimos la tabla considerando la variable "Género" en las filas y "Región" en las columnas:



Región	Costa	Sierra	Selva	TOTAL
Masculino				
Femenino				
TOTAL				

Segundo.- Completamos cada celda de la tabla con el número de veces que aparece cada dato bivariado. Para ello contamos la cantidad de estudiantes de género masculino que viajarán a la costa, que viajarán a la sierra y a la selva. Hacemos lo propio con las estudiantes de género femenino. Luego sumamos para calcular los totales de fila y columna.

Región	Costa	Sierra	Selva	TOTAL
Masculino	3 <i>12.50%</i>	7 <i>29.17%</i>	4 <i>16.67%</i>	14 <i>58.33%</i>
Femenino	2 <i>8.33%</i>	0 <i>0.00%</i>	8 <i>33.33%</i>	10 <i>41.67%</i>
TOTAL	5 <i>20.83%</i>	7 <i>29.17%</i>	12 <i>50.00%</i>	24 <i>100.00%</i>

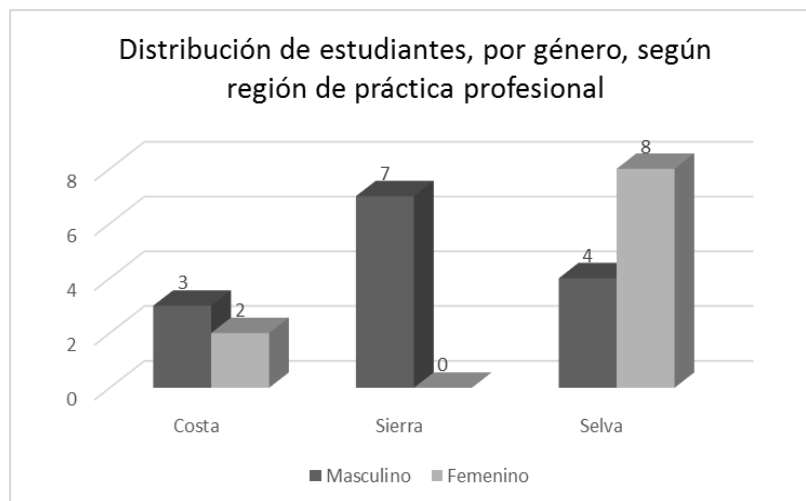
$$\left(\frac{4}{24} \right) \times 100 = 16,67\%$$

Tercero.- Elaboramos las tablas de frecuencias marginales:

GÉNERO	REGIÓN
TOTAL	TOTAL
Masculino	Costa
Femenino	Sierra
TOTAL	Selva
TOTAL	TOTAL

14	5
<i>58.33%</i>	20,83%
10	7
<i>41.67%</i>	29,17%
24	12
<i>100,00%</i>	50,00%
	24
	100,00%

Cuarto.- Se elabora el gráfico de barras agrupadas.



EJERCICIOS PROPUESTOS:

1. La empresa PLAZA VEA como parte de su política de prestaciones de salud a sus colaboradores, ha iniciado una campaña de prevención del cáncer pulmonar entre fumadores. Para ello aplicó una



encuesta en la que se preguntó, entre otras cosas, el género y si es fumador. Las respuestas se muestran a continuación:

GÉNERO	FUMADOR	GÉNERO	FUMADOR	GÉNERO	FUMADOR
M	SÍ	F	SÍ	F	SÍ
F	NO	M	NO	M	SÍ
F	NO	M	NO	F	NO
F	SÍ	F	SÍ	M	NO
M	SÍ	M	NO	F	NO

- 1.1. Elabore una tabla de contingencia para organizar los datos, considerando frecuencias absolutas y relativas conjuntas.
- 1.2. Elabore la distribución de frecuencias marginales para las variables "Género" y "Fumador".
- 1.3. Construya el gráfico de barras agrupadas para presentar dicha tabla de contingencia.

2. Treinta estudiantes de nuestra universidad se identificaron y clasificaron al azar según dos variables: Género (M/F) y Especialidad: Ingeniería (I), Administración (A), Derecho (D)

GENERO	ESPECIALIDAD	GENERO	ESPECIALIDAD	GENERO	ESPECIALIDAD
F	I	F	A	F	I
F	A	F	I	F	A
F	A	M	D	M	I
F	D	M	A	M	A
M	I	M	I	F	I

Elabore una tabla con tabulación cruzada o tabla de contingencia expresándola en porcentajes del gran total, y conteste las siguientes preguntas:

- a) ¿Qué porcentajes de estudiantes estudian derecho?
- b) ¿Cuántos estudiantes estudian administración?
- c) ¿Cuántos estudiantes estudian ingeniería y son del género femenino?
- d) ¿Qué porcentaje de estudiantes estudian administración y son del género masculino?

3. Se está estudiando la relación que existe entre el grado de instrucción y el número de hijos que tienen las mujeres de Huancayo. Para ello se ha entrevistado a un grupo de pobladoras y los resultados se muestran a continuación:

GDO. INST.	Nº HIJOS	GDO. INST.	Nº HIJOS	GDO. INST.	Nº HIJOS
SIN INSTRUCCIÓN	6	SUPERIOR	1	PRIMARIA	3
SECUNDARIA	4	SECUNDARIA	2	SUPERIOR	1
SECUNDARIA	3	SECUNDARIA	2	PRIMARIA	2
SUPERIOR	2	SECUNDARIA	3	SUPERIOR	2
SIN INSTRUCCIÓN	5	SUPERIOR	2	PRIMARIA	3
SUPERIOR	1	SUPERIOR	2	SECUNDARIA	3

- 3.1. Construya una tabla de contingencia para organizar dicha información, mostrando frecuencias absolutas y relativas conjuntas.
- 3.2. Elabore la distribución de frecuencias marginales para cada variable.
- 3.3. Elabore el gráfico de barras agrupadas que represente a la tabla creada.
- 3.4. ¿Qué conclusión puede usted obtener al analizar la información ahora que está organizada?

4. Se han aplicado 3 métodos diferentes (Métodos: A, B y C) para la enseñanza de Análisis Matemático en la facultad de Ciencias de la Empresa, luego de lo cual se aplicó una prueba para medir el tiempo (en minutos) que los alumnos empleaban en resolver un conjunto de 20 ejercicios, siendo los resultados los siguientes:



Método	A	B	B	A	C	B	B	A	C	B	B	C	B	A	B
Tiempo	15	8	10	18	15	11	9	10	11	8	10	10	12	15	8

Método	B	A	C	C	B	C	A	A	B	C	A	B	B	A	C
Tiempo	11	14	10	11	11	10	10	15	9	14	17	9	10	11	12

- 4.1. Construya una tabla de contingencia que organice los datos.
 - 4.2. Elabore la distribución de frecuencias marginales para cada variable en estudio.
 - 4.3. ¿Qué conclusión puede usted obtener al analizar la tabla, respecto a la eficacia de los métodos aplicados?
5. En una muestra de 50 viviendas familiares de la ciudad de Huancayo, se considera como primera variable (X) el número de personas por vivienda, y como segunda variable (Y) el número de ambientes que ocupan por cada vivienda. Los valores observados fueron:

X	6	4	6	5	5	6	5	6	8	7	4	7
Y	2	2	3	4	3	3	3	3	4	4	5	3

X	8	4	4	6	5	5	4	9	6	5	8	5
Y	5	2	4	3	3	3	4	5	2	2	6	2

Ordenar la información proporcionada en un cuadro de distribución de frecuencias e indique: Las frecuencias absolutas y relativas conjuntas, las frecuencias marginales. Luego elabore el diagrama de dispersión que represente dicha información.

6. Se ha observado detenidamente el número de horas que dedican al estudio un grupo de estudiantes de Ingeniería de la Universidad Continental y sus calificaciones en el curso de Matemática. Los datos se muestran a continuación. Ordenar la información proporcionada en un cuadro de distribución de frecuencias e indique: Las frecuencias absolutas y relativas conjuntas, las frecuencias marginales. Finalmente elabore el diagrama de dispersión correspondiente.

HORAS DE ESTUDIO	3	2	4	4	3	5	2	6	7	1	3	6	4	5	2
CALIFICACIÓN	12	9	16	15	12	16	10	18	19	8	13	17	15	15	9

HORAS DE ESTUDIO	1	4	5	1	1	2	5	4	3	3	7	5	6	6	2
CALIFICACIÓN	6	16	17	5	5	9	14	13	13	14	18	15	18	16	10

7. Un comisario de la policía en Huancayo, clasifica los delitos por edad (en años) del malhechor y si el delito cometido fue con violencia o no. Según se muestra a continuación, al comisario se le informó de un total de 150 delitos cometidos durante el pasado año, parte de la información se ha perdido. Se pide completar la tabla. De acuerdo a ello conteste:
- | Tipo de Delito | Edad en años | | | Total |
|----------------|--------------|---------|----------|-------|
| | Menos de 20 | 20 a 39 | 40 o más | |
| Con violencia | | 41 | 14 | 82 |
| Sin violencia | 12 | 34 | | |
| Total | | | | |
- a) ¿Qué porcentaje de los crímenes se realizaron con violencia?
 - b) ¿Qué porcentaje de los crímenes se realizaron con violencia y lo cometieron los malhechores menores de 20 años?

REFERENCIAS BIBLIOGRÁFICAS Y ENLACES

- Kuby, Johnson. ESTADÍSTICA ELEMENTAL. Cengage Learning. México, 2012.

Semana N°4: GRÁFICA ESTADÍSTICOS

1. **TEMA: Barras, histograma, polígono, ojiva, grafico de puntos, hojas y tallos, gráfico de Pareto, circulares, dispersión, pirámide de población y caja y bigotes.**

2. **PROPÓSITO:**

- Representa e interpreta gráficos estadísticos.

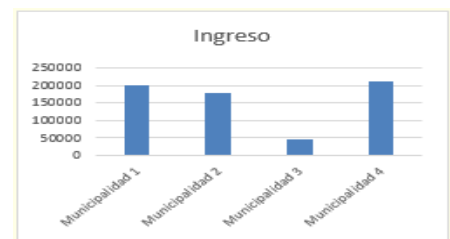
1. GRÁFICO ESTADÍSTICO

Un gráfico estadístico es una representación visual de una serie de datos estadísticos. Es una herramienta muy eficaz, ya que un buen gráfico:

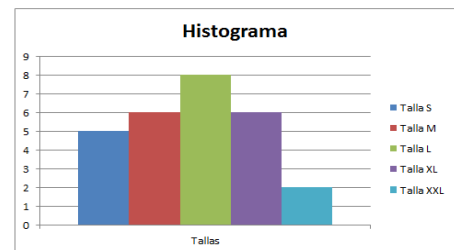
- Capta la atención del lector, ilustrando el mensaje o tema que acompaña.
- Presenta la información de forma rápida, sencilla, clara y precisa. Facilita la comparación de datos.
- Destaca las tendencias y el comportamiento de los datos.

2. TIPOS DE GRÁFICA DE DATOS:

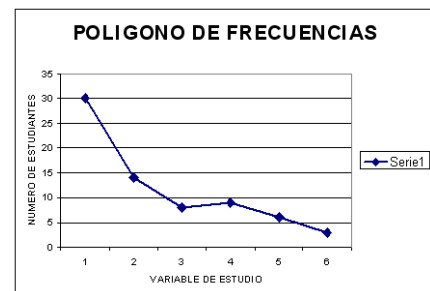
- **GRÁFICO DE BARRAS:** Cada barra rectangular corresponde a una modalidad, tiene una base constante, y su altura puede ser medida en unidades de frecuencia relativa, absoluta o porcentual.



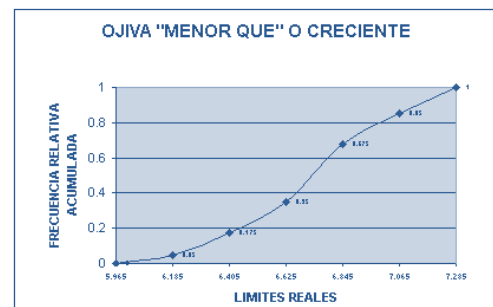
- **HISTOGRAMA:** Es una gráfica de barras donde la escala horizontal representa clases de valores de datos y la escala vertical representa frecuencias. Las alturas de las barras corresponden a los valores de frecuencia y no existe separación entre las barras.



- **POLÍGONO DE FRECUENCIAS:** Utiliza segmentos lineales conectados a puntos que se localizan directamente por encima de los valores de las marcas de clase. Las alturas de los puntos corresponden a las frecuencias de clase; en tanto que los segmentos lineales se extienden hacia la derecha y hacia la izquierda, de manera que la gráfica inicia y termina sobre el eje horizontal.

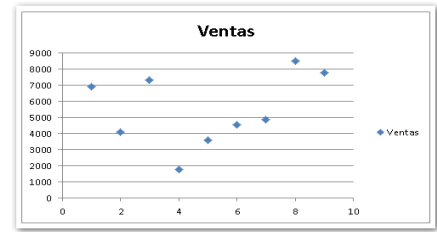


- **OJIVA:** Es una gráfica lineal que representa frecuencias acumulativas. La ojiva utiliza fronteras de clase a lo largo de la escala horizontal, y que la gráfica comienza con la frontera inferior de la primera clase y termina con la frontera superior de la última clase. Las ojivas son útiles para determinar el número de valores que se encuentran por debajo de un valor específico.





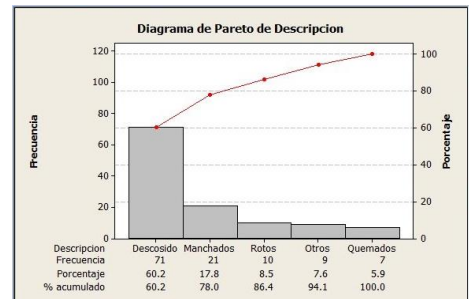
- **GRÁFICA DE PUNTOS:** Es aquella donde se marca cada valor de un dato como un punto a lo largo de una escala de valores. Los puntos que representan valores iguales se apilan.



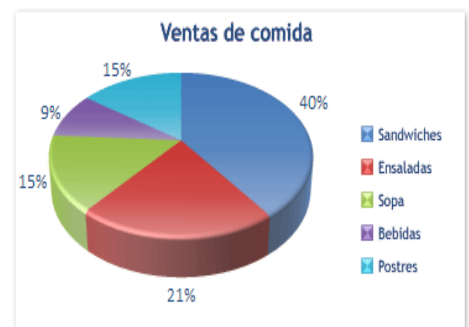
- **GRÁFICA DE TALLO Y HOJAS:** Representa datos que separan cada valor en dos partes: el tallo (el dígito ubicado en el extremo izquierdo) y la hoja (el dígito del extremo derecho). Las hojas se ordenan de forma creciente y no en el orden en que aparecen en la lista original. Una gran ventaja de la gráfica de tallo y hojas radica en que nos permite ver la distribución de los datos y, al mismo tiempo, retener toda la información de la lista original. Otra ventaja es que la construcción de una gráfica de tallo y hojas implica una forma fácil y rápida de *ordenar* datos (acomodarlos en orden), y algunos procedimientos estadísticos requieren de un ordenamiento (como el cálculo de la mediana o de los percentiles).

Tallo (decenas)	Hojas (unidades)
2	12445555666677778888999999
3	0011122333334444555555677888899
4	01111223569
5	04 ← Los valores son 50 y 54
6	013
7	4
8	0 ← El valor es 80.

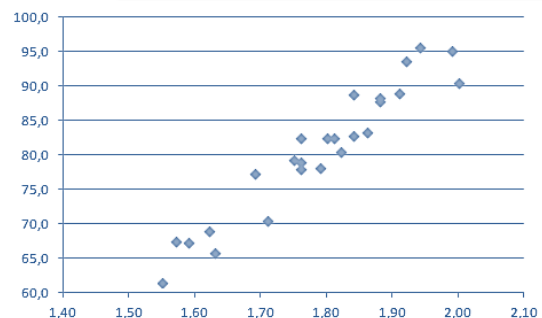
- **GRÁFICA DE PARETO:** Es una gráfica de barras para datos cualitativos, donde las barras se ordenan de acuerdo con las frecuencias. Las escalas verticales de las gráficas de Pareto representan tanto frecuencias absolutas como frecuencias relativas. La barra más alta se coloca a la izquierda y las más pequeñas a la derecha. Al ordenar las barras por frecuencias, esta gráfica enfoca la atención en las categorías más importantes.



- **LAS GRÁFICAS CIRCULARES:** También se utilizan para visualizar datos cualitativos. Presenta datos cualitativos como si fueran rebanadas de un pastel. Para construir una gráfica circular, se divide el círculo en las proporciones adecuadas. Cada sector corresponde a una modalidad y su correspondiente ángulo en el centro.

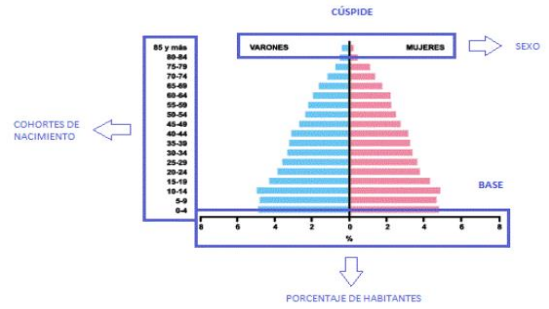


- **DIAGRAMAS DE DISPERSIÓN:** es una gráfica de datos apareados (x, y) , con un eje x horizontal y un eje y vertical. Los datos se aparean de tal forma que cada valor de un conjunto de datos corresponde a un valor de un segundo conjunto de datos. Para elaborar manualmente un diagrama de dispersión, construya un eje horizontal para los valores de la primera variable, construya un eje vertical para los valores de la segunda variable y después grafique los puntos. El patrón de los puntos graficados suele ser útil para determinar si existe alguna relación entre las dos variables.

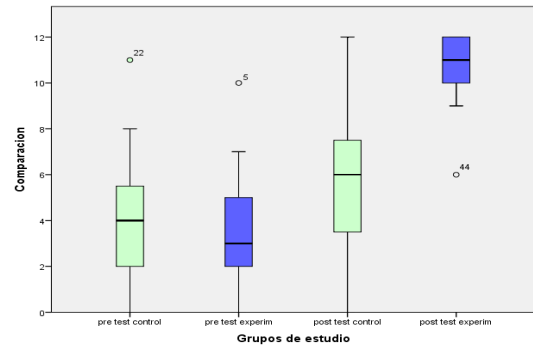




- PIRAMIDE DE POBLACION:** Es una representación gráfica de la distribución por edad y sexo de una población en un momento determinado. Nos pueden brindar información sobre migración de la población, mortalidad, guerras, epidemias y muchas otras situaciones que se presentan en una población. Además, que nos ayuda a comparar los resultados de diversos fenómenos.



- CAJA Y BIGOTES:** Son una representación visual que describe varias características importantes, al mismo tiempo, tales como la dispersión y simetría. Para su realización se representan **los tres cuartiles** y los valores **mínimo** y **máximo** de los datos, sobre un rectángulo, alineado horizontal o verticalmente.

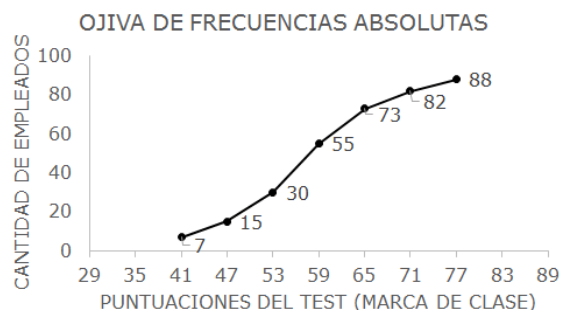
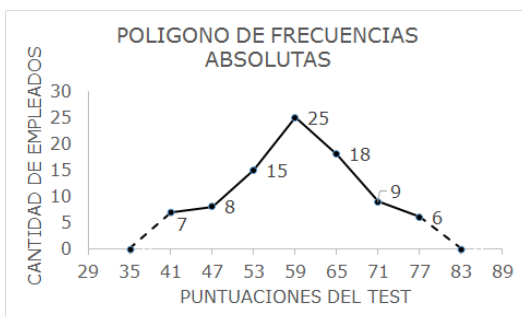


EJERCICIOS RESUELTOS:

- La empresa "La Grande" registra las horas extras de los colaboradores en un año determinado, obteniendo la siguiente tabla:

Intervalos	Xi	fi	Fi	Hi	Hi	Pi	Pi
[38-44>	41	7	7	0.0795	0.0795	7.95%	7.95%
[44-50>	47	8	15	0.0909	0.1705	9.09%	17.05%
[50-56>	53	15	30	0.1705	0.3409	17.05%	34.09%
[56-62>	59	25	55	0.2841	0.6250	28.41%	62.50%
[62-68>	65	18	73	0.2045	0.8295	20.45%	82.95%
[68-74>	71	9	82	0.1023	0.9318	10.23%	93.18%
[74-80]	77	6	88	0.0632	1	6.82%	100.00%
TOTAL		88		1		100.00%	

Se pide elaborar: polígono de frecuencia y ojiva, e interpreta **Solución:**



INTERPRETACION:

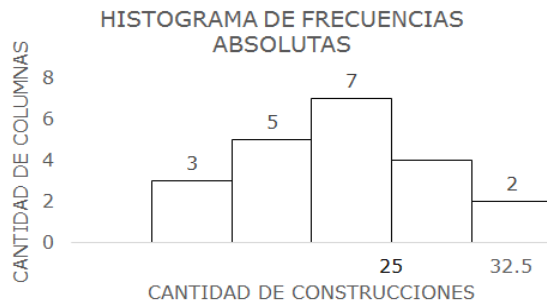
De los 88 colaboradores de la empresa "La Grande", 25 trabajan de 56 a menos de 62 horas haciendo el porcentaje en un 28,41%.

- Determine la tabla de distribución de frecuencia del histograma que se muestra, donde se observa la cantidad de columnas que tienen 21 construcciones:



Solución:

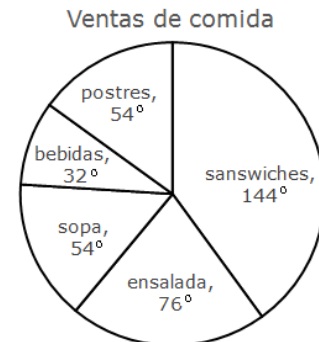
Intervalos	Xi	Fi	Fi	hi	Hi	pi%	Pi%
[10-15>	12.5	3	3	0.1429	0.1429	14.29	14.29
[15-20>	17.5	5	8	0.2381	0.3810	23.81	38.10
[20-25>	22.5	7	15	0.3333	0.7143	33.33	71.43
[25-30>	27.5	4	19	0.1905	0.9048	19.05	90.48
[30-35]	32.5	2	21	0.0952	1	9.52	100.00
TOTAL		21		1		100.00	



3. A partir del grafico que se muestra elabore su tabla de distribución de frecuencia, donde se muestran el consumo de 300 comensales de "Rustica":

Solución:

Ventas de comida	Fi	Fi	hi	Hi	pi%	Pi%	ángulo
sándwiches	120	120	0.4	0.4	40	40	144
ensalada	63	183	0.21	0.61	21	61	76
sopa	45	228	0.15	0.76	15	76	54
bebidas	27	255	0.09	0.85	9	85	32
postres	45	300	0.15	1	15	100	54
TOTAL	300		1		100		360



EJERCICIOS PROPUESTOS:

1. El siguiente cuadro muestra el total de inasistencia de los alumnos del mes, de tres facultades distintas. Elabora un gráfico de barras. (elige otro gráfico que te parezca conveniente)

MESES	FACULTADES		
	INGENIERIA	DERECHO	MEDICINA
JULIO	30	24	20
AGOSTO	35	30	38
SEPTIEMBRE	19	25	25
OCTUBRE	20	19	27
NOVIEMBRE	15	20	32
DICIEMBRE	18	22	38



2. En una encuesta sobre el tipo de información que se leía en las secciones de un periódico, se obtuvieron los datos siguientes: deporte 75; cultural y viajes 125; economía 94; política 56; noticias locales 114; noticias internacionales 45 y farándula 220. Elabora con los datos una tabla y su gráfico respectivo.

3. La Cámara de Comercio de Huancayo está interesada en conocer de qué manera vienen desarrollando sus actividades los restaurantes y las pollerías del centro de la ciudad. Para ello han tomado una muestra de 50 de estos establecimientos y ha revisado el libro de reclamaciones de cada uno para contabilizar el número de quejas que presentaron los clientes. Los datos se muestran a continuación:

Número de quejas	f_i	F_i
1	2	2
2	4	6
3	21	27
4	15	42
5	6	48
6	1	49
7	1	50
TOTAL	50	

Construya el gráfico estadístico correspondiente tanto para las frecuencias absolutas (diagrama de bastones) y las acumuladas (diagrama de escalones).

4. La empresa ELECTROCENTRO S.A. está llevando a cabo un estudio minucioso acerca de los salarios que perciben los obreros de esta institución, con la finalidad de realizar mejoras económicas entre su personal. La siguiente tabla muestra los salarios que perciben una muestra de 26 de estos obreros:

Salarios S/.	[750-900>	[900-1050>	[1050-1200>	[1200-1350>	[1350-1500>	[1500-1650>	[1650-1800>
Cantidad de obreros	2	4	6	7	3	3	1

Se pide que grafique el histograma, polígono de frecuencia y ojiva "Menor que" de dicha tabla.

5. El peso en gramos de 30 objetos de un mismo tipo fue como sigue:

21,3	15,8	18,4	22,7	19,6	15,8	26,4	17,3	11,2	23,9
26,8	22,7	18,0	20,5	11,0	18,5	23,0	24,6	20,1	16,2
08,3	21,9	12,3	22,3	13,4	17,9	12,2	13,4	15,1	19,1

Construir un diagrama de tallo y hojas para los datos indicados, indicar las características de la distribución.

6. En la siguiente tabla se muestran los resultados después de las evaluaciones a un grupo de estudiantes de la carrera profesional de contabilidad. (utilizar dos tipos de gráficos)

CONDICION	ESTUDIANTES (GENERO)	
	VARONES	MUJERES
APROBADO	65	96
DESAPROBADO	25	32
RETIRADO	10	8

7. Dado el siguiente gráfico de sectores referente al presupuesto mensual de un trabajador. Si gana mensualmente s/. 1800, ¿cuánto gasta en educación y en otras actividades?

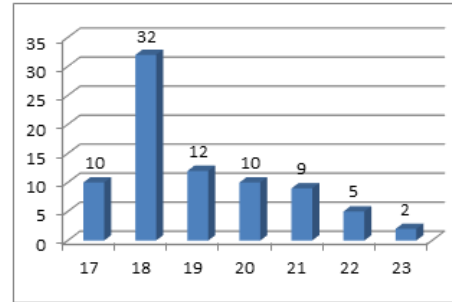




8. El gráfico muestra las edades de los estudiantes ingresantes a la carrera de INGENIERA CIVIL de la Universidad Continental para el año 2015-2.

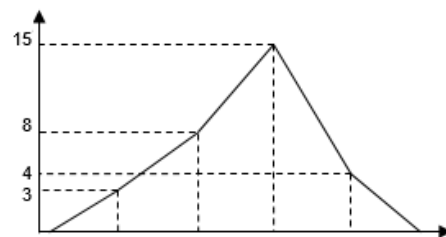
Marcar como Verdadero o Falso.

- El 40 % de los ingresantes tienen la edad de 18 años. ()
- La muestra corresponde a 80 estudiantes ingresantes de la carrera de Ingeniería Civil. ()
- En un 90% de los estudiantes son mayores de edad. ()
- El diagrama de barras muestra la variable edad y la frecuencia relativa. ()

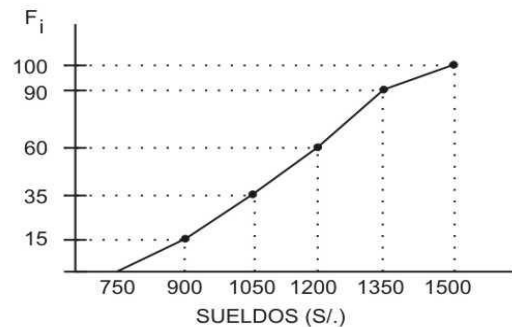


9. El polígono de frecuencias muestra el peso de un grupo de niños de 9 años de edad, quienes son los hijos menores de los empleados de la empresa de giros y encomiendas CARGO 1. Reconstruya la tabla de frecuencias a partir de la información mostrada en el gráfico.

Polígono de frecuencias de los pesos de los niños



10. Como parte de un informe que deberán presentar al Ministerio de Trabajo, se ha tomado los datos referentes a los sueldos mensuales de una muestra de empleados de la Municipalidad Distrital de Huancayo. Los datos se muestran en la siguiente ojiva. Se pide reconstruir la tabla y contestar: ¿Qué porcentaje de la muestra representan los empleados que perciben de S/. 900 a menos de S/. 1200?



11. El ingeniero de control de calidad de la fábrica de aluminios G&A ha recibido constantes quejas sobre las fallas que presentan dichas láminas. Por ello le ha encargado a usted investigar cuáles son las fallas a las que hay que prestar mayor atención para mejorar la calidad de dichas láminas de aluminio. Construya el diagrama de Pareto.

TIPO DE FALLA LA LÁMINA DE ALUMINIO	TOTAL
Grietas	34
No cumple con la longitud establecida	16
Otros	5
Rugosidad	12
Deformaciones	3

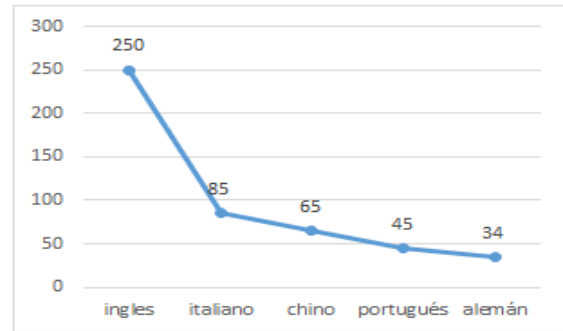
12. Los sistemas de cómputo fallan por muchas razones, entre ellas las fallas de hardware o software, errores del operador, sobrecargas del sistema mismo y a otras causas. La tabla siguiente muestra los resultados obtenidos en un estudio acerca de las causas de fallas en una muestra de 98 sistemas de cómputo. Usted debe priorizar entre las dos principales

TIPO DE FALLA	FRECUENCIA
Hardware	9
Operador	20
Sobrecarga	55
Software	8
Otras	6



causas de falla de los sistemas de cómputo. Elabore el gráfico apropiado que permita visualizar dicho propósito.

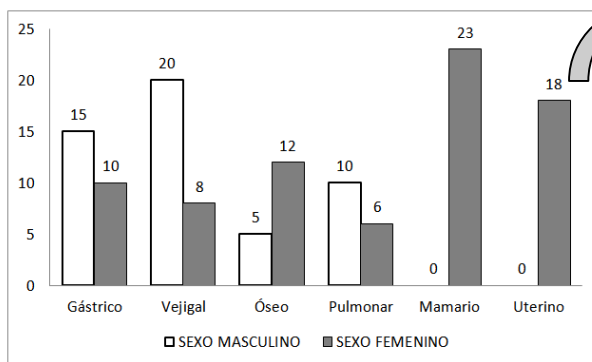
13. Se ha tomado el número de estudiantes que egresaron del centro de idiomas en el año 2016, representado en un polígono de frecuencias. Reconstruya la tabla de frecuencias de dicho gráfico.



14. A continuación se presentan las edades (en años) de un grupo de varones y mujeres, pacientes de la Clínica Dental IMAGEN. Recupere los datos originales y elabore su tabla de frecuencias, agrupando en intervalos. Luego grafique el histograma, polígono de frecuencias y ojiva de dichas edades.

HOJAS (VARÓN)	TALLOS	HOJAS (MUJER)
732	4	67
36	5	2
2683228	6	6679
461	7	5183752
0	8	2

15. Complete la tabla de contingencia a partir de la información del gráfico de barras agrupadas.



TIPO DE CÁNCER	SEXO		TOTAL
	MASCULINO	FEMENINO	
TOTAL			

REFERENCIAS BIBLIOGRÁFICAS Y ENLACES

- Triola Mario F. ESTADÍSTICA. Pearson Educación. México 2009
- Díaza Mata, Alfredo. ESTADÍSTICA APLICADA A LA ADMINISTRACIÓN Y LA ECONOMÍA. Mc. Graw Hill. México 2012
- Martínez C. (2014). ESTADÍSTICA BASICA APLICADA. Colombia: Editorial ECOE.



GUÍA DE PRÁCTICA DE PROBABILIDADES Y ESTADÍSTICA

SEGUNDA UNIDAD ESTADÍSTICOS PARA DESCRIBIR, EXPLORAR Y COMPARAR DATOS

- GUÍA DE PRÁCTICA N° 5: Medidas de tendencia central
- GUÍA DE PRÁCTICA N° 6: Medidas de dispersión
- GUÍA DE PRÁCTICA N° 7: Medidas de posición relativa
- GUÍA DE PRÁCTICA N° 8: Medidas de forma



Semana N°5: MEDIDAS DE TENDENCIA CENTRAL

1. TEMA: Medidas de Tendencia Central: Media, mediana, moda.

2. PROPÓSITO:

- Define las medidas de tendencia Central.
- Calcula e interpreta las medidas de tendencia central.

Existen tres medidas de tendencia central, que son:

1. Media Aritmética: Es el **promedio** de los datos, y su objetivo principal es encontrar el valor que debería estar al centro.

- **Datos No Agrupados:** Se utiliza cuando hay datos sueltos y se halla la suma de dichos valores dividida entre el número de valores, así como se expresa en la fórmula:

Media Aritmética: Datos No Agrupados	
$\bar{x} = \frac{\sum x_i}{n}$ (Muestral)	$\mu = \frac{\sum x_i}{N}$ (Poblacional)

- **Datos Agrupados:** Se utiliza cuando agrupamos datos en una tabla de frecuencias por intervalos o sin intervalos, se halla multiplicando la marca de clase por la frecuencia absoluta de cada categoría (cuando se encuentran en intervalos de clase) y sin intervalos se multiplica el valor de la variable por la frecuencia absoluta, después se debe sumar todos los resultados y dividirlos entre el número total de datos, así como se expresa en la fórmula:

Media Aritmética: Datos Agrupados	
$\bar{x} = \frac{\sum x_i \cdot f_i}{n}$ (Muestral)	$\mu = \frac{\sum x_i \cdot f_i}{N}$ (Poblacional)

• PROPIEDADES DE LA MEDIA ARITMETICA

Propiedad 1: La media aritmética de una constante es igual a la constante.
Demostración.-

$$\bar{x} = x_1, x_2, x_3, \dots, x_n$$

Dados los valores: a, a, a, \dots, a

$$\bar{x} = \frac{a + a + a + \dots + a}{n} = \frac{na}{n} = a$$

Por lo tanto: $\bar{x} = a$

Propiedad 2: La media aritmética de una variable más una constante es igual a la media aritmética de la variable más la constante.

Demostración.-

$$\bar{x} = x_1 + x_2 + x_3 + \dots + x_n$$

$$\bar{y} = x_1 + c + x_2 + c + x_3 + c + \dots + x_n + c$$

Hallamos la media aritmética (con este cambio):



$$\bar{y} = \frac{(x_1+c)+(x_2+c)+\dots+(x_n+c)}{n}$$

Factorizamos:

$$\bar{y} = \frac{x_1+x_2+x_3+\dots+x_n+nc}{n}$$

Que es lo mismo que separarlo en dos grupos:

$$\bar{y} = \frac{x_1+x_2+x_3+\dots+x_n}{n} + \frac{nc}{n}$$

Por lo tanto:

$$= \bar{x} + c$$

Propiedad 3: La media aritmética de una variable por una constante es igual al producto de la constante por la media de la variable

Demostración.-

$$\begin{aligned} \bar{x} &= x_1 + x_2 + x_3 + \dots + x_n \\ \bar{y} &= x_1 \cdot c + x_2 \cdot c + x_3 \cdot c + \dots + x_n \cdot c \end{aligned}$$

Hallamos la media aritmética (con este cambio):

$$\bar{y} = \frac{(x_1 \cdot c) + (x_2 \cdot c) + \dots + (x_n \cdot c)}{n}$$

Factorizamos:

$$\bar{y} = \frac{c(x_1+x_2+x_3+\dots+x_n)}{n}$$

Por lo tanto:

$$= c \cdot \bar{x}$$

EJEMPLOS:

En tres cursos que llevan los alumnos de Ingeniería, los promedios de las calificaciones fueron 5,6; 6,1 y 4,9; si los cursos tenían respectivamente 34; 30 y 36 alumnos, determine la calificación promedio de los tres cursos.

Solución:

$$\bar{x} = \frac{(5,6)34 + (6,1)30 + (4,9)36}{34 + 30 + 36} = \frac{549,8}{100} = 5,498 \approx 5,5$$

RESPUESTA: El promedio de las calificaciones de los tres cursos es 5,5

2. Mediana (Me): Es el valor central, el que limita al 50% de los datos, es decir, el valor que se encuentra en la mitad de los datos.

- **Datos No Agrupados:** Cuando hay un problema sin cuadros o uno en cuadros pero sin intervalos de clase. Los pasos a seguir son:
 1. Ordenar los datos de menor a mayor
 2. Determine la Posición (L):

Mediana: Datos No Agrupados	
$L = \frac{n+1}{2}$ (número impar de datos)	$L = \frac{n}{2}$ y el valor de la siguiente posición (número par de datos)

3. Determine la mediana



- **Datos Agrupados:** Cuando se tienen cuadros con intervalos de clase, se utiliza la fórmula:

$$Me = L_i + \left[A \left(\frac{\frac{n}{2} - F_{j-1}}{f_j} \right) \right]$$

- 3. Moda (Mo):** Es el valor más frecuente, el que se observa mayor número de veces.
- **Datos No Agrupados o agrupados sin intervalos de clase:** Después de ordenar los datos buscamos el valor que más se repite.
 - **Datos Agrupados con intervalos de clase:**
 - En cuadros con intervalos de clase, se localiza la clase modal buscando la frecuencia más alta y después se aplica la siguiente fórmula:

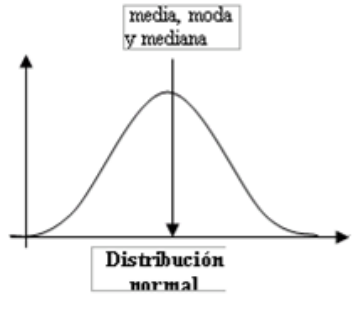
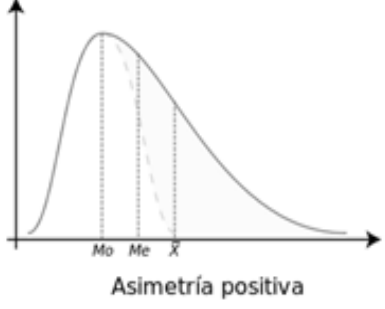
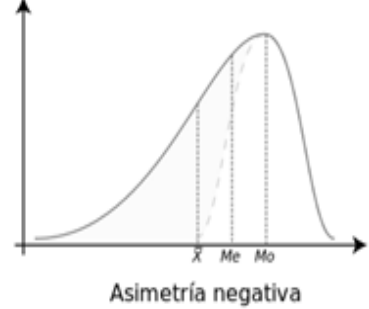
$$Mo = L_i + \left[A \left(\frac{d_1}{d_1 + d_2} \right) \right]$$

Donde:

$$d_1 = f_j - f_{j-1}$$

$$d_2 = f_j - f_{j+1}$$

4. Relación entre Media Aritmética, Mediana y Moda:

Si la distribución de frecuencias de los datos es simétrica , entonces la media, la mediana y la moda tienen el mismo valor:	Si la distribución es asimétrica de cola derecha, entonces, la moda es menor que la mediana y esta a su vez es menor que la media:	Si la distribución es asimétrica de cola a la izquierda, entonces la media es menor que la mediana y esta a su vez menor que la moda:
 <p>media, moda y mediana</p> <p>Distribución normal</p>	 <p>Mo Me X</p> <p>Asimetría positiva</p>	 <p>X Me Mo</p> <p>Asimetría negativa</p>

EJERCICIOS RESUELTOS:

1. Se desea estimar el rendimiento promedio de las llantas de cierta marca. Para ello se toma una muestra de cuatro automóviles a los que se les coloca esta marca de llanta.

Número de Auto	Recorrido (kms)
1°	56 000
2°	42 000
3°	23 000
4°	73 000

Solución:



$$\bar{X} = \frac{(56\,000 + 42\,000 + 23\,000 + 73\,000)}{4} = 48\,500 \text{ Km}$$

-Por tanto, se puede concluir que el rendimiento promedio de las llantas de esta marca (vida útil) es de 48 500 kilómetros.

2. Se desea estimar el número de productos vendidos de MAESTRO en una semana, el cual se muestra en el siguiente cuadro:

Intervalos	x_i	f_i	$x_i \cdot f_i$
[10- 20>	15	1	15
[20- 30>	25	8	200
[30- 40>	35	10	350
[40- 50>	45	9	405
[50- 60>	55	8	440
[60-70>	65	4	260
[70- 80>	75	2	150
		42	$\sum x_i \cdot f_i = 1820$

Entonces:

$$\bar{X} = \frac{1820}{42} = 43,33$$

INTERPRETACION: El número promedio de productos vendidos en MAESTRO en una semana es de 43 productos aproximadamente.

3. Estimar la mediana de los siguientes datos agrupados de la edad de los trabajadores de la empresa Sedam-Huancayo y están divididos por grupos, el cual se muestra en el siguiente cuadro:

Formula:

$$Me = li + A \left(\frac{\frac{n}{2} - F_{i-1}}{f_i} \right)$$

Procedimiento:

$$\text{Posición} = \frac{n}{2} = \frac{100}{2} = 50^a$$

Clase mediana: [66, 69)

Hallamos la amplitud:

Amplitud de clase: 3

Hallamos la mediana:

$$Me = 66 + 3 \left(\frac{50 - 23}{42} \right) = 67,93$$

$$Me = 66 + 3 \left(\frac{50 - 23}{42} \right) = 67,93$$

	f_i	F_i
[60- 63)	5	5
[63- 66)	18	23
[66- 69)	42	65
[69- 72)	27	92
[72- 75)	8	100
	100	



INTERPRETACION:

- Al menos el 50% de los trabajadores de Sedam-Huancayo tienen una edad menor o igual a 68 años
- Al menos el 50% de los estudiantes tienen una edad mayor o igual a 68 años

4. Del ejercicio anterior, estime la moda:

Formula:

$$Mo = L_{mo} + \left(\frac{d_1}{d_1 + d_2} \right) * a$$

	f_i
[60- 63)	5
[63- 66)	18
[66- 69)	42
[69- 72)	27
[72- 75)	8
	100

Procedimiento:

Clase mediana: [66, 69)

Amplitud de clase: 3

Hallando d1 y d2:

$$d1 = 42 - 18 = 24$$

$$d2 = 42 - 27 = 15$$

Hallando la moda:

$$Mo = 66 + \left(\frac{24}{24 + 15} \right) 3 = 67,846$$

INTERPRETACION:

La edad más frecuente de los 100 trabajadores de Sedam-Huancayo es de 68 años.

EJERCICIOS PROPUESTOS:

PROBLEMA 01.

- Una encuesta realizada a un grupo de familias del distrito de El Tambo, sobre el pago mensual por servicio de internet y televisión por cable, arrojó los datos que se muestran a continuación:

Pago de servicios	F_i
[140-220>	100
[220-300>	60
[300-380>	90
[380-460>	120
[460-540]	80
Total	



- a) Calcule la media, mediana y moda.
b) Interprete los resultados.
2. Una empresa ha realizado un test físico entre todos sus empleados para comprobar la capacidad de esfuerzo que posee cada uno de ellos. Una de las medidas que componen el test es el número de pulsaciones después de una determinada actividad física que está altamente relacionada con las que se realizan a lo largo de una jornada laboral. Los datos conseguidos han sido distribuidos en la siguiente tabla:

Numero de pulsaciones	Número de empleados
[65-70>	12
[70-75>	15
[75-80>	10
[80-85>	28
[85-90]	30
[90-95]	5

- a) ¿Qué porcentaje de empleados tuvo menos de 85 pulsaciones?
b) Calcule la media aritmética, la mediana y la moda. Interprete los resultados.

3. Los nacimientos en miles por sexo durante un año han sido registrados en la siguiente tabla:

Mes	E	F	M	A	M	J	J	A	S	O	N	D
H	15	14	16	15	16	16	16	16	16	16	15	15
M	15	13	15	14	15	14	15	15	15	15	14	14

- a) Calcule la media, mediana y moda.
b) Interprete los resultados.

4. Los ingresos en dólares de 18 hombres elegidos al azar del Banco BBVA CONTINENTAL (entre un total de 1000) se muestran a continuación:

45,16	83,61	79,85	22,07	76,91	65,73
88,91	99,49	62,59	34,20	88,61	41,50
68,89	92,22	54,33	53,20	16,60	62,59

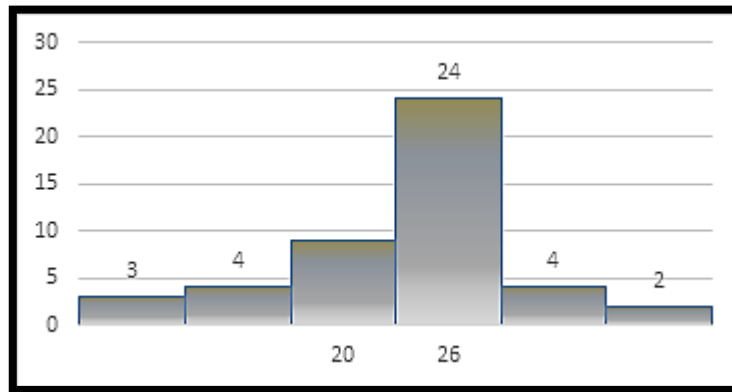
- Calcule la media aritmética empleando la tabla de frecuencias.
- Halle la mediana y moda e interprete (en termino de dólares).
- ¿Se puede considerar que la población de 1000 personas tendrán la misma media que la muestra de 18 personas? ¿Por qué?



5. El gráfico tallo y hoja muestra los productos vendidos de la tienda "La Moderna" en un día. Calcular las medidas de tendencia central y graficar el sesgo y, ¿la media aritmética es significativa en los productos vendidos?

TALLO	HOJAS
2	2
3	122
4	2333
5	228
6	7

6. De los 46 productos vendidos de la tienda "Casa Sueldo" un día domingo. Calcular las medidas de tendencia central, asimetría e interpretar el sesgo. Elabora una tabla de frecuencias a partir del histograma.



7. La tabla muestra las notas de un examen de Estadística de una muestra de 30 estudiantes:

Notas	fi
[05 - 08 >	1
[08 - 11 >	3
[11 - 14 >	
[14 - 17 >	6
[17 - 20 >	5
Total	30

Halle e interprete la media aritmética, la mediana, la moda y evalúe el sesgo (Graficar sesgo).

8. Se tiene el siguiente cuadro que corresponde quesos producidos por trabajador para la empresa SERRANITA:

Litros de lácteos producidos	[5-11>	[11-17>	[17-23>	[23-29>	[29-35>
Cantidad de trabajadores	12	18	13	9	10

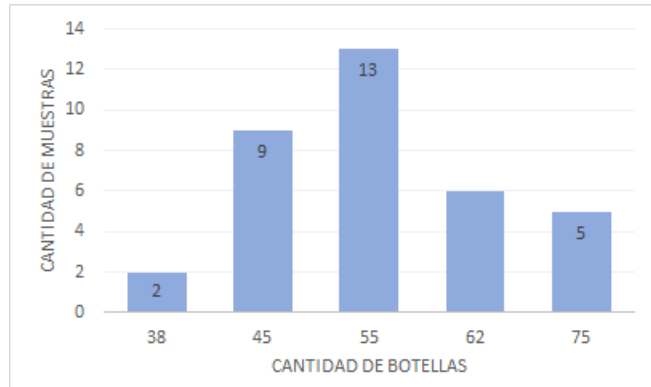
Con los datos se pide:

- Grafica el histograma
- Determine media, mediana y moda e interpreta cada uno de ellos



9. La siguiente grafica nos muestra la cantidad de botellas de vidrio encontradas en el carro recolector de basura de la Municipalidad de Huancayo evaluadas en 35 muestras.

De acuerdo a ello, se pide que realices un análisis en función de las medidas de tendencia central, interprete y determine el tipo de distribución que presentan los datos.



10. Un estudio a una muestra de 12 barras de acero, se realizó para ver la cantidad de cementita (en gr) que tiene cada uno de ellos, siendo los siguientes resultados:

8	7	9	20	18	17
17	20	26	28	28	28

Se pide que determinen sus medidas de tendencia central e interprete los resultados.

REFERENCIAS BIBLIOGRÁFICAS Y ENLACES

- Devore J. (2008) Probabilidad y Estadística para Ingeniería y Ciencias. Editorial Latinoamericana. Séptima Edición.
- Triola F. (2009) Estadística Elemental. Editorial Mexicana. Décima Edición.



Semana N°6: MEDIDAS DE DISPERSIÓN

1. **TEMA: Medidas de variación: Varianza, desviación estándar, coeficiente de variación.**

2. **PROPÓSITO:**

- Calcula e interpreta medidas de variación o dispersión. De datos relacionados a su formación profesional.

MEDIDAS DE DISPERSIÓN

a) **Varianza:** La varianza mide la mayor o menor representatividad de la media aritmética.

Muestral (s^2)	Poblacional (σ^2)
DATOS NO AGRUPADOS	
$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$	$\sigma^2 = \frac{\sum (x_i - \mu)^2}{N}$
DATOS AGRUPADOS	
$s^2 = \frac{\sum (x_i - \bar{x})^2 \times f_i}{n-1}$	$\sigma^2 = \frac{\sum (x_i - \mu)^2 \times f_i}{N}$

Propiedades de la Varianza:

1. $Var(x) = S_x^2 \geq 0$
Es decir, la varianza será siempre un valor positivo o cero, en el caso de que las puntuaciones sean iguales.
2. $Var(x) = 0$ si $x = constante$
La varianza de una constante es cero
3. $Var(ax) = a^2 Var(x)$
Si se tiene la varianza de un conjunto de datos y a cada observación se multiplica por una constante, entonces la nueva varianza de los datos se obtiene multiplicando a la varianza de los datos por a^2
4. $Var(x + b) = Var(x)$
Si a todos los valores de la variable se les suma un mismo número, la varianza no varía.
5. $Var(ax + b) = a^2 Var(x)$
Si todos los valores de la variable se multiplican por un mismo número y luego se le suma otro número, la varianza queda multiplicada por el cuadrado del número multiplicado.
6. Las unidades de medida de la varianza son las unidades al cuadrado de los datos.

b) **Desviación estándar o típica:**

Para eliminar el problema de la elevación al cuadrado de la varianza, se realiza una transformación consistente en calcular la raíz cuadrada de la varianza con lo que obtendríamos la desviación estándar o típica

Muestral (s)	Poblacional (σ)
DATOS NO AGRUPADOS	
$s = \sqrt{s^2}$	$\sigma = \sqrt{\sigma^2}$
DATOS AGRUPADOS	
$s = \sqrt{s^2}$	$\sigma = \sqrt{\sigma^2}$



Con lo que la desviación estándar o típica vendrá dada en las mismas unidades que los valores de la variable.

- La desviación estándar o típica siempre es positiva porque la varianza también lo es.
- La desviación estándar o típica es la medida de dispersión óptima, más exacta, más estable y más utilizada, sirviendo de base para las medidas de asimetría, estadísticos típicas y correlación.
- Cuanto más se acerca a cero la desviación más concentrada es la serie.
- Suele decirse que cuando la desviación estándar o típica es menor que la media aritmética la serie es concentrada y sí la desviación estándar o típica es mayor que la media aritmética la serie es dispersa.

Es la medida de dispersión óptima, más exacta, más estable, y más utilizada. La desviación típica o estándar será siempre un valor positivo o cero, en el caso de que las puntuaciones sean iguales.

Observaciones de la Desviación Estándar:

- La desviación típica o estándar, al igual que la media y la varianza, es un índice muy sensible a las puntuaciones extremas.
- En los casos que no se pueda hallar la media aritmética tampoco será posible hallar la desviación típica.
- Cuanta más pequeña sea la desviación típica mayor será la concentración de datos alrededor de la media.

c) Rango o alcance: Es la medida de variabilidad más sencilla entre todas las mencionadas. El rango de un conjunto de datos es la diferencia entre dato mayor y el dato menor:

$$R = x_{\max} - x_{\min}$$

Donde:

- Xmax : Valor máximo observado de la variable.
- Xmin : Valor mínimo observado de la variable.

Coefficiente de Variación: Es una medida relativa que se usa para comparar la variación en diferentes conjuntos de datos que no vienen dadas en las mismas unidades o que las medias no son iguales, el cual mide la magnitud de la desviación estándar en relación a la media aritmética, expresada como porcentaje:

$$CV = \left(\frac{s^2}{\bar{x}}\right)100\%$$

Valor de CV Interpretación	
0 ≤ CV < 5	Los datos son muy homogéneos
5 ≤ CV < 10	Los datos son homogéneos
10 ≤ CV < 15	Los datos son regularmente homogéneos
15 ≤ CV < 20	Los datos son regularmente heterogéneos
20 ≤ CV < 25	Los datos son heterogéneos
25 ≤ CV	Los datos son muy heterogéneos

Fuente: Estadística descriptiva y probabilidades, Universidad de Lima pag. 59



EJERCICIOS RESUELTOS

1. A continuación se presentan las puntuaciones que registra en minutos el trámite de cierto documento en una institución privada. Calcular la varianza de la distribución:

9	3	8	8	9	8	9	18
---	---	---	---	---	---	---	----

Solución:

-Primero hallamos la media=9

-Usamos la fórmula, donde X_i es cada dato del problema:

$$s^2 = \frac{(9 - 9)^2 + (3 - 9)^2 + (8 - 9)^2 + (8 - 9)^2 + (9 - 9)^2 + (8 - 9)^2 + (9 - 9)^2 + (18 - 9)^2}{7}$$

$$s^2 = \frac{119}{7}$$

Entonces: $s^2 = 17$

2. Un pediatra de la clínica Ortega obtuvo la siguiente tabla sobre los meses de edad de todos los 40 niños de su consulta en el momento de andar por primera vez en Huancayo. Calcular la **desviación estándar**:

$\mu = 12,05$
Luego: $9 - 12,05 = -3,05$

x_i	f_i	$x_i \cdot f_i$	$(x_i - \mu)$	$(x_i - \mu)^2$
9	2	18	-3.05	9.3025
10	4	40	-2.05	4.2025
11	6	66	-1.05	1.1025
12	13	156	-0.05	0.0025
13	9	117	0.95	0.9025
14	5	70	1.95	3.8025
15	1	15	2.95	8.7025
	40	482	-0.35	28.0175

Solución:

-Para resolverlo se debe tener en cuenta que la palabra TODOS_hace referencia al total, es decir, la población por lo que se usara σ

-Luego, hallamos el promedio de los datos como hemos aprendido anteriormente. Así tenemos:



$$\mu = \frac{482}{40} = 12,05$$

- Usamos el promedio para obtener $x_i - \mu$, y después se lleva al cuadrado cada resultado
- Se suma cada fila para obtener un total igual a 28,02
- Finalmente, reemplazamos en la fórmula:

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (x_i - \mu)^2}{N}}$$

$$\sigma^2 = \frac{28,02}{40} = 0,70$$

$$\sigma = \sqrt{0,70} = 0,84$$

3. Del ejercicio anterior, calcular el Coeficiente de variación:

Solución:

-Primero, hallamos σ y μ

$$\sigma = 0,84$$

$$\mu = \frac{482}{40} = 12,05$$

-Luego, usamos la fórmula para hallar el CV:

$$CV = \frac{\sigma}{\mu} \cdot 100\% = \frac{0,84}{12,05} \times 100$$

$$CV = 6,97\%$$

-Finalmente, interpretamos:

Los datos sobre los meses de edad de 40 trabajadores de una empresa son homogéneos porque su CV que es 6,97% se encuentra entre 5% y 10%.

EJERCICIOS PROPUESTOS:

1. Calcular todas las medidas de dispersión para la siguiente distribución:

X_i	5	10	15	20	25
n_i	3	7	5	3	2

2. Calcular todas las medidas de dispersión para los datos de la siguiente distribución:

X	0-100	100-200	200-300	300-800
N	90	140	150	120

3. Un artículo reportó los siguientes datos sobre consumo de oxígeno (ml/kg/min) para una muestra de diez bomberos que realizaron un simulacro de supresión de incendio.

29,5	49,3	30,6	28,2	28,0
26,3	33,9	29,4	23,5	31,6

Calcule lo siguiente:

- El rango muestral.
- La varianza muestral (s^2) a partir de la definición (es decir, calculando primero las desviaciones y luego elevándolas al cuadrado, etcétera).
- La desviación estándar muestral.
- S^2 utilizando el método más corto. (con ayuda de la formula)



4. Se determinó el valor de cinco libros (S/.) en una feria y se obtuvieron las siguientes observaciones muestrales:

116,4	115,9	114,6	115,2	115,8
-------	-------	-------	-------	-------

- a) Calcule la \bar{X}
b) Use la \bar{X} calculada en el inciso a) para obtener la varianza muestral y la desviación estándar muestral.
5. Las observaciones adjuntas de viscosidad estabilizada (cP) realizadas en probetas de un cierto grado de asfalto con 18% de caucho agregado se tomaron de un artículo:

2781	2900	3013	2856	2888
------	------	------	------	------

- a) ¿Cuáles son los valores de la media y mediana muestrales?
b) Calcule la varianza muestral por medio de la fórmula de cálculo.
6. Calcule e interprete los valores de la mediana muestral, la media muestral, la desviación estándar muestral y varianza de los siguientes datos las llamadas realizadas expresadas en Minutos:

87	93	96	98	105	114	128	131	142	168
----	----	----	----	-----	-----	-----	-----	-----	-----

7. La distribución de edades del Censo Electoral para los distritos de Huancayo y El Tambo, en tantos por cien es la siguiente:

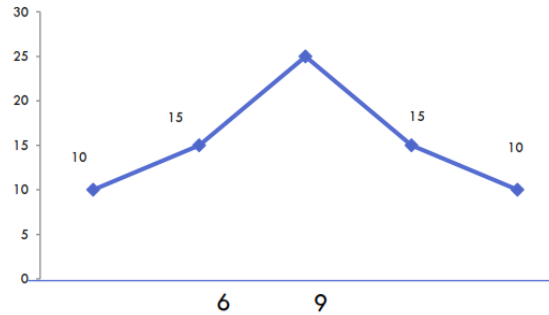
Edades	Huancayo	El Tambo
16-18	3.54	4.35
18-30	21.56	29.99
30-50	31.63	35.21
50-70	28.14	21.97
70-90	15.12	8.48

- a) Calcule la edad mediana para los dos distritos. Compáralas. ¿Que indican estos resultados?
b) ¿Qué comunidad tiene mayor variabilidad en la distribución de su edad?
8. Una compañía requiere los servicios de un técnico especializado. De los expedientes presentados, se han seleccionado 2 candidatos: A y B, los cuales reúnen los requisitos. Para decidir cuál de los 2 se va a contratar se toman siete pruebas a cada uno de ellos. Los resultados se dan a continuación:

	Prueba						
	1	2	3	4	5	6	7
Puntaje obtenido por A	57	55	54	52	62	55	59
Puntaje obtenido por B	80	40	62	72	46	80	40



- a) Halle e interprete todas las medidas de dispersión de los dos candidatos.
b) Estadísticamente, ¿Cuál de los candidatos debe ser contratado? Fundamente su respuesta.
9. Sean las edades de 75 estudiantes de educación virtual de la Universidad Continental, que se muestra en el gráfico. Elabora tablas de frecuencia y calcula las medidas de dispersión (varianza, desviación estándar, coeficiente de variación) interpretar cada medida estadística y graficar la distribución normal e interpreta los resultados a dos desviaciones estándar. ¿Los datos son homogéneos? ¿Por qué?



10. Se tiene dos muestras de obreros, cuyos ingresos diarios son:

Muestra 1: S/.138; S/.136; S/.146; S/.140 y S/.145,

Muestra 2: S/.134; S/.147; S/.147; S/.145 y S/.137

Compare el coeficiente de variabilidad de ambas muestras y determine que muestra de obreros presenta ingresos diarios más homogéneos (Fundamente su respuesta).

REFERENCIAS BIBLIOGRÁFICAS Y ENLACES

- Bejarano, M. (1995) Estadística descriptiva, probabilidades y lineamientos para la elaboración del protocolo de investigación en ciencias de la Salud. Universidad Peruana Cayetano Heredia.
- Triola F. (2009) Estadística Elemental. Editorial Mexicana. Décima Edición.
- Jorge Chue, Emma Barreno, Carlos Castillo, Rosa Millones, Félix Vásquez (2008) Estadística Descriptiva y Probabilidades. Universidad de Lima. Fondo Editorial



Semana N°7: MEDIDAS DE POSICIÓN RELATIVA

1. **TEMA: Medidas de posición relativa: Cuartiles y percentiles. Análisis exploratorio de datos.**

2. **PROPÓSITO:**

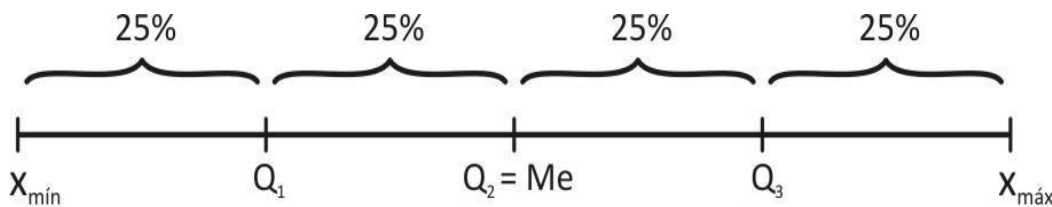
- Calcula las medidas de posición e interpreta el diagrama de cajas. De datos discretos y continuos.

MEDIDAS DE POSICIÓN RELATIVA: CUARTILES Y PERCENTILES

Son medidas que resultan para comparar valores de diferentes conjuntos de datos o para comparar valores dentro del mismo conjunto de datos.

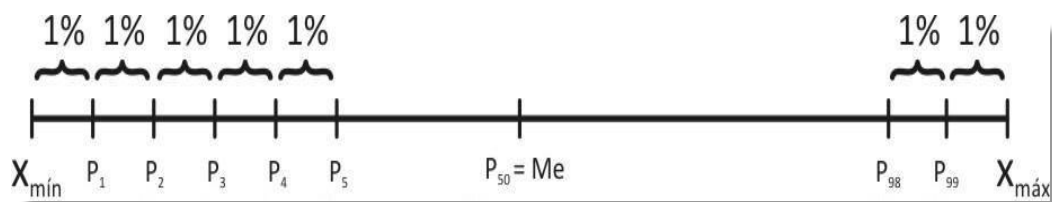
1. **Cuartiles (Q_i) y percentiles (P_i):**

Cuartiles: Los cuartiles dividen el conjunto de datos en cuatro partes iguales y las observaciones arriba del tercer cuartil constituyen el cuarto superior del conjunto de datos, el segundo cuartil es idéntico a la mediana y el primer cuartil separa el cuarto inferior de los tres cuartos superiores.



PARA CALCULAR CUARTILES(DATOS NO AGRUPADOS):	PARA CALCULAR CUARTILES(DATOS AGRUPADOS):
$Q_k = \frac{(\text{Nro. de cuartil}) \times (n+1)}{4}$	$Q_k = L_i + \left[A \left(\frac{\frac{kn}{4} - F_{j-1}}{f_j} \right) \right]$

Percentiles: Asimismo, un conjunto de datos (muestra o población) puede ser incluso más finamente dividido por medio de percentiles, el 99° percentil separa el 1% más alto del 99% más bajo, y así sucesivamente. A menos que el número de observaciones sea un múltiplo de 100, se debe tener cuidado al obtener percentiles.



Equivalencias:

$$\begin{aligned} Q_1 &= P_{25} \\ Q_2 &= P_{50} = \text{Me} \\ Q_3 &= P_{75} \end{aligned}$$

EL PERCENTIL K-ÉSIMO (DATOS NO AGRUPADOS):	EL PERCENTIL K-ÉSIMO (DATOS AGRUPADOS):
--	---



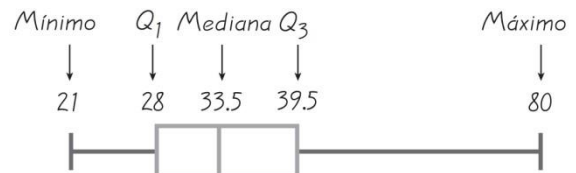
$P_k = \frac{(Nro. de percentil) \times (n+1)}{100}$	$P_k = L_i + \left[A \left(\frac{\frac{kn}{100} - F_{j-1}}{f_j} \right) \right]$
--	--

2. Análisis exploratorio de datos:

Es el proceso de utilizar herramientas estadísticas (como gráficas, medidas de tendencia central y medidas de variación) con la finalidad de investigar conjuntos de datos para comprender sus características importantes.

• **Graficas de cuadro (caja y bigote)**

En años recientes, se ha utilizado con éxito un resumen gráfico llamado gráfica de caja para describir varias de las características más prominentes de un conjunto de datos. Estas características incluyen 1) el centro, 2) la dispersión, 3) el grado y naturaleza de cualquier alejamiento de la simetría y 4) la identificación de las observaciones "extremas o apartadas" inusualmente alejadas del cuerpo principal de los datos. Como incluso un solo valor extremo puede afectar drásticamente los valores de \bar{x} y s , una gráfica de caja está basada en medidas "resistentes" a la presencia de unos cuantos valores apartados, la mediana y una medida de variabilidad llamada dispersión de los cuartos.



EJERCICIOS RESUELTOS:

- Dentro de un programa de capacitación de personal, en una evaluación de los cursos de Economía y Finanzas se obtuvieron las medias de 13 y 17, respectivamente.
- Para los siguientes datos, calcular el cuartil 1.

265	120	91	115	185	255	233	190	160	170
-----	-----	----	-----	-----	-----	-----	-----	-----	-----

Solución:

Primer paso: Ordenar los datos en forma ascendente:

91 - 115 - 120 - 160 - 170 - 185 - 190 - 233 - 255 - 265

Segundo paso: Se sabe que $Q_1 = P_{25}$, por lo tanto, calcularemos el percentil 25.

Calcular el localizador L: ($k = 25$) $L = \frac{25}{100} \times 10 = 2,5$

Tercer paso: Redondeamos L con las reglas del redondeo que hemos aprendido anteriormente, en este caso $L=3$

Cuarto paso: Calculamos el valor de P_{25} (el dato que ocupa la posición 3 empezando del menor dato).

$91 - 115 - \boxed{120} - 160 - 170 - 185 - 190 - 233 - 255 - 265$

$P_{25}=120$

Interpretación: El 25% de los datos son menores o iguales a 120 y el 75% restante son mayores o iguales que este valor.

- Los datos representan el peso en kilogramos de 12 cajas enviadas por encomienda a través de la Empresa "Cargo 1". Calcule el tercer cuartil.

9	10	12	3	5	7	15	10	9	11	13	11
---	----	----	---	---	---	----	----	---	----	----	----

Solución:

Primer paso: Ordenar los datos en forma ascendente:

3 5 7 9 9 10 10 11 11 12 13 15



Segundo paso: Se sabe que $Q_3 = P_{75}$, por lo tanto, calcularemos el percentil 75.

Calcular el localizador L: ($k = 75$)
$$L = \frac{75}{100} \times 12 = 9$$

Tercer paso: L sí es un número entero, por lo tanto se toman los datos de la posición 9 y 10, empezando del menor dato.

Cuarto paso: Calculamos el valor de P_{75} (promediamos los datos que ocupan las posiciones 9 y 10).

3 5 7 9 9 10 10 11 11 12 13 15

$$P_{75} = (11+12)/2 = 11,5$$

Interpretación: El 75% de los pesos son menores o iguales a 11,5 Kg. y el 25% restante son mayores o iguales que este valor.

4. Dada la siguiente tabla de distribución, calcule P_{45} e interprete.

Sueldos (\$)	Nº Trabajad.	F _i
(x _i)	(f _i)	
[90 - 120>	11	11
[120 - 150>	13	24
[150 - 180>	20	44
[180 - 210>	17	61
[210 - 240>	15	76
[240 - 270>	3	79
[270 - 300>	1	80
	n = 80	

Solución:

Primer paso: Calcular el localizador L: ($k = 45$)
$$L = \frac{45}{100} \times 80 = 36$$

Segundo paso: Se busca el valor de $L = 36$ en las frecuencias absolutas acumuladas (F_i), como no figura se toma el valor inmediato superior (44), lo cual indica que el intervalo [150 - 180> es el intervalo para el P_{45} .

Tercer paso: Calculamos el valor de P_{45} aplicando la fórmula:

$$P_{45} = 150 + 30 \left(\frac{36 - 24}{20} \right) = 168$$

Interpretación: El 45% de los sueldos a lo más son iguales a \$168 y el 55% restante son como mínimo iguales a este valor.

EJERCICIOS PROPUESTOS

1. Diga usted con sus propias palabras ¿qué es un valor o puntuación "z"?
2. Si la edad promedio de un grupo de trabajadores de Electrocentro S.A. es de 30 años y la desviación estándar es igual a 4 años, un trabajador recién contratado de 26 años ¿a cuántas desviaciones estándar de la edad promedio se ubicaría?
3. Una persona A mide 1,65 m y reside en una ciudad donde la estatura media es de 1,60 m y la desviación típica es de 20 cm. Otra persona B mide 1,80 m y vive en una ciudad donde la estatura



media es de 1,70 m y la desviación típica es de 15 cm. ¿Cuál de las dos será más alta respecto a sus conciudadanos?

- 4. Un profesor ha realizado dos tests a un grupo de 40 alumnos del turno tarde, obteniendo los siguientes resultados: Para el primer test la media es 6 y la desviación típica 1,5. Para el segundo test la media es 4 y la desviación típica 0,5. Un alumno obtiene un 6 en el primero y un 5 en el segundo. En relación con el grupo, ¿en cuál de los dos tests obtuvo mejor puntuación?
- 5. Diga usted con sus propias palabras ¿qué son los cuartiles? ¿Y qué son los percentiles? ¿Cuántos cuartiles y percentiles hay?
- 6. Establezca el valor de verdad (V) o falsedad (F) de las siguientes proposiciones:

$$Q_1 = P_{50} \quad () \qquad Q_3 = P_{75} \quad ()$$

$$Me = P_5 \quad () \qquad Q_2 = P_{50} = Me \quad ()$$

$$Q_1 = P_{25} \quad () \qquad P_{75} = Me \quad ()$$

- 7. Los tiempos (en minutos) empleados por un grupo de mecánicos para armar un motor fueron los siguientes:

11	18	15	13	14	16	10	9	14	17	18	20	25
----	----	----	----	----	----	----	---	----	----	----	----	----

- ¿Cuánto tiempo le tomó como máximo al 35% de los mecánicos más experimentados?
- ¿Cuánto tiempo como mínimo le tomó al 18% de los mecánicos con menos experiencia?

- 8. Los siguientes datos corresponden al peso de 15 estudiantes:

48	50	52	49	58
48	52	60	57	61
55	48	63	65	54

- ¿Cuál es el peso máximo del 75% de los alumnos? ¿Cuál es el peso mínimo del 63% de los alumnos?

- 9. Los ingresos diarios (en Nuevos Soles) de un grupo de supervisores de la obra de construcción del nuevo centro comercial se han resumido en la siguiente tabla de frecuencias:

Ingresos	x_i	f_i	F_i
[48 - 52>	50	12	
[52 - 56>	54	25	
[56 - 60>	58	58	
[60 - 64>	62	32	
[64 - 68>	66	10	
		n =	

- ¿Cuánto tiene como ingreso máximo el 32% de los supervisores?
- ¿Cuánto tiene como ingreso mínimo el 25% de los supervisores?

- 10. Dada la tabla siguiente, referente a los pesos de cierto número de pacientes de un hospital:

Pesos	f_i
[0 - 12>	5
[12 - 24>	24
[24 - 36>	18
[36 - 48>	36
[48 - 60>	17
[60 - 72>	10

- Calcule media, mediana y moda de los pesos.
- Calcule: Cuartil 1, 2 y 3. (Compare el cuartil 2 con la mediana). Comente.
- Calcule: Percentil 28 y 50. (Compare el percentil 50 con la mediana). Comente.
- ¿Cuánto pesa como máximo el 19% de pacientes?
- ¿Cuánto pesa como mínimo el 42% de pacientes?
- Calcule el rango intercuartil (RIC). Interprete.

- 11. En la siguiente tabla se muestra la distribución del tiempo (en horas) de duración de los componentes electrónicos de las marcas Alpha y Beta, sometidos a un trabajo continuo:

Tiempo de duración (Horas)	Marca ALPHA		Marca BETA	
	f_i	h_i	f_i	h_i
[0 - 100>	2	0,025	12	0,171
[100 - 200>	4	0,050	16	0,229
[200 - 300>	22	0,275	25	0,357
[300 - 400>	26	0,325	10	0,143
[400 - 500>	20	0,250	4	0,057
[500 - 600>	5	0,063	2	0,029
[600 - 700>	1	0,013	1	0,014
Total	80		70	

- Calcule e interprete la media y la moda para el tiempo de duración de los componentes Alpha y Beta, respectivamente.
- Se decide descartar el 15% de los componentes menos durables, ¿cuál debería ser el tiempo mínimo de duración en el componente Beta para no ser descartado?
- Por el contrario, se decide que el 23% de los componentes más durables de la marca Alpha sean los que se envíen primero al mercado. ¿Cuál debería ser el tiempo mínimo de duración en el componente Alpha para ser considerado en el primer lote que sea trasladado a los puntos de venta?



12. La granja que provee de pollos a EL MEZON, registra la siguiente tabla de distribución de los pesos (en gramos) de los pollos beneficiados:

Para efectos de venta y marketing la empresa los clasifica en tres categorías, de acuerdo a su peso:

Peso	Nº pollos
[1800 - 1900>	60
[1900 - 2000>	160
[2000 - 2100>	280
[2100 - 2200>	260
[2200 - 2300>	160
[2300 - 2400>	80
	1000

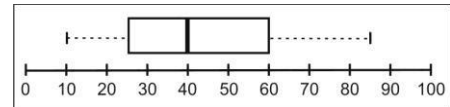
- El 20% de los pollos menos pesados pertenecen a la categoría de "pollos tiernos".
- El 60% de los pollos menos pesados siguientes pertenecen a la categoría de "super pollos".
- El resto (20% restante), pertenecen a la categoría de "polli pavos"

Responda:

- ¿Cuáles son los límites de peso entre las categorías referidas?
- ¿Cuántos pollos pertenecen a la categoría de polli pavos?

13. Se da el siguiente diagrama de caja:

- ¿Cuál es la mediana, el valor mínimo y el máximo, el primer y tercer cuartil?
- ¿Estaría usted de acuerdo en que la distribución es simétrica? ¿Por qué? Explique.



14. Los pesos de un grupo de estudiantes del Centro Preuniversitario de la Universidad Continental han dado los siguientes datos: $Q_1 = 50$ Kg. ; $Q_3 = 70$ Kg. ; $Me = 55$ Kg. Elabore el diagrama de caja y bigote para dichos pesos. En el gráfico, diga usted ¿cómo se consideraría a un peso de 89 Kg?

15. Elabore el diagrama de caja y bigote de los tiempos empleados por los mecánicos para armar un disco de embrague (ejercicio 9). Responda: Un tiempo de 28 minutos, ¿cómo sería considerado? ¿Y uno de 36 minutos, cómo sería considerado?

16. Las estaturas de los jugadores del equipo de futbol de la Universidad Continental viene dada por la siguiente tabla:

Halle e interprete el Q_1 y P_{70}

Talla (cm)	fi
[172 - 175>	5
[175 - 178 >	7
[178 - 181>	9
[181 - 184>	3
Total	

17. Calcule el P_{25} y P_{75} de las edades de un grupo de clientes que asistieron el día de inauguración de tiendas "Oeschle" - Huancayo.

Tallo	Hojas
1	8999
2	001234
3	01235
4	012

Interprete sus respuestas.

18. Calcular e interpretar los percentiles 30 y 75 de los siguientes datos.

42	52	56	60	64	69	72	73	79	80	84	86	88	90
----	----	----	----	----	----	----	----	----	----	----	----	----	----

REFERENCIAS BIBLIOGRÁFICAS Y ENLACES

- Triola F. (2009) Estadística Elemental. Editorial Mexicana. Décima Edición.
- Devore J. (2008) Probabilidad y Estadística para Ingeniería y Ciencias. Editorial Latinoamericana. Séptima Edición

Semana N°8: MEDIDAS DE FORMA

1. **TEMA: Medidas de forma: Asimetría y Curtosis**
2. **PROPÓSITO:**

- Calcula e interpreta la curtosis

MEDIDAS DE DISTRIBUCIÓN O FORMA: ASIMETRÍA Y CURTOSIS

Son medidas que resultan útiles para cuantificar la semejanza que tiene la distribución de los datos con respecto a la distribución simétrica y unimodal conocida como "distribución normal". Estas medidas son:

4. Coeficiente de Asimetría de Pearson:

$$As = \frac{3(\bar{X} - Me)}{s}$$

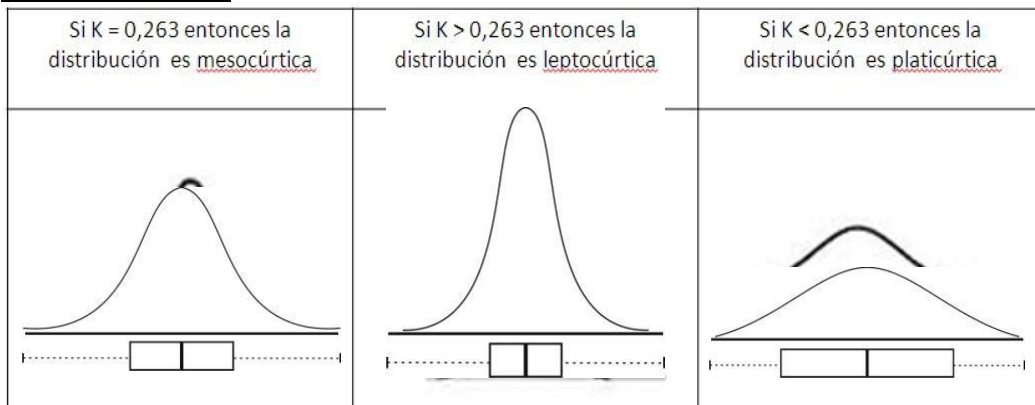
Interpretación:

- Si $As = 0$, La distribución es simétrica, esto es:
$$\bar{X} = Me = Mo$$
- Si $As > 0$, La distribución es asimétrica positiva, esto es:
$$Mo < Me < \bar{X}$$
- Si $As < 0$, La distribución es asimétrica negativa, esto es:
$$\bar{X} < Me < Mo$$

5. Coeficiente de curtosis: Compara la dispersión de los datos observados cercanos al valor central con la dispersión de los datos cercanos a ambos extremos de la distribución. Se calcula mediante:

$$K = \frac{P_{75} - P_{25}}{2(P_{90} - P_{10})}$$

Interpretación:



Al igual que el caso de los coeficientes de asimetría de una distribución, los que representan a curtosis o apuntamiento se utilizan para ayudar a describir las características de una distribución y no precisamente como medidas, ya que a veces el valor de la curtosis se contradice con la realidad por estar relacionado con la distribución normal (distribución teórica). Este coeficiente solo se calcula para distribuciones simétricas o ligeramente asimétricas.

EJERCICIOS RESUELTOS:

1. Sean los pesos de 8 niños (en Kg):



6	9	9	12	12	12	15	17
---	---	---	----	----	----	----	----

Solución:

- Asimetría: Calculamos los estadígrafos necesarios:

$$\bar{X} = \frac{6+9+9+12+12+12+15+17}{8} = 11,5 \text{ Kg}$$

Me=12 Kg Mo=12 Kg s=3,505 Kg

Luego:

$$As = \frac{3(\bar{X} - Me)}{s} = \frac{3(11,5 - 12)}{3,505} = -0,14$$

Se puede observar que la asimetría es negativa (cola a la izquierda).

- Curtosis: Calculamos los estadígrafos necesarios:

P₂₅=9Kg P₇₅=13,5Kg
P₁₀=6Kg P₉₀=17Kg

Luego:

$$K = \frac{P_{75} - P_{25}}{2(P_{90} - P_{10})} = \frac{13,5 - 9}{2(17 - 6)} = 0,205$$

Se puede observar que la distribución tiende a ser platicúrtica (mayor dispersión).

2. El departamento de Recursos Humanos de "Plaza Veá" hizo un estudio de 120 empleados, obteniéndose los siguientes datos:

Salarios(S/.)	Nº de empleados
[180-220>	12
[220-260>	30
[260-300>	36
[300-340>	24
[340-380>	18

- a) Hallar el primer y segundo coeficiente de asimetría de Pearson.
b) Hallar el coeficiente de curtosis.

Solución:

Construimos la tabla de frecuencias:

Salarios(S/.)	X _i	f _i	Fi	h _i	p%
[180-220>	200	12	12	0,10	0,10
[220-260>	240	30	42	0,25	0,35
[260-300>	280	36	78	0,30	0,65
[300-340>	320	24	102	0,20	0,85
[340-380>	360	18	120	0,15	1
		120		1	



- Asimetría: Calculamos los estadígrafos necesarios:

$$\bar{X} = \frac{33840}{120} = 282$$

$$Me = 260 + 40 \left(\frac{60-42}{36} \right) = 280$$

$s=48,33$

Luego:

$$As = \frac{3(\bar{X}-Me)}{s} = \frac{3(282-280)}{48,33} = 0,124$$

Se puede observar que la asimetría es ligeramente positiva (cola a la derecha).

- Curtosis: Calculamos los estadígrafos necesarios:

$$P_{25} = 220 + 40 \left(\frac{30-12}{30} \right) = 244 \qquad P_{75} = 300 + 40 \left(\frac{90-78}{24} \right) = 320$$

$$P_{10} = 180 + 40 \left(\frac{12-0}{12} \right) = 220 \qquad P_{90} = 340 + 40 \left(\frac{108-102}{18} \right) = 353,33$$

Luego:

$$K = \frac{P_{75} - P_{25}}{2(P_{90} - P_{10})} = \frac{320 - 244}{2(353,33 - 220)} = 0,285$$

Se puede observar que la distribución tiende a ser leptocúrtica (menor dispersión).

EJERCICIOS PROPUESTOS:

- Un grupo de jóvenes de la facultad de Ingeniería han sido encuestados en relación a su estatura. Los datos se han procesado mediante el IBM SPSS. Los resultados se muestran en la tabla adjunta. Calcule los coeficientes de asimetría de Pearson e indique si la distribución de estaturas es simétrica o asimétrica.
- Una muestra de digitadoras de textos reveló que su rapidez media de tecleo es de 87 palabras por minuto, con una mediana de 73 palabras. La desviación estándar es de 16,9 palabras por minuto. ¿Cuál es el coeficiente de asimetría? Interprete la respuesta, luego conteste: ¿La mayoría de las secretarías son rápidas tecleando o son lentas?
- Dadas las siguientes mediciones de la emisión diaria (en toneladas) de óxido de azufre en una planta industrial:

ESTATURA DEL ENCUESTADO	
Nº Validos	50
Perdidos	0
Media	1,6330
Mediana	1,6400
Moda	1,64
Desv. Tip.	,06270
Varianza	,004

Construya una distribución de frecuencias y luego calcule los coeficientes de asimetría y curtosis:

8,3	15,8	16,2	18,5	22,7	23,0
11,0	13,4	17,3	19,1	22,7	23,9
11,2	15,1	17,9	19,6	22,3	24,6
12,2	15,8	18,0	20,1	21,9	26,4
12,3	13,4	18,4	20,5	21,3	26,8

- Se ha medido pulsaciones de un equipo de atletas después de una carrera. Los datos obtenidos son:

Pulsaciones	[70-75>	[75-80>	[80-85>	[85-90>	[90-95>	[95-100>
Nº Atletas	3	3	7	10	12	8



Se pide:

- Hallar el primer coeficiente de asimetría de Pearson e indica que tipo de asimetría presenta la distribución.
 - Indicar que tipo de apuntamiento o curtosis corresponde a la distribución.
5. En una prueba de Análisis Matemático aplicado a 20 estudiantes de la carrera de Administración , se obtuvo la siguiente distribución:

Puntaje	[35-45>	[45-55>	[55-65>	[65-75>	[75-85>	[85-95>
N° estudiantes	1	3	8	3	3	2

Se pide:

- Hallar la media aritmética y la mediana.
 - Hallar la el coeficiente de asimetría
 - Calcule el puntaje mínimo que debería obtener para estar considerado en el tercio superior.
6. Se han presentado los datos de rendimiento académico de 17 estudiantes, estos datos reportan las siguientes estadísticas:
 $\bar{x} = 14$, Mediana = 15, Moda = 17, Cuartil 1 = 10,5, Cuartil 3 = 17 Percentil 10 = 4 y Percentil 90 = 35
- a. Determine el valor de la curtosis. Interprete el resultado.
 - b.

REFERENCIAS BIBLIOGRÁFICAS Y ENLACES

- Bejarano, M. (1995) Estadística descriptiva, probabilidades y lineamientos para la elaboración del protocolo de investigación en ciencias de la Salud. Universidad Peruana Cayetano Heredia.
- Triola F. (2009) Estadística Elemental. Editorial Mexicana. Décima Edición.



GUÍA DE PRÁCTICA DE PROBABILIDAD Y ESTADÍSTICA

TERCERA UNIDAD PROBABILIDAD

- GUÍA DE PRÁCTICA N° 10: Probabilidades
- GUÍA DE PRÁCTICA N° 11: Regla de la suma y de la multiplicación
- GUÍA DE PRÁCTICA N° 12: Probabilidad total y teorema de Bayes

Semana N°10: PROBABILIDADES

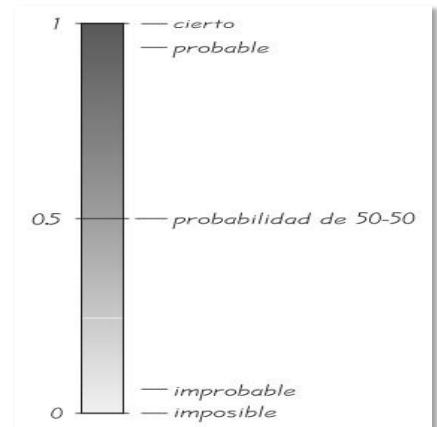
1. TEMA: Fundamentos: Definiciones, Notación, Reglas, Ley de los números grandes, Sucesos, Probabilidades.

2. PROPÓSITO:

- Define la probabilidad de un evento.
- Analiza y resuelve problemas sobre probabilidades, aplicando las propiedades más importantes

1. Conceptos básicos:

- **Experimento Aleatorio:** Es todo proceso que consiste en la ejecución de un acto una o más veces, cuyo resultado en cada prueba no se puede predecir con certeza y tiene dos o más resultados posibles
- **Espacio muestral:** Se compone de todos los resultados posibles de un experimento aleatorio, es decir, está formado por todos los resultados que ya no pueden desglosarse más. Se simboliza con la letra griega Ω
- **Evento:** Es un subconjunto de resultados de un espacio muestral. Un evento simple es un resultado o un evento que ya no puede desglosarse en componentes más simples. Se denota con letras mayúsculas.



Tipos de Eventos

1. **Evento Elementales:** Es el formado por un solo elemento. Son los elementos del espacio muestral Ω

2. **Evento Compuesto:** Es el formado por dos o más elementos.

3. **Evento Seguro:** Es el que va a ocurrir siempre, por lo tanto debe ser el mismo espacio muestral Ω

4. **Evento Imposible:** Es el que no va a ocurrir nunca, se simboliza con ϕ

5. **Eventos Incompatibles o Disjuntos:** Son los que no pueden ocurrir a la vez. $E \cap F = \emptyset$

6. **Evento Contrario o Complementario:** Dado un suceso A, se denomina suceso contrario o complementario A' al suceso que se verifica o ocurre cuando no lo hace A, está formado por los elementos del espacio muestral que no pertenecen a A



OPERACIONES CON EVENTOS:

	UNIÓN \cup	INTERSECCIÓN \cap
Asociativa	$(A_1 \cup A_2) \cup A_3 = A_1 \cup (A_2 \cup A_3)$	$(A_1 \cap A_2) \cap A_3 = A_1 \cap (A_2 \cap A_3)$
Commutativa	$A_1 \cup A_2 = A_2 \cup A_1$	$A_1 \cap A_2 = A_2 \cap A_1$
Idempotencia	$A \cup A = A$	$A \cap A = A$
Simplificativa	$A \cup (B \cap A) = A$,	$A \cap (B \cup A) = A$
Distributiva	$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$	
De los complementarios	$A \cup A^c = \Omega$,	$A \cap A^c = \emptyset$
Del suceso imposible	$A \cup \emptyset = A$,	$A \cap \emptyset = \emptyset$
Del suceso seguro	$A \cup \Omega = \Omega$,	$A \cap \Omega = A$
Involución del contrario	$(A^c)^c = A$	
Contrarios de los sucesos seguro e imposible	$\Omega^c = \emptyset$	$\emptyset^c = \Omega$
Leyes de Morgan	$\overline{A \cup B} = \overline{A} \cap \overline{B}$	$\overline{A \cap B} = \overline{A} \cup \overline{B}$

PROBABILIDADES

- **Probabilidad:** Es la medida numérica de la posibilidad de que un evento pueda ocurrir. Su valor esta entre 0 y 1.
- **Notación de probabilidades:**

P: denota una probabilidad.

P(A): denota la probabilidad de que ocurra el evento A.

Formas de calcular una probabilidad: Puede ser:

Aproximación de la probabilidad por frecuencias relativas
Valor fijo al cual tiende a ocurrir después del experimento.
$P(A) = \frac{\text{número de veces que ocurrió } A}{\text{número de veces que se repitió el ensayo}}$
Método clásico de la probabilidad
Método son efectuar el experimento, te puedes anticipar a lo que va a suceder.
$P(A) = \frac{\text{número de formas en que puede ocurrir } A}{\text{número de sucesos simples diferentes}} = \frac{S}{n}$
Probabilidades subjetivas
Probabilidad asignada bajo un criterio personal, basado en cualquier tipo de evidencia disponible. Implica un grado de creencia personal
Ejemplos de la probabilidad subjetiva son estimar la probabilidad de que un equipo de fútbol gane el campeonato este año.



Ley de los grandes números: Esta ley engloban varios teoremas de la teoría de probabilidades, que describen el comportamiento del promedio de una sucesión de variables aleatorias conforme aumenta su número de ensayos. Lo que explica que es "La **probabilidad de un evento** es la constante a la que se aproxima la frecuencia relativa cuando el experimento se repite muchísimas veces".

Axiomas de probabilidad

- La probabilidad de un evento A, $P(A) \geq 0$.
- La probabilidad de un evento que ocurrirá con certeza es 1.
- Para toda colección de eventos incompatibles $\{A_i\}$ con $A_i \cap A_j = \emptyset$ para $i \neq j$, debe ser:

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

PROPIEDADES ELEMENTALES DE LA PROBABILIDAD

Toda probabilidad cumple una serie de propiedades, las cuales se obtienen como consecuencia de los axiomas que debe de cumplir:

1. $P(\emptyset) = 0$.
2. Se cumple la *aditividad* finita para sucesos incompatibles.

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i) \quad \text{si } A_i \cap A_j = \emptyset, i \neq j$$

3. La probabilidad del complementario de un suceso A es
 $P(A') = 1 - P(A)$
4. Si dos sucesos son tales que $A \subset B$, entonces $P(A) \leq P(B)$.
5. La probabilidad de todo suceso A es un número entre 0 y 1:
 $0 \leq P(A) \leq 1$.
6. Si dos sucesos no son incompatibles, la probabilidad de su unión debe calcularse por la siguiente regla:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Algunos otros teoremas de probabilidades

1. Probabilidad del suceso seguro

La probabilidad del suceso seguro es 1, es decir $P(S) = 1$

2. Diferencia de eventos

$$P(A - B) = P(A) - P(A \cap B)$$

3. Desigualdad de Boole

$$P(A \cap B) \geq 1 - P(\bar{A}) - P(\bar{B})$$

4. Probabilidad de la unión de tres o más sucesos

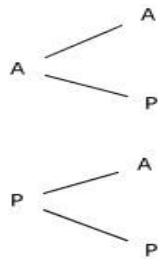
Vamos a probar que siendo A_1, A_2 y A_3 sucesos cualesquiera

$$P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) - P(A_1 \cap A_3) + P(A_2 \cap A_3) - P(A_1 \cap A_2 \cap A_3)$$



EJERCICIOS RESUELTOS:

Construya el espacio muestral para los siguientes experimentos:



- Verificar el estado de dos transistores (apagado=A, prendido=P).

$$\Omega = \{AA, AP, PA, PP\}$$



1. ¿Cuál es la probabilidad de obtener 1 sello si arrojamos una moneda una vez?

Solución:

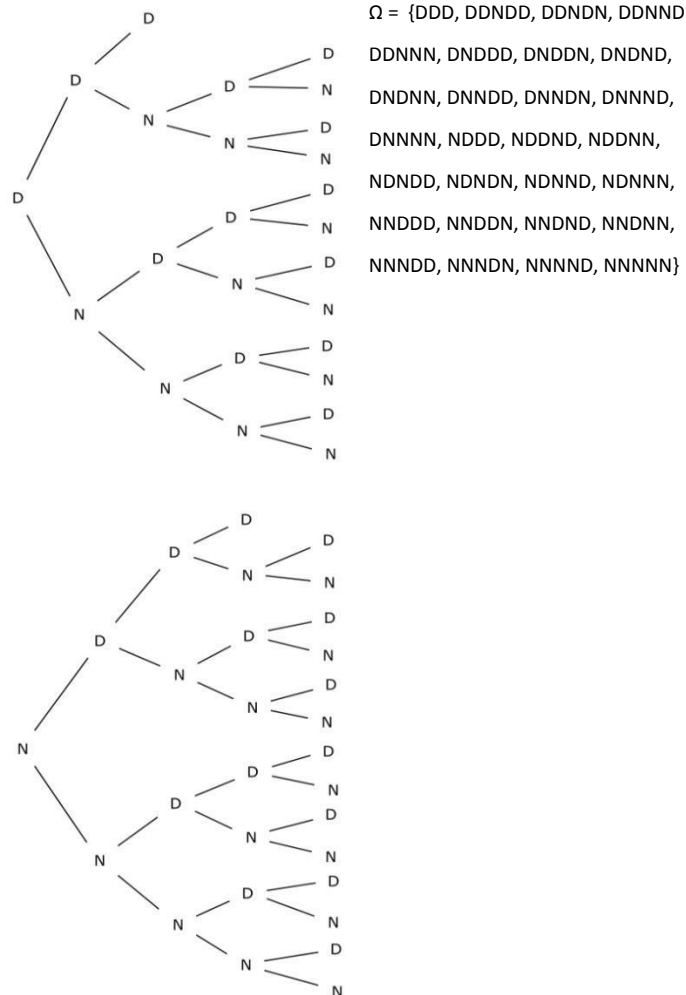
- EXPERIMENTO: Arrojamos una moneda.
- EVENTO: Obtener un sello.
- ESPACIO MUESTRAL. = $P(\Omega) = S = \{\text{cara, sello}\}$ #S=2

$$Prob = \frac{\text{Numero de resultados favorables}}{\text{Numero de resultados posibles}}$$

$$\frac{\{s\}}{\{c, s\}} = \frac{1}{2} = 0,5$$

2. Una línea de producción clasifica sus productos en defectuosos "D" o no defectuosos "N". De un almacén donde guardan la producción diaria de esta línea, se extraen artículos hasta observar tres defectuosos consecutivos o hasta que se hayan verificado cinco artículos (sin importar que sean defectuosos o no defectuosos).

Solución:



3. ¿Cuál es la probabilidad de obtener 1 cara si arrojamos una moneda tres veces?



Solución:

- EXPERIMENTO: Arrojamus una moneda tres veces.
- EVENTO: Obtener una cara.
- ESPACIO MUESTRAL. = $P(\Omega) = S = \{ccc, ccs, csc, css, scc, scs, ssc, sss\} = \#S=8$

$$Prob = \frac{\text{Numero de resultados favorables}}{\text{Numero de resultados posibles}}$$

$$\frac{\{css, scs, ssc\}}{\{ccc, ccs, csc, css, scc, scs, ssc, sss\}} = \frac{3}{8} = 0,375$$

EJERCICIOS PROPUESTOS:

1. Se lanza un dado. Encuentre la probabilidad de que el número de la cara superior sea:
 - a) 3
 - b) Impar
 - c) Menor que 5
 - d) No mayor que 3
2. En el examen de Costos de opción múltiple, con cinco posibles respuestas para cada pregunta ¿Cuál es la probabilidad de responder una pregunta correctamente si usted la elige al azar?
3. En la sección de control de calidad de la empresa "LALO'S", se encontró 3 productos en mal estado, en una partida de 95 productos tomados aleatoriamente de la producción de un día. Estima la probabilidad de producir uno en mal estado.
4. Se selecciona una familia que posee dos automóviles, y para el más nuevo y el más viejo observamos si fue fabricado en Estados Unidos, Europa o Asia.
 - ¿Cuáles son los posibles resultados de este experimento?
 - ¿Cuáles resultados están contenidos en el evento de que por lo menos uno de los dos automóviles sea extranjero?
5. En la empresa Scotiabank, se tiene una muestra de cuatro hipotecas para vivienda está clasificada como de tasa Fija (F) o tasa Variable (V).
 - a) ¿Cuáles resultados están en el evento de que exactamente tres de las hipotecas seleccionadas sea de tasa variable?
 - b) ¿Cuáles resultados están en el evento de que las cuatro hipotecas sean del mismo tipo?
 - c) ¿Cuáles resultados están en el evento de que a lo sumo una de las cuatro hipotecas sea de tasa Variable?
6. Construya el espacio muestral del lanzamiento de una moneda.
7. Construya el espacio muestral del lanzamiento de tres monedas.
8. Construya el espacio muestral del siguiente experimento: Lanzar una moneda hasta que ocurra "cara".
9. Suponga que una caja contiene un número igual de canicas rojas y de amarillas, el doble de canicas verdes que de rojas. Se extrae una canica de la caja y se observa su color. Asigne la probabilidad de obtener una canica verde.
10. Construya el espacio muestral para el siguiente experimento: En un laboratorio químico, el volumen producido por día de un producto "X" varía desde un valor mínimo "a" hasta un valor máximo "b", los cuales corresponden a la cantidad de insumo con que se cuente. Se escoge un día aleatoriamente y se mide el volumen producido.
11. Los artículos provenientes de una línea de producción para la empresa MAESTRO se clasifican en defectuosos (D) y no defectuosos (N), se observan los artículos y se anota su condición. Este proceso se continúa hasta observar dos defectuosos consecutivos o hasta que se observen tres artículos no defectuosos. Construya el espacio muestral asociado a este experimento.



12. Una línea de producción para PROMART clasifica sus productos en defectuosos (D) y no defectuosos (N). De un almacén que guarda la producción diaria de esta línea se extraen artículos hasta observar dos defectuosos consecutivos o hasta que se hayan verificado cuatro artículos (sin importar que sean defectuosos o no defectuosos). Construya el espacio muestral asociado a este experimento.
13. Un tazón contiene dos tipos de huevos de chocolate para pascua de apariencia idéntica, envueltos en aluminio. Todos menos 42 son chocolates de leche y todos menos 35 son de chocolate oscuro.
 - ¿Cuántos de cada tipo hay en el tazón?
 - ¿Cuántos chocolates hay en el tazón?
14. Si ruedas un dado 40 veces y 9 de las rodaduras resultan en un "5", ¿qué probabilidad empírica observas para el evento?
15. En un estudio de 420 000 usuarios de teléfono celular de Movistar, se encontró que 135 desarrollaron cáncer cerebral o del sistema nervioso. Estime la probabilidad de que un usuario de teléfono celular que se seleccionó al azar desarrolle un cáncer de este tipo. ¿Qué sugiere el resultado acerca de los teléfonos celulares como causantes de cáncer de este tipo?
16. En la fabricación de un cierto tipo de clavos, aparecen un cierto nº de ellos defectuosos. Se han estudiado 200 lotes de 500 clavos cada uno obteniendo:

Clavos defectuosos	1	2	3	4	5	6	7	8
Nº de lotes	5	15	38	42	49	32	17	2

- Se selecciona al azar un lote, determine las siguientes probabilidades:
- ¿Cuál es la probabilidad de se encuentre 7 clavos defectuosos?
 - ¿Cuál es la probabilidad de encontrar al menos 3 clavos defectuosos?
 - ¿Cuál es la probabilidad de encontrar a lo más 5 clavos defectuosos?
 - ¿Cuál es la probabilidad de encontrar menos de tres clavos y más de 7 clavos defectuosos?
17. Determine la probabilidad de que exactamente dos de los tres hijos de una pareja sean varones. Suponga que es igualmente probable dar a luz un niño que una niña, y que el género de cualquier hijo no influye en el género de otro.
 18. En una carrera de natación intervienen sólo tres estudiantes: Matías, Carlos y Sandro. Matías y Sandro tienen la misma probabilidad de ganar y el doble que la de Carlos. Hallar la probabilidad de que gane Sandro.
 19. El 35% de los trabajadores de la planta de producción de la Empresa "SAZON LOPESA" están conformes con la Dirección. De manera inopinada se aplica una encuesta a 15 de estas personas.
 - ¿Cuál es la probabilidad de que exactamente 4 de ellas estén conformes con la Dirección?
 - ¿Cuál es la probabilidad de que a lo más 2 personas estén conformes con la Dirección?
 20. Según una encuesta, los trabajadores de oficina de la empresa RIPLEY 28 horas por semana trabajando en la tienda. Si la desviación estándar es de 8 horas y se sabe que los tiempos siguen una distribución normal.
 - ¿Cuál es la probabilidad de que un empleado este trabajando entre 26 a 33 horas semanales?
 - ¿Qué porcentaje de trabajadores estén trabajando menos de 24 horas semanales?

REFERENCIAS BIBLIOGRÁFICAS Y ENLACES

- Triola Mario F. ESTADÍSTICA. Pearson Educación. México 2009
- Robert Johnson, Patricia Kuby. Estadística Elemental. Thomson Editorial. México 2002



Semana N°11: REGLA DE LA SUMA Y DE LA MULTIPLICACIÓN

1. TEMA: Regla de la suma y de la multiplicación

2. PROPÓSITO:

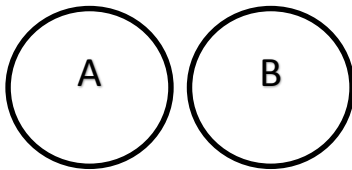
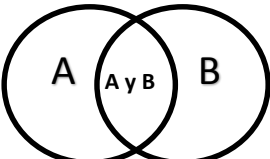
- Reconoce y aplica la regla de la suma y la multiplicación, identificando si un evento es mutuamente excluyente o no mutuamente excluyente, dependiente o independiente en situaciones de su contexto.
- Resuelve ejercicios de probabilidades haciendo uso de la regla de la suma y multiplicación, dando énfasis a una probabilidad condicional.

REGLA DE LA SUMA

1. Definición

Un suceso compuesto es cualquier suceso que combine dos o más sucesos simples.

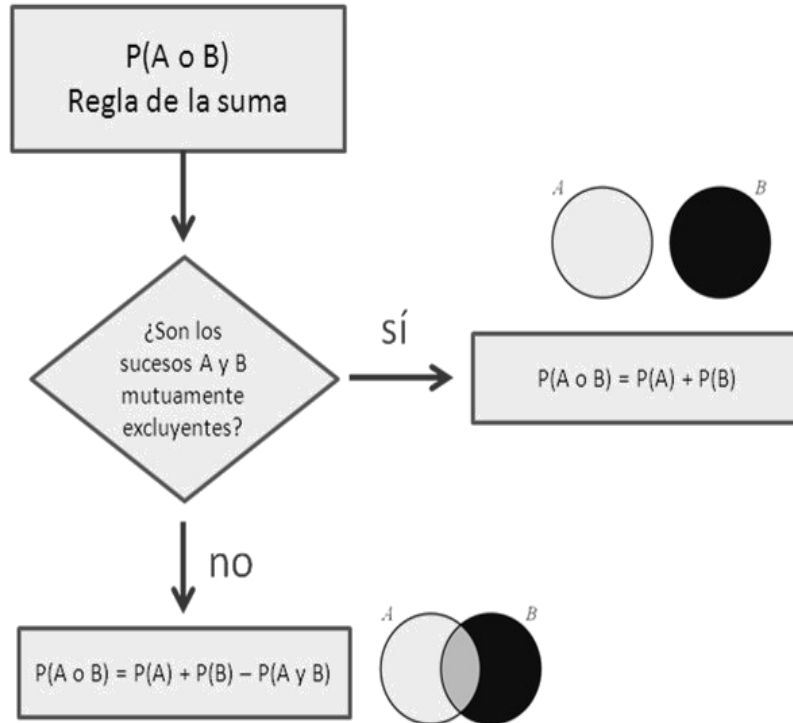
2. Clases de eventos

<p>Eventos Mutuamente Excluyentes</p>	<p>Dos o más eventos que no pueden ocurrir al mismo tiempo.</p> <p>Si A y B son sucesos mutuamente excluyentes, habremos de sumar la Probabilidad de suceso A a la Probabilidad del Suceso B.</p>	<p>$P(A \text{ o } B) = P(A \cup B) = P(A) + P(B)$</p> 
<p>Eventos NO Mutuamente Excluyentes</p>	<p>Dos o más eventos que si pueden ocurrir al mismo tiempo.</p> <p>Si A y B son sucesos NO mutuamente excluyentes, habremos de sumar la Probabilidad de suceso A a la Probabilidad del suceso B y restar la probabilidad conjunta de los sucesos A y B.</p>	<p>$P(A \text{ y } B) = P(A \cap B) = P(A) + P(B) - P(A \text{ y } B)$</p> 

3. Notación de la regla de la suma

$P(A \text{ o } B)$: P (en un solo ensayo, ocurre el suceso A u ocurre el suceso B o ambos ocurren)

Para calcular la probabilidad de que ocurra el suceso A o el suceso B, primero debemos obtener el número total de maneras en que puede ocurrir A y de maneras en que puede ocurrir B, pero calculamos ese total sin contar cada resultado más de una vez. La palabra clave en esta sección es "o", que significa: uno o el otro o ambos.



REGLA DE LA MULTIPLICACIÓN

1. Definición:

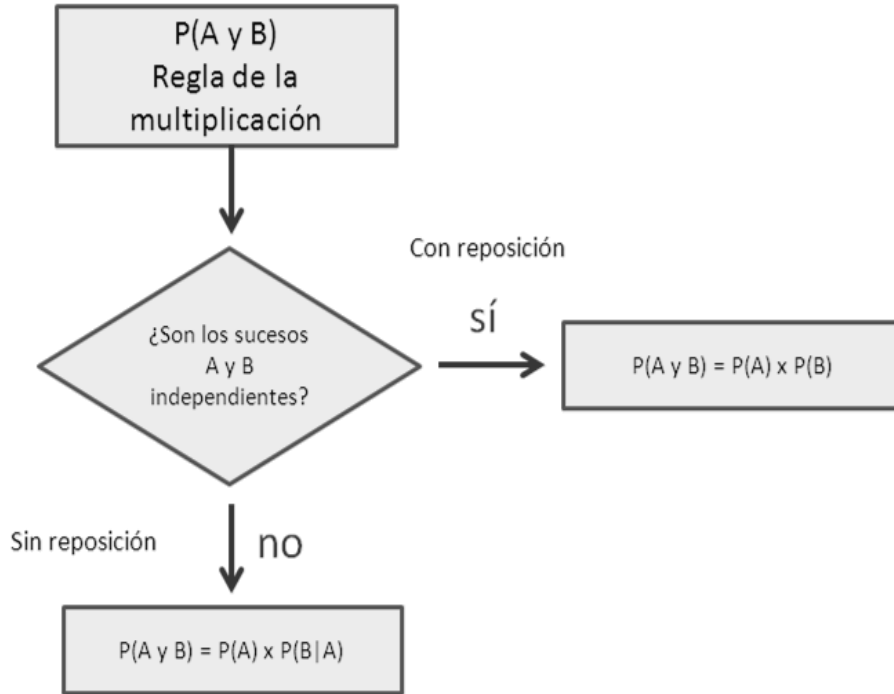
La regla para el cálculo de $P(A \text{ y } B)$ se denomina regla de la multiplicación porque implica multiplicar la probabilidad del suceso A por la probabilidad del suceso B (donde la probabilidad del suceso B se ajusta por el resultado del suceso A)

2. Notación de la regla de la multiplicación:

$P(A \text{ y } B)$: P (el suceso A ocurre en un primer ensayo y el suceso B en un segundo ensayo)

La **regla de la multiplicación** requiere que:

Dos eventos A y B sean INDEPENDIENTES cuando la ocurrencia de uno no afecta las probabilidades de la ocurrencia de los demás).	$P(A \text{ y } B) = P(A \cap B) = P(A) P(B)$.
Dos eventos A y B sean DEPENDIENTES cuando A y B no son independientes	$P(A \text{ y } B) = P(A \cap B) = P(A) P(B/A)$



PROBABILIDAD CONDICIONAL

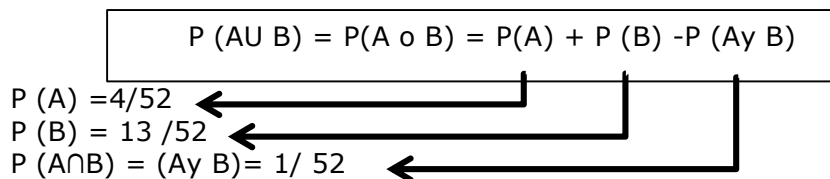
La **probabilidad condicional** de un suceso es una probabilidad obtenida con la información adicional de algún otro evento que ya ocurrió. $P(B|A)$ denota la probabilidad condicional de que el suceso B ocurra, dado que el suceso A ya ocurrió, y puede calcularse dividiendo la probabilidad de que ambos sucesos A y B ocurran entre la probabilidad del suceso A :

$$P(B | A) = \frac{P(AyB)}{P(A)}$$

EJERCICIOS RESUELTOS:

1. De una baraja de cartas, ¿Cuál es la probabilidad de que al *extraer una carta, que sea As o espada*?

Solución:



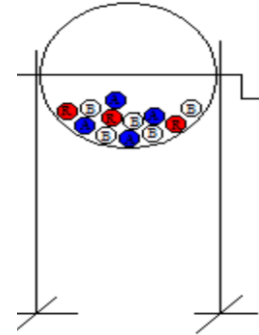
Entonces:

$$P(AU B) = P(A o B) = 4/52 + 13/52 - 1/52$$

$$P(AU B) = P(A o B) = 16/52 = 4/13 = 0.308$$



2. En una caja hay 7 bolas azules y 3 rojas. Se sacan dos bolas al azar de una en una. Hallar la probabilidad de que la primera sea azul (Evento A) y la segunda también (Evento B):



Solución:

- a) Entonces, si la primera bola no se devuelve, la probabilidad del evento B depende del evento A
 $P(A \cap B) = P(A) \times P(B/A) = (7/10)(6/9) = 0.47$
- b) Pero, si la primera bola se devuelve, la probabilidad del evento B no depende del evento A (Caso hipotético solo para hacer la comparación, pues no pide en el ejercicio)
 $P(A \cap B) = P(A) \times P(B) = (7/10)(7/10) = 0.49$

3. Se realizó una prueba de diagnóstico de prueba de embarazo a 30 mujeres, los resultados se muestran en la siguiente tabla:

Embarazadas	10	6	16
No embarazadas	3	11	14
TOTAL	13	17	30

Si seleccionamos una mujer al azar, ¿cuál es la probabilidad de seleccionar a una mujer embarazada o que su resultado salió positivo?

Solución:

$$P(A \cup B) = P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$P(A) = 16/30$

$P(B) = 13/30$

$P(A \cap B) = (A \cap B) = 10/30$

Entonces:

$P(A \cup B) = P(A \cup B) = 16/30 + 13/30 - 10/30$

$P(A \cup B) = 0.53 + 0.43 - 0.33$

$P(A \cup B) = P(A \cup B) = \mathbf{0.6333}$

4. Suponga que hay 10 focos en una caja, y se sabe que 3 están defectuosos. Se extrae primero uno de los focos y antes de extraer el segundo, se devuelve el primero a la caja. Calcule la probabilidad de que ambos sean defectuosos.

Solución:

$$P(A \cap B) = \frac{3}{10} \times \frac{3}{10} = 0,09$$

5. En la evaluación de 800 personas (entre varones y mujeres) para el ascenso de gerentes en Interbank se determinó que: 182 de los postulantes eran mujeres, de las cuales solo 53 lograron el ascenso, por otro lado 117 varones no lograron su ascenso. ¿Cuál es la probabilidad de que un gerente:

- a) Seleccionado al azar, haya ascendido, dado que es varón?

Solución:

- ✓ Primero, elaboramos la tabla de contingencia para que el ejercicio sea de fácil entendimiento:



	VARÓN	MUJER	TOTAL
ASCENDIÓ	501	53	554
NO ASCENDIÓ	117	129	246
TOTAL	618	182	800

✓ Segundo, comenzamos a resolver cada ítem:

La **intersección** va en la parte va como numerador en la fracción

Identifica la condicional, luego ubicamos en la base de la fracción

a) Seleccionado al azar, haya ascendido, dado que es varón

$$P(\text{Ascendió} \setminus \text{Varon})$$

	VARÓN	MUJER	TOTAL
ASCENDIÓ	501	53	554
NO ASCENDIÓ	117	129	246
TOTAL	618	182	800

Se busca Hombre que haya ascendido (**la intersección**)

Por la condicional buscamos siempre el **TOTAL** que es varón=618

Entonces:

$$P(\text{Ascendió} \setminus \text{Varon}) = \frac{501}{618} = 0,8107 = 81,07\%$$

EJERCICIOS PROPUESTOS

- Se lanzan dos dados, ¿Cuál es la probabilidad de que al menos sumen 10 ambos casos?
Identifica:
Experimento: _____
Evento: _____
Espacio muestral: _____
Resultado: _____
- Considere una baraja de 52 cartas y se desea extraer una carta. ¿Cuál es la probabilidad de obtener J o corazón?
- Se compraron 30 lápices de diferentes colores: 12 azules, 8 amarillos y 10 verdes. ¿Cuál es la probabilidad al extraer un lápiz de que sea: a) azul, b) azul o amarillo, c) amarillo o verde?
- Suponga que se tienen 30 fichas de tres colores, así: Amarillo = 15 fichas, negro = 10 fichas, azul = 5 fichas. Al mezclarlas, ¿cuál es la probabilidad, al sacar una de ellas de que sea: a) azul, b) azul o negra, c) amarilla o negra?



5. Consideremos el lanzamiento de un dado. Usted gana el juego si el resultado es impar o divisible por dos. ¿Cuál es la probabilidad de ganar?
6. La empresa Cruz del Sur sortea una AUDI R8 0km entre sus 180 mejores clientes. De ellos, 72 son mujeres, 90 son solteros y 60 son mujeres casadas. Se pide:
 - a) ¿Cuál es la probabilidad de que le toque la camioneta a una mujer o este casada?
 - b) ¿Cuál es la probabilidad de que sea una mujer, si es casada?
 - c) ¿Cuál es la probabilidad de que le toque a un varón, si no es casado?
7. Si planteamos el ejercicio anterior con la condición de ganar obteniendo un resultado par o divisible por 3. ¿Cuál es la probabilidad de ganar?
8. Se realizó una prueba de diagnóstico de prueba de embarazo a 35 mujeres , los resultados se muestran en la siguiente tabla:

	Positivo	Negativo	TOTAL
Embarazadas	11	8	19
No embarazadas	3	13	16
TOTAL	14	21	35

Si seleccionamos una mujer al azar cual es la probabilidad de:

- Seleccionar a una mujer embarazada o que su resultado salió positivo.
 - Seleccionar a una mujer embarazada o que su resultado salió negativo.
 - Seleccionar a una mujer embarazada o que no esté embarazada.
 - Seleccionar a una mujer embarazada y que su resultado sea positivo.
9. Una tienda vende 3 marcas distintas de celulares. El 50% de sus ventas son de la marca A, el 30% son de la marca B y el 20% son de la marca C. Se ofrece 1 año de garantía y el 25% de los aparatos de marca A regresan por reparación, el 20% de la marca B y el 10% de la marca C.
 - a) ¿Cuál es la probabilidad de que se seleccione un comprador al azar que haya comprado la marca A y necesite reparación en el periodo de garantía?
 - b) ¿Cuál es la probabilidad de que se seleccione un comprador al azar cuyo celular necesite reparación en el periodo de garantía?
 10. En una caja hay 10 focos defectuosos y 6 buenos. Se sacan dos a la vez (sin reposición). Se prueba uno de ellos y se encuentra que es bueno.
 - a) ¿Cuál es la probabilidad de que el otro sea defectuoso?
 - b) ¿Cuál es la probabilidad de que el otro también sea bueno?
 11. Considere el experimento en el que se lanzan dos dados:
 - a) ¿Cuál es la probabilidad de sacar dos números que me resulte sumados 7?
 - b) ¿Cuál es la probabilidad de que la suma de los dados resulte 10?
 - c) ¿Cuál es la probabilidad de obtener 8 puntos?
 - d) ¿Cuál es la probabilidad de que el número que aparece en un dado sea igual al otro?
 12. En una bolsa se tiene 10 caramelos de naranja y 14 de chicha morada. Si se extraen dos caramelos uno tras otro con devolución.
 - a) ¿Cuál es la probabilidad de que salgan dos caramelos de naranja?
 - b) ¿Cuál es la probabilidad de que salgan dos caramelos de chicha morada?
 - c) ¿Cuál es la probabilidad de que salgan uno de cada uno?
 13. Bitel sortea un viaje a Cancún entre sus 160 mejores clientes durante el año pasado. De ellos, 65 son mujeres, 82 están casados y 45 son mujeres casadas. Se pide:
 - a) ¿Cuál será la probabilidad de que le toque el viaje a un hombre soltero?
 - b) Si del afortunado se sabe que es casado, ¿Cuál será la probabilidad de que sea mujer?
 - c) ¿Cuál es la probabilidad de que le toque el viaje a una mujer o este casada?
 14. Calcular la probabilidad de que al arrojar al aire tres monedas, salgan:
 - a) Tres caras
 - b) Al menos dos caras
 - c) Dos caras
 - d) Tres sellos



- e) Un sello
- f) Al menos un sello

15. Se seleccionó una muestra de 300 clientes de Ripley para determinar varios aspectos acerca del comportamiento de compra de sus clientes. Se obtuvo la tabla de contingencia que se indica a continuación:

DISFRUTA COMPRANDO ROPA	VARON	MUJER	TOTAL
SI	84	118	
NO	51	36	
TOTAL			

- a) Suponga que el cliente elegido es mujer. ¿Cuál es la probabilidad de que no disfrute comprando ropa?
 - b) Suponga que el participante elegido disfruta comprando ropa. ¿Cuál es la probabilidad de que la persona sea hombre?
16. En una encuesta realizada en la provincia de Huancayo, se halló que el 40% de las personas son fumadoras de tabaco. De acuerdo a estudios anteriores, se encontró que la probabilidad de muerte causada por cáncer de pulmón en una persona fumadora de tabaco es del 38%, mientras que en una persona no fumadora esta probabilidad sólo alcanza el 9%.
- Si se elige al azar a un poblador huancaíno, ¿cuál es la probabilidad de que sea una persona que no padece de cáncer de pulmón?
 - Si se eligió a un poblador que resultó estar enfermo de cáncer al pulmón, ¿cuál es la probabilidad que no sea fumadora?
17. En la evaluación de 800 personas (entre varones y mujeres) para el ascenso de gerentes de una municipalidad se determinó que: 181 de los postulantes eran mujeres, de las cuales solo 51 lograron el ascenso, por otro lado 118 varones no lograron su ascenso. ¿Cuál es la probabilidad de que un gerente:
- A. Seleccionado al azar sea varón y fue ascendido?
 - B. Seleccionado al azar sea masculino o fue ascendido?
 - C. Seleccionado al azar sea varón y no fue ascendido?
 - D. Seleccionado al azar sea masculino o no fue ascendido?
 - E. Seleccionado al azar sea femenina y fue ascendido?
 - F. Seleccionado al azar sea mujer o fue ascendido?
 - G. Seleccionado al azar sea mujer y no fue ascendido?
 - H. Ascendido dado que es hombre.
 - I. Mujer dado que no ascendió.
 - J. Ascendido dado que es mujer.

REFERENCIAS BIBLIOGRÁFICAS Y ENLACES

- Triola Mario F. ESTADÍSTICA. Pearson Educación. México 2009



Semana N°12: PROBABILIDAD TOTAL Y TEOREMA DE BAYES

1. TEMA: Probabilidad total y teorema de bayes

2. PROPÓSITO/OBJETIVO/LOGRO/HIPÓTESIS:

- Calcula la probabilidad de eventos aleatorios aplicando la probabilidad Total y los interpreta.
- Calcula la probabilidad de eventos aleatorios aplicando el Teorema de Bayes y los interpreta.

PROBABILIDAD TOTAL

Si conocemos la probabilidad de B en cada uno de los componentes de un sistema exhaustivo y excluyente de sucesos donde participan los sucesos A_1, A_2, A_3 y A_4 , entonces podemos calcular la probabilidad de B como la suma:

$$P(B) = P(B \cap A_1) + P(B \cap A_2) + P(B \cap A_3) + P(B \cap A_4)$$

$$P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + P(B|A_3)P(A_3) + P(B|A_4)P(A_4)$$

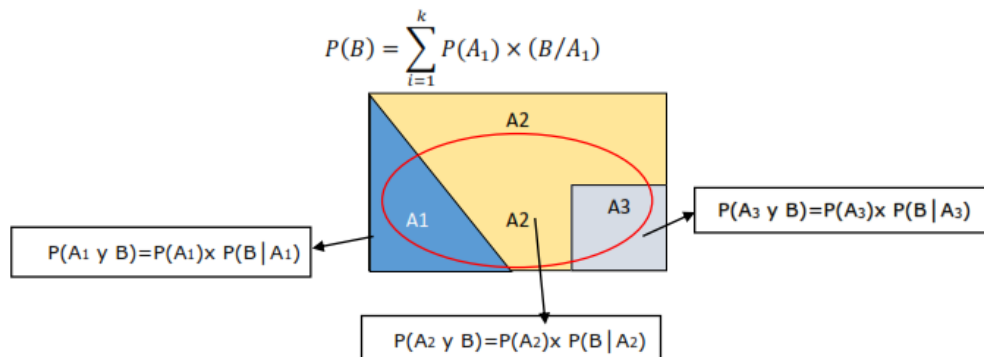
Por lo tanto si generalizamos, para la Ley de la Probabilidad Total:

Supongamos que A_1, A_2, \dots, A_n son una partición de E ,

Es decir que los sucesos son mutuamente excluyentes entre sí ($A_i \cap A_j = \emptyset$ para todo par) y su unión es E entonces se cumple

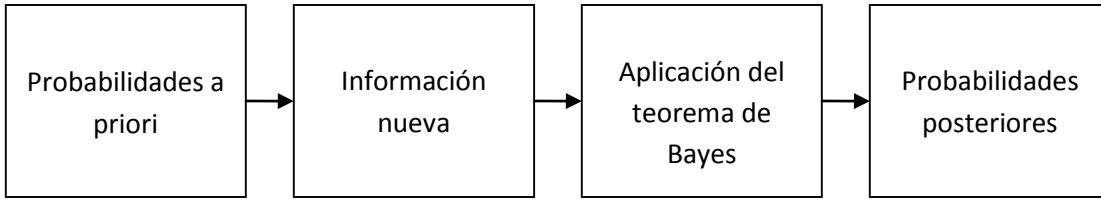
$$P(B) = \sum_{i=1}^n P(B | A_i)P(A_i) \quad P(B) = \sum_{i=1}^n P(B \cap A_i)$$

Si "k" eventos o sucesos A_1, A_2, \dots, A_k constituyen una partición del espacio muestral Ω , entonces para cualquier suceso B (B es un subconjunto de Ω) en Ω se tiene:



TEOREMA DE BAYES

Una etapa importante del análisis de las probabilidades condicionales consiste en revisar dichos valores cuando se obtiene nueva información. Con esta nueva información actualizamos los valores iniciales de probabilidad encontrando probabilidades revisadas, conocidas como *probabilidades posteriores*. El teorema de Bayes proporciona un medio para hacer estos cálculos de probabilidad. Los pasos en este proceso de revisión de probabilidades se muestran en la Figura siguiente.



Es un método que nos permite calcular la probabilidad de que un suceso que ya ocurrió (“efecto”) sea resultante de alguna “causa”.

Sean A_1, A_2, \dots, A_k un conjunto de sucesos incompatibles entre si. Sea B el suceso del cual se conocen como el teorema de Bayes y su valor se determina mediante la expresión:

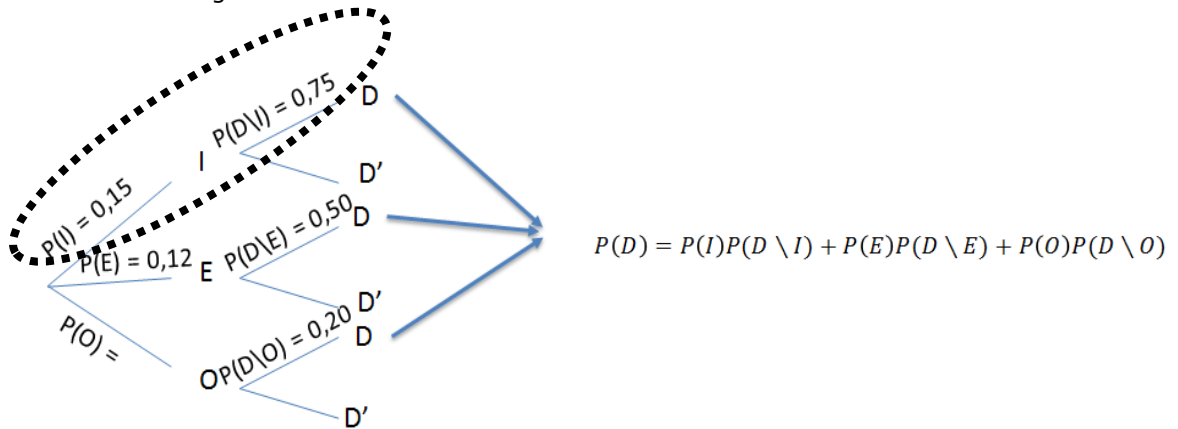
$$P(A_i|B) = \frac{P(A_i) \times P(B|A_i)}{P(A_1) \times P(B|A_1) + P(A_2) \times P(B|A_2) + P(A_3) \times P(B|A_3) + \dots + P(A_k) \times P(B|A_k)}$$

EJERCICIOS RESUELTOS

- El 15% de los empleados de una empresa son ingenieros y otro 12% son economistas. El 75% de los ingenieros ocupan un puesto directivo y el 50% de los economistas también, mientras que los no ingenieros y los no economistas solamente el 20% ocupa un puesto directivo. ¿Cuál es la probabilidad de que un empleado elegido al azar ocupe un puesto directivo?

Solución:

Hacemos un diagrama de árbol con los datos:



Entonces, multiplicamos, las tres elipses nos ayudan a reconocer lo que multiplicaremos:

$$P(D) = (0,15)(0,75) + (0,12)(0,50) + (0,73)(0,20)$$

$$P(D) = 0,1125 + 0,06 + 0,146$$

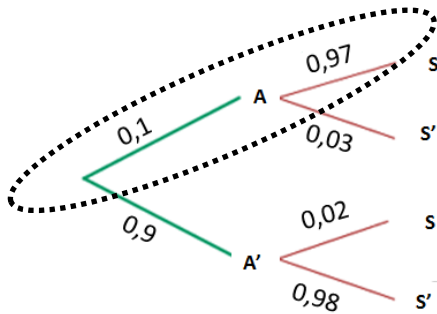
$$P(D) = \boxed{0,3185}$$

- La probabilidad de que haya un incidente en “La Grande” que dispone de alarma es de 0,1. La probabilidad de que suene esta sí se ha producido algún incidente es de 0,97 y la probabilidad de que suene si no ha sucedido ningún incidente es 0,02. En el supuesto de



que haya funcionado la alarma, ¿cuál es la probabilidad de que no haya habido ningún incidente?

Solución:



$$P(A'/S) = \frac{P(A' \cap S)}{P(S)}$$

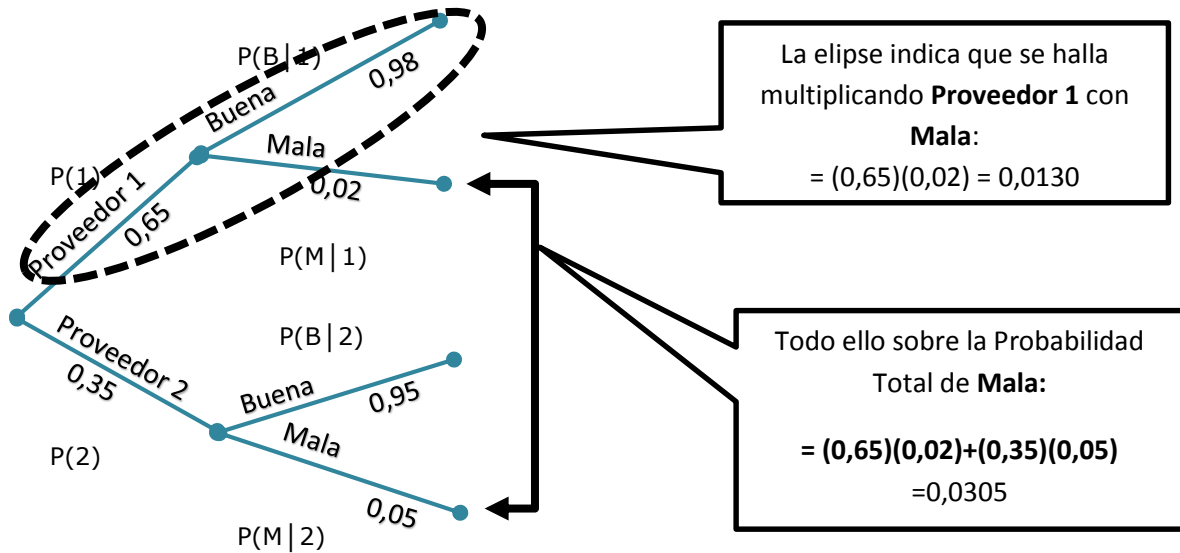
$$P(S) = (0,1)(0,97) + (0,9)(0,02) = 0,115$$

$$P(A'/S) = \frac{(0,9)(0,02)}{0,115} = 0,1565$$

3. "Maestro" recibe productos de dos proveedores. El 65% de los productos adquiridas provienen del proveedor 1 y resto del porcentaje faltante del proveedor 2. Según la estadística existente, de las 100 últimos productos recibidos del proveedor 1. Estuvieron buenas 98 y 2 estuvieron malas. En el caso del proveedor 2, de 100 productos 95 estuvieron buenas y el resto malas. Suponga que una de los productos recibidos de los dos proveedores esta malo, ¿Cuál es la probabilidad que provenga del proveedor 1?

Solución:

Podemos identificar que se trata de teorema de Bayes porque la interrogante es una condicional. Luego, elaboramos el diagrama del árbol:



Entonces, tenemos:

$$P(1|M) = \frac{P(1, M)}{P(M)} = \frac{0,0130}{0,0305} = 0,4262$$

EJERCICIOS PROPUESTOS:



1. Tres máquinas A, B, C producen el 45%, 30% y 25% respectivamente, del total de las piezas producidas para la empresa "Toyota". Los porcentajes de producción defectuosa de estas máquinas son del 3%, 4% y 5%.
 - a) Seleccionamos una pieza al azar, calcula la probabilidad de que sea defectuosa
 - b) Tomamos, al azar, una pieza y resulta ser defectuosa, calcula la probabilidad de haber sido producida por la máquina B
 - c) ¿Qué máquina tiene la mayor probabilidad de haber producido la citada pieza defectuosa?
2. La empresa "Sazón Lopesa" distribuye sus productos a cuatro mayoristas. El porcentaje de producción total que se obtiene en cada mayorista es del 30%, 40%, 10% y 20% respectivamente, y además el porcentaje de envasado incorrecto en cada mayorista es del 2%, 1%, 7% y 4%. Tomamos un producto de la empresa al azar en almacén principal.
 - a) ¿Cuál es la probabilidad de que se encuentre defectuosamente envasado?
 - b) ¿Cuál es la probabilidad de que se encuentre bueno?
3. En un laboratorio hay tres jaulas. En la jaula I hay 4 conejos negros y 2 blancos, la jaula II tiene 4 conejos negros y 4 blancos y la jaula III contiene 5 conejos negros y 6 blancos. Se selecciona al azar una de las jaulas y se saca un conejo al azar de esa jaula. ¿Cuál es la probabilidad que el conejo escogido sea blanco?
4. El 70% del ganado es inyectado con una vacuna para combatir una enfermedad grave. La probabilidad de recuperarse de la enfermedad es 1 en 20 si no ha habido tratamiento y de 1 en 5 si hubo tratamiento. Si un animal infectado se recupera, ¿Cuál es la probabilidad de que haya recibido la vacuna preventiva?
5. Una unidad de producción de enlatados de atún marca "Fanny" produce 5000 envases diarios. La máquina A produce 3000 de estos envases, de los que 2% son defectuosos y la máquina B produce el resto, se sabe que de la máquina B, el 4% son defectuosos. Calcular:
 - a) La probabilidad de que un envase elegido al azar este defectuoso.
 - b) La probabilidad de que proceda de la máquina B, si el envase tiene defectos.
 - c) La probabilidad de que proceda de la máquina A, si el envase tiene defectos.
6. De 300 estudiantes de Ciencias de la Empresa, 100 cursan Finanzas y 80 cursan Economía. Estas cifras incluyen 30 estudiantes que cursan ambas materias.
 - a) ¿Cuál es la probabilidad de que un estudiante elegido aleatoriamente curse Finanzas o Economía?
 - b) ¿Qué probabilidad hay de que al elegir un estudiante al azar curse Economía, dado que cursa Finanzas?
7. En la Universidad Continental se selecciona tres carreras que pueden cursarse del siguiente modo: el 20% estudian economía, el 35% contabilidad y el 45% administración. El porcentaje de alumnos que finalizan sus estudios en cada caso es del 5%, 12% y del 18% respectivamente. Elegido un alumno al azar determinar la probabilidad de que haya acabado sus estudios.
8. Un profesor de Estadística sabe que el 80% de sus estudiantes terminan los problemas asignados como tarea. Determinó además, que de los estudiantes que cumplen con la tarea, el 90% aprobará el curso. De aquellos estudiantes que no terminan su tarea, el 60%



aprobará. Si un alumno recibió una calificación aprobatoria, ¿Cuál es la probabilidad de que sí haya hecho sus tareas?

9. Suponga que GESA adquiere sus productos de tres proveedores: A, B y C. El proveedor A suministra el 60% de sus productos, el proveedor B el 30% y el proveedor C el 10%. La calidad de sus productos varía entre los proveedores, siendo las tasas defectuosas del proveedor A 0,25%, del proveedor B 1% y del proveedor C 2%.
 - a) ¿Cuál es la probabilidad de que el producto sea defectuoso?
 - b) Cuando se encuentre un producto defectuoso. ¿Cuál será el proveedor más probable?
10. Los registros de los delitos en la Comisaría de Huancayo muestran que el 20% de ellos son violentos y el 80% son no violentos. Se señala también que son denunciados, el 90% de los delitos violentos y solo el 70% de los delitos no violentos.
 - a) ¿Cuál es la proporción global de delitos que se denuncian?
 - b) Si no se denuncia un delito ante la policía, ¿Cuál es la probabilidad de que el delito sea no violento?

REFERENCIAS BIBLIOGRÁFICAS Y ENLACES

- Triola Mario F. ESTADÍSTICA. Pearson Educación. México 2009



GUÍA DE PRÁCTICA DE PROBABILIDADES Y ESTADÍSTICA

CUARTA UNIDAD DISTRIBUCIONES DE PROBABILIDAD

- GUÍA DE PRÁCTICA N° 13: Variables aleatorias
- GUÍA DE PRÁCTICA N° 14: Distribuciones de probabilidad discretas
- GUÍA DE PRÁCTICA N° 15: Distribuciones de probabilidad continuas
- GUÍA DE PRÁCTICA N° 16: Aplicaciones de la distribución normal



Semana N°13: VARIABLES ALEATORIAS

1. TEMA: Variables aleatorias

2. PROPÓSITO:

- Identifica el tipo de variable aleatoria y construye la distribución de la función de probabilidad.

VARIABLES ALEATORIAS

1. Definiciones

Una **variable aleatoria** es aquella (casi siempre representada por x) que tiene un solo valor numérico determinado por el azar, para cada resultado de un procedimiento.

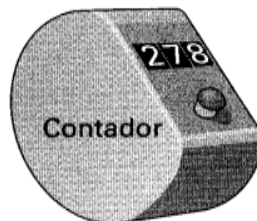
Una **distribución de probabilidad** es una distribución que indica la probabilidad de cada valor de la variable aleatoria. A menudo se expresa como gráfica, tabla o fórmula.

VARIABLE ALEATORIA DISCRETA

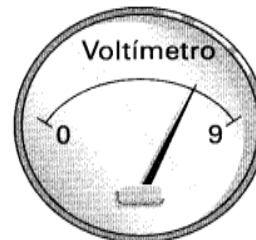
Una variable aleatoria discreta tiene un número finito de valores o un número de valores contable, donde "contable" se refiere al hecho de que podría haber un número infinito de valores, pero que pueden asociarse con un proceso de conteo.

VARIABLE ALEATORIA CONTINUA

Una variable aleatoria continua tiene un número infinito de valores, y esos valores pueden asociarse con mediciones en una escala continua, de manera que no existan huecos o interrupciones.



a) Variable aleatoria discreta: contador del número de asistentes al cine.



b) Variable aleatoria continua: voltaje medido de una batería de un detector de humo.

EJEMPLO Los siguientes son ejemplos de variables aleatorias discretas y continuas:

1. Sea x = número de huevos que una gallina pone en un día. Ésta es una variable aleatoria *discreta* porque sus únicos valores posibles son 0 o 1 o 2, etcétera. Ninguna gallina puede poner 2,343115 huevos, lo que sería posible si los datos provinieran de una escala continua.
2. El conteo del número de estudiantes de estadística que asisten a una clase es un número entero y, por lo tanto, una variable aleatoria discreta. a) es capaz de indicar únicamente



un número finito de valores, por lo que se utiliza para obtener valores de una variable aleatoria *discreta*.

3. Sea x = cantidad de leche que produce una vaca en un día. Ésta es una variable aleatoria *continua*, ya que puede tomar cualquier valor en un tramo continuo. En un solo día, una vaca produce una cantidad de leche cuyo valor puede ser cualquiera entre 0 galones y 5 galones. Es posible obtener 4,123456 galones, ya que la vaca no está restringida a las cantidades discretas de 0; 1; 2; 3; 4 o 5 galones.
4. La medida del voltaje de una batería de un detector de humo puede ser cualquier valor entre 0 y 9 volts. Por lo tanto, se trata de una variable aleatoria continua. El voltímetro indica valores en una escala continua, de manera que permite obtener valores de una variable aleatoria *continua*.

2. Requisitos de una distribución de probabilidad

1. $\sum P(x) = 1$ donde x asume todos los valores posibles. (Es decir, la suma de todas las probabilidades debe ser 1).

El primer requisito surge del simple hecho de que la variable aleatoria x representa todos los sucesos posibles en el espacio muestral completo, de manera que tenemos la certeza (con probabilidad 1) de que uno de los sucesos ocurrirá.

2. ~~$0 \leq P(x) \leq 1$ para cada valor individual de x . (Es decir, cada valor de probabilidad debe ubicarse entre 0 y 1, inclusive).~~

Asimismo, la regla de probabilidad que establece que $0 \leq P(x) \leq 1$ para cualquier suceso A , implica que $P(x)$ debe estar entre 0 y 1 para cualquier valor de x .

EJEMPLO ¿La tabla describe una distribución de probabilidad?

Probabilidades de una variable aleatoria	
X	P(x)
0	0,2
1	0,5
2	0,4
3	0,3

Solución: Para ser una distribución de probabilidad, $P(x)$ debe satisfacer los dos requisitos anteriores.

$$\begin{aligned} \sum P(x) &= P(0) + P(1) + P(2) + P(3) \\ &= 0,2 + 0,5 + 0,4 + 0,3 \\ &= 1,4 \text{ (lo que demuestra que } \sum P(x) \neq 1) \end{aligned}$$

Pero como no se satisface el primer requisito, concluimos que la tabla no describe una distribución de probabilidad.

3. Cálculo de la Media, Varianza y Desviación Estándar

A menudo podemos calcular la media, la varianza y la desviación estándar de los datos, los cuales brindan información acerca de otras características. La media, la varianza y la desviación estándar de una distribución de probabilidad se calcula al aplicar las siguientes fórmulas:

- **Fórmula 1** $\mu = \sum [x \cdot P(x)]$ media de una distribución de probabilidad



- **Fórmula 2** $\sigma^2 = \sum[(x - \mu)^2 \cdot P(x)]$ varianza de una distribución de probabilidad
- **Fórmula 3** $\sigma^2 = \sum[x^2 \cdot P(x)] - \mu^2$ varianza de una distribución de probabilidad
- **Fórmula 4** $\sigma = \sqrt{\sum[x^2 \cdot P(x)] - \mu^2}$ desviación estándar de una distribución de probabilidad

Precaución: Evalúe $\sum[x^2 \cdot P(x)]$ elevando al cuadrado primero cada valor de x , multiplicando después cada cuadrado por la probabilidad $P(x)$ correspondiente y después sumando.

EJERCICIOS RESUELTOS

1. El gerente de personal de la empresa de taxi "ARCHI", estudia el número de accidentes ocurridos durante un mes. Elaboró la siguiente distribución probabilística. ¿Cuántos accidentes se espera tener en un mes?

Número de accidentes	0	1	2	3	4
Probabilidad	0,30	x	0,2	0,08	0,02

Solución:

- Si queremos saber, cuántos accidentes se espera tener en un mes en dicha empresa, necesitamos calcular el valor esperado, media o también llamado esperanza matemática.
- Para calcular el valor de "X". En la distribución de la probabilidad la suma de todas sus probabilidades siempre es a la unidad o al 100%. Entonces haciendo una suma e igualando ello al 100% tenemos el valor de $x = 40\%$
- Finalmente, hallamos el valor esperado:

$$\mu = \sum[x \cdot P(x)] = (0 \times 0,3) + (1 \times 0,4) + (2 \times 0,2) + (3 \times 0,08) + (4 \times 0,02) = 1,12 = 1$$

Interpretación: En dicha empresa se espera tener un accidente al mes. Para entender mejor presentamos una tabla.

X	P(X)	P(X) * X
0	0,3	0
1	0,4	0,4
2	0,2	0,4
3	0,08	0,24
4	0,02	0,08
	1	1,12



Nota: La esperanza matemática es 1,12 entonces, se redondea por ser una variable discreta, para poder interpretar, entonces diremos que se espera tener un accidente al mes en dicha empresa.

EJERCICIOS PROPUESTOS:

1. Si lanzamos dos monedas, cual es la distribución de la probabilidad, la esperanza matemática o media, varianza y desviación estándar. Si analizamos la ocurrencia de las caras.
2. Expresar $P(X) = 1/6$, para $X = 1;2;3;4;5$ ó 6 en forma de distribución. Calcular la media, varianza y desviación típica.
3. ¿Es $P(X) = X/10$, para $X = 1;2;3$ ó 4 , una función de probabilidad? Graficar un histograma; Calcular el valor esperado, varianza y desviación estándar.
4. Los datos del censo a menudo se utilizan para obtener distribuciones de probabilidades de algunas variables aleatorias. Los datos del censo en el Perú, de familias con un ingreso combinando de 5000 o más, en un departamento en particular, muestra que el 20% no tiene hijos, 30% tienen un hijo, 40% tienen dos hijos y 10% tienen tres hijos. Con base a esta información, elaborar la distribución de probabilidad, calcular la esperanza matemática, varianza y desviación estándar.
5. Compruebe si la siguiente función es una función de probabilidad. $P(X) = (X^2 + 5)/50$, para $X = 1,2,3$ ó 4 . Enumere la distribución de probabilidades, graficar un polígono de frecuencias, mencione cuáles son el valor esperado, varianza y desviación estándar.
6. Explique si es una distribución de probabilidad la siguiente tabla:

Edad	18	20	21	22	23	24
Porcentaje que fuman	1,7	4,9	8,9	16,3	25,2	37,2

7. Asumiendo que un dado se lanza en un juego de azar no está cargado ni quedara de canto, calcular la probabilidad:
 - A. De cada resultado posible calcular la media aritmética, la varianza y la desviación estándar de la variable aleatoria y graficar la distribución de la probabilidad.
 - B. Que el resultado sea un número impar.
 - C. Que el resultado sea un número menor igual a dos.
 - D. Que el resultado sea un número mayor que seis.
8. La variable aleatoria A tiene la siguiente probabilidad de distribución

A	1	2	3	4	5	Total
P(A)	0.6	0.1	x	0.1	0.1	

- A. Encuentre la media y la desviación estándar de A.
- B. ¿Cuánto de esta distribución de probabilidad se encuentra a menos de dos desviaciones estándar de la media?



9. El jefe personal de la empresa "CHIFA CENTRO" estudia el número de accidentes en la cocina ocurridos durante un semestre. Elaboró la siguiente distribución probabilística. Calcule la media, la varianza y la desviación estándar del número de accidentes en dicho período.

Número de accidentes	0	1	2	3	4
Probabilidad	0,40	0,20	0,20	0,10	0,10

10. El número de llamadas, x , que llegan al centro de servicio al cliente de la empresa "BITEL" durante cualquier periodo de un minuto es una variable aleatoria con la siguiente probabilidad.

X	0	1	2	3	4
P(x)	0.1	0.2	0.4	0.2	0.1

- Encuentre la media y la desviación estándar de x . y elabore el histograma de $p(x)$

REFERENCIAS BIBLIOGRÁFICAS Y ENLACES

- Triola Mario F. ESTADÍSTICA. Pearson Educación. México 2009



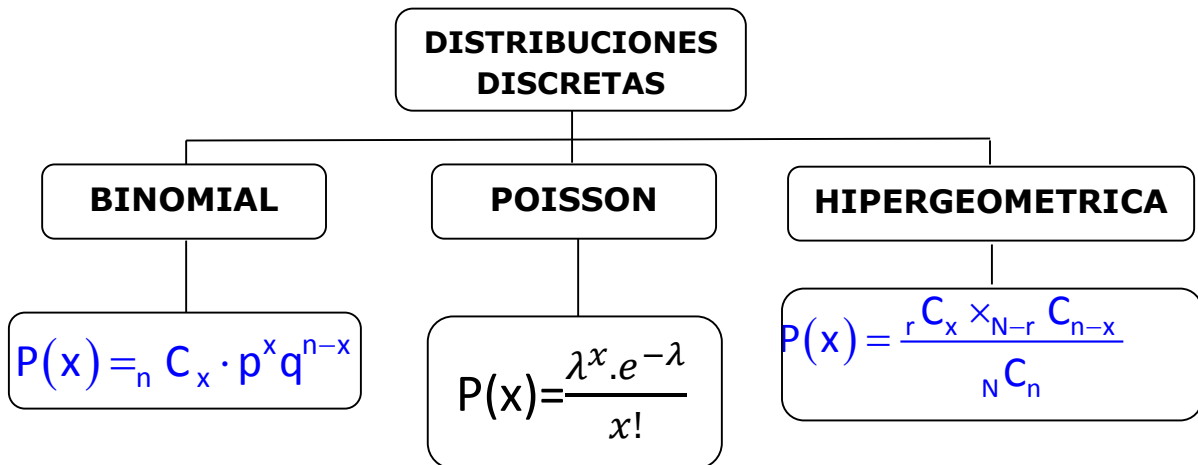
Semana N°14: DISTRIBUCIONES DE PROBABILIDAD DISCRETAS

1. TEMA: Distribuciones de probabilidad discretas

2. PROPÓSITO:

- Explica y diferencia las principales distribuciones de probabilidad para variables aleatorias discretas, como son: la distribución binomial, hipergeométrica, geométrica y poisson; identificando las condiciones que se debe de tomar en cuenta para su aplicación.
- Aplica e interpreta las distribuciones de probabilidades para variables aleatorias discretas en el desarrollo de prácticas y ejercicios.

DISTRIBUCIONES DISCRETAS



DISTRIBUCION BINOMIAL

Una distribución de probabilidad binomial resulta de un procedimiento que cumple con todos los siguientes requisitos:

1. El procedimiento tiene un número fijo de ensayos.
2. Los ensayos deben ser independientes. (El resultado de cualquier ensayo individual no afecta las probabilidades de los demás ensayos).
3. Todos los resultados de cada ensayo deben estar clasificados en dos categorías (generalmente llamadas éxito y fracaso).
4. La probabilidad de un éxito permanece igual en todos los ensayos.

FÓRMULA:

$$P(x) = {}_n C_x \cdot p^x q^{n-x}$$

n: es el número de pruebas



- k:** es el número de éxitos
- p:** es la probabilidad de éxito
- q:** es la probabilidad de fracaso

El número combinatorio:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

LA MEDIA, VARIANZA Y DESVIACION ESTANDAR:

La media, varianza y desviación estándar de la variable aleatoria binomial son, respectivamente:

$$\mu = np$$

$$\sigma^2 = npq$$

$$\sigma = \sqrt{npq}$$

DISTRIBUCIONES DE POISSON

Una distribución de probabilidad Poisson resulta de un procedimiento que cumple con todos los siguientes requisitos:

1. El experimento consiste en contar el número "x" de veces que ocurre un evento en particular durante una unidad de tiempo dada, o en un área o volumen dado.
2. La probabilidad de que un evento ocurra en una unidad dada de tiempo, área o volumen es la misma para todas las unidades.
3. El número de eventos que ocurren en una unidad de tiempo, área o volumen es independiente del número de los que ocurren en otras unidades.
4. El número medio (o esperado) de eventos en cada unidad se denota por la letra griega λ ("lambda")

FÓRMULA:

$$P(x) = \frac{\lambda^x \cdot e^{-\lambda}}{x!} ; e = 2,7183$$

λ : denota el número medio de eventos que ocurren en una unidad dada de tiempo, área o volumen (promedio=μ).

x: denota un número específico de eventos que ocurren durante una unidad dada de tiempo, área o volumen (variable aleatoria).

e: denota la constante matemática "épsilon"

x!: denota el factorial de "x"

P(x): denota la probabilidad de que ocurran "x" eventos en una unidad dada de tiempo, área o volumen.

LA MEDIA Y VARIANZA:



La media y varianza de la variable aleatoria Poisson son respectivamente:

$$\mu = \lambda \quad ; \quad \sigma^2 = \lambda$$

DISTRIBUCIÓN HIPERGEOMÉTRICA

Una distribución de probabilidad hipergeométrica resulta de un procedimiento que cumple con todos los siguientes requisitos:

1. El procedimiento tiene un número fijo de ensayos.
2. Los ensayos deben ser dependientes. (El resultado de cualquier ensayo individual sí afecta las probabilidades de los demás ensayos).
3. Todos los resultados de cada ensayo deben estar clasificados en dos categorías (generalmente llamadas éxito y fracaso).
4. La probabilidad de un éxito no permanece igual en todos los ensayos.

FÓRMULA:

$$P(x) = \frac{{}_r C_x \times {}_{N-r} C_{n-x}}{{}_N C_n} \quad ; \quad (x = 0; 1; 2; \dots; n)$$

E y F (éxito y fracaso) denotan las dos categorías posibles de todos los resultados:

r :denota el número fijo de éxitos en la población.

x :denota un número específico de éxitos en la muestra de tamaño " n ", de manera que " x " puede ser cualquier número entero entre 0 y " n " inclusive.

N :denota el número de elementos de la población.

n : denota el número de elementos de la muestra.

$P(x)$: denota la probabilidad de obtener exactamente " x " éxitos al extraer la muestra de tamaño " n ".

LA MEDIA, VARIANZA:

La media y varianza de la variable aleatoria hipergeométrica son, respectivamente:

$$\mu = np \quad \left(p = \frac{r}{N} ; q = 1 - p \right) \quad ; \quad \sigma^2 = npq \frac{N-n}{N-1}$$

EJERCICIOS RESUELTOS

Distribución binomial:



1. La probabilidad de que un cliente potencial de la pastelería "PARIS", elegido al azar, realice una compra es de 0,20. Si su agente de ventas visita a 6 clientes, ¿cuál es la probabilidad de que realice exactamente 4 ventas?

Éxito	Cliente que realiza la compra
Probabilidad de éxito (p)	0,20
Probabilidad de fracaso (q)	0,80
Número de ensayos (n)	6
Número de éxitos en los "n" ensayos (x)	4

$$P(x) = {}_n C_x \cdot p^x q^{n-x}$$

Solución:

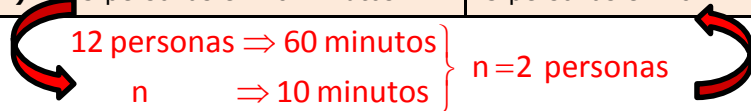
$$P(x = 4) = {}_6 C_4 \cdot (0,20)^4 (0,80)^{6-4}$$

$$P(x = 4) = 0,0154$$

Distribución de Poisson:

1. En tiendas "EFE", en la sección de electrodomésticos, un promedio de 12 personas por hora le hacen preguntas al encargado. ¿Cuál es la probabilidad de que exactamente 3 personas se acerquen al encargado a hacer preguntas en un periodo de 10 minutos?

Media (λ)	12 personas en una hora	2 personas en 10 minutos
Número de eventos esperados (x)	3 personas en 10 minutos	3 personas en 10 minutos



Solución:

$$P(X = 3) = \frac{5^3 \cdot e^{-5}}{3!}$$

$$P(X = 3) = 0,1404$$

Distribución Hipergeométrica:

1. En la tienda "PROMART", se tiene una caja de 10 bombillas la cual contiene 2 defectuosas y 8 no defectuosas. Si se eligen al azar y sin reposición 3 bombillas de la caja, ¿cuál es la probabilidad de que la muestra contenga exactamente una bombilla defectuosa?

Éxito	Bombilla defectuosa
Número de elementos de la población (N)	10
Número de elementos de la muestra (n)	3



Éxitos en la población (r)	2
Éxitos en la muestra (x)	1

Solución:

$$P(x) = \frac{{}^r C_x \times {}^{N-r} C_{n-x}}{{}^N C_n}$$

$$P(X = 1) = \frac{{}^2 C_1 \times {}^{10-2} C_{3-1}}{{}^{10} C_3} = 0,4667$$

EJERCICIOS PROPUESTOS

Distribución binomial:

1. Se tiene información que en la comisaria de Huancayo, el 25% de los detenidos se encuentran allí por delitos menores. Se selecciona una muestra aleatoria de 20. ¿Cuál es la probabilidad de que el número de detenidos por delitos menores en esa muestra sea 4?
2. Se sabe que en el Centro de Huancayo el 70% de los establecimientos comerciales pequeños no entregaban boleta al momento de efectuar una transacción económica. ¿Cuál es la probabilidad de que en una muestra de 15 negocios, a lo más la mitad entregue boleta?
3. La probabilidad de que una persona se muera por consumir bebidas alcohólicas es 0.4. Después de un estudio realizado en un sector de Huancayo, se conoce que a la semana 30 personas consumen bebidas alcohólicas, cual es la probabilidad que:
 - a) Sobrevivan al menos 14 personas
 - b) Sobrevivan exactamente 5
 - c) Cuantos, en promedio, sobrevivirán
 - d) ¿Cuál es la varianza de la v.a. en cuestión?
4. La encargada de los préstamos del banco "BBVA CONTINENTAL" estima, con base en sus años de experiencia, que la probabilidad de que un solicitante no sea capaz de pagar su préstamo es 0.035. El mes pasado realizó 45 préstamos. ¿Cuál es la probabilidad de que 3 préstamos no sean pagados oportunamente?
5. En "Western Union" se evalúa el número de envíos de dinero al extranjero que se realizó en un mes determinado. Suponga que 45% de todas sus transacciones fueron envíos de dinero al extranjero y considere una muestra de 50 transacciones. ¿Cuál es la probabilidad de que:
 - a) Cuando mucho 6 de las transacciones sean envíos de dinero al extranjero?
 - b) Por lo menos 6 de las transacciones sean envíos de dinero al extranjero?
 - c) Más de 6 de las transacciones sean un envío de dinero al extranjero?
6. "CAJA CENTRO" informó que el 30% de los préstamos garantizados con joyería vencieron. Si se toma una muestra aleatoria de 4 préstamos, ¿cuál es la probabilidad de que ninguno esté vencido?
7. El 18% de los tornillos de la última producción para la empresa "MAESTRO" están defectuosos. Determinar la probabilidad de que de 4 tornillos seleccionados aleatoriamente por lo menos 3 estén defectuosos.

Distribución de Poisson:

8. El número promedio de Buses que llegan cada día al Terminal Terrestre de Huancayo es 11. Las facilidades que tiene el Terminal hace a que pueden manejar hasta 22 buses por día. ¿Cuál es la probabilidad de que en un día dado se tenga que rechazar el ingreso de buses al Terminal Terrestre de Huancayo?
9. En Caja Centro se analiza la cantidad de clientes que se atiende en las ventanillas, se obtuvo que la cantidad de clientes promedio que se atiende en esta institución financiera en un lapso



- de 30 minutos es de 10 clientes. ¿Cuál es la probabilidad de que se atiendan a 7 clientes en 20 minutos?
10. En una investigación realizada en Huancayo, se ha determinado que el número de muertes debido a accidentes de tránsito ha sido en promedio 4 por en dos meses. Suponiendo que el número de muertes sigue una distribución de Poisson, hallar la probabilidad de que en un determinado mes mueran 3 personas.
 11. En la oficina de ventas de "RIPLEY" se analiza el número de ventas al día, se tiene la información que tiene un promedio de 9,5 ventas al día. Suponiendo que los pedidos siguen una distribución de Poisson, ¿cuál es la probabilidad de que se realicen menos de 5 pedidos en un día determinado?
 12. Un libro de 450 páginas de la biblioteca tiene 199 errores de impresión distribuidos aleatoriamente. Calcule la probabilidad de que cualquier página elegida al azar tenga un error.

Distribución Hipergeométrica:

13. En un salón de clase de Estadística de la Universidad Continental conformado por 45 alumnos, se ha observado que 11 de ellos están con la gripe. Si se extrae un grupo de 25 alumnos sin remplazo y al azar. ¿Cuál es la probabilidad que ninguno esté con la gripe?
14. En una encuesta realizada durante la semana de olimpiadas internas de la Universidad Continental, se observó que el fútbol y voleibol eran los deportes preferidos. Si en un grupo de 45 alumnos, 17 prefieren fútbol y 28 prefieren voleibol, se toma una muestra aleatoria de 5 de estas personas:
 - a) ¿Cuál es la probabilidad de que exactamente 2 prefieran el fútbol?
 - b) ¿De que la mayoría prefiera fútbol?
15. En el cafetín de la Universidad Continental se tiene en el refrigerador 15 botellas de Coca Cola y 10 de Inca Cola. Luego de una clase en el taller de Basquetbol un grupo de 4 amigas piden gaseosas heladas. ¿Cuál es la probabilidad de que todas las gaseosas atendidas sean Coca Cola?, ¿Cuál es probabilidad de que 3 de ellas sean Inca Cola?
16. Para la construcción de un departamento en San Carlos se contrata a 12 obreros. De estos obreros, 5 cuentan con equipos de protección personal durante toda la jornada laboral. Si de este grupo de 14 se elige aleatoriamente una cuadrilla de 5 obreros, ¿qué probabilidad hay de que a lo más 3 cuenten con equipos de protección personal?

REFERENCIAS BIBLIOGRÁFICAS Y ENLACES

- Triola Mario F. ESTADÍSTICA. Pearson Educación. México 2009
- Robert Johnson, Patricia Kuby. Estadística Elemental. Thomson Editorial. México 2002



Semana N°15: DISTRIBUCIONES DE PROBABILIDAD CONTINUAS

1. TEMA: Distribución normal estándar.

2. PROPÓSITO:

- Identifica las principales características de la curva normal estándar.
- Calcula la estandarización y ubica correctamente el área bajo la curva normal estándar.

DISTRIBUCIÓN NORMAL ESTÁNDAR

La distribución normal o distribución gaussiana es una de las distribuciones teóricas mejor estudiadas y más utilizadas en los diferentes campos de la investigación. Carl Gauss formuló la ecuación de la curva normal, de ahí que también se le conoce, más comúnmente, como la "Campana de Gauss".

La importancia de esta distribución radica fundamentalmente por la frecuencia con la que muchas variables asociadas a fenómenos naturales y cotidianos se asemejan a esta distribución. Entre ellas se tienen:

- Los caracteres morfológicos de las personas, animales o plantas, como por ejemplo: las tallas, pesos, envergaduras, diámetros, perímetros, etc.
- Los caracteres fisiológicos, como por ejemplo el efecto de una misma dosis de un fármaco a un grupo de persona, o de una misma cantidad de abono o fertilizante aun conjunto de plantas.
- Los caracteres sociológicos, como por ejemplo el consumo de cierto producto por un mismo grupo de individuos, o las puntuaciones por un grupo de estudiantes en una evaluación.
- Los caracteres psicológicos, como por ejemplo el cociente intelectual (CI) de una muestra de niños, el grado de adaptación de las persona a un medio, etc.
- Los errores que se comenten al realizar mediciones de ciertas magnitudes físicas o químicas.
- Los valores estadísticos muestrales, como por ejemplo la media aritmética, la mediana, la moda, la varianza, la desviación estándar, etc.
- El comportamiento de muchos parámetros en el campo de la salud también pueden ser descritos mediante una distribución normal.



Carl Friedrich Gauss
(1777 – 1855)

Definición: Una variable aleatoria continua tiene una **distribución uniforme** si sus valores se dispersan uniformemente a través del rango de posibilidades. La gráfica de una distribución uniforme tiene forma rectangular.

Definición: Una **curva de densidad** es una gráfica de una distribución de probabilidad continua. Debe satisfacer las siguientes propiedades:

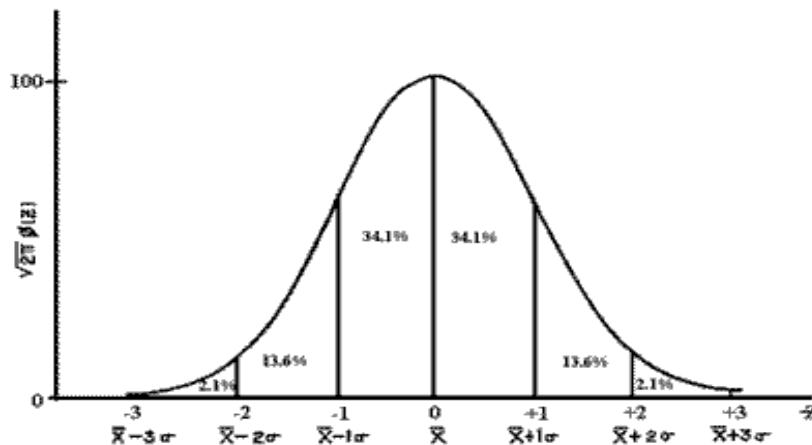
1. El área total de la curva debe ser igual a 1
2. Cada punto de la curva debe tener una altura vertical igual o mayor que 0. (es decir, la curva no puede estar por debajo del eje x).

Si una variable aleatoria continua tiene una distribución con una gráfica simétrica y una forma de campana y puede expresarse por medio de la fórmula que se muestra a continuación se dice que se tiene una distribución normal.

$$y = \frac{e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}}$$



Triola, M. (2009) sostiene que: **La distribución normal estándar** es una distribución normal de probabilidad con $\mu=0$ y $\sigma=1$, y el área total debajo de su curva de densidad es igual a 1. (p.249), como se muestra en el siguiente gráfico:



a) Propiedades de la distribución normal

La distribución normal posee las siguientes propiedades:

- Tiene una única moda, que coincide con su media aritmética y su mediana.
- La curva normal es asintótica al eje de las abscisas. Por ello, cualquier valor entre $-\infty$ y $+\infty$ es teóricamente posible. El área total bajo la curva es igual a 1.
- Es simétrica con respecto a su media μ , según esto, para este tipo de variables existe una probabilidad de un 50% de observar un dato mayor que la media, y un 50% de observar un dato menor.
- La distancia entre la línea trazada en la media y el punto de inflexión de la curva es igual a una desviación estándar (σ). Cuanto mayor sea σ , más aplanada será la curva de la densidad.
- El área bajo la curva comprendida entre los valores situados aproximadamente a dos desviaciones estándar de la media es igual a 0,95. Es decir existe un 95% de posibilidades de observar un valor comprendido en el intervalo $(\mu -1,96\sigma ; \mu +1,96\sigma)$.
- La forma de la campana de Gauss depende de los parámetros μ y σ .

A partir de las propiedades ya explicadas se afirma que no existe una única distribución normal, sino una familia de distribuciones con una forma común, diferenciadas por los valores de su media y su varianza. De entre todas ellas, la más utilizada es la **distribución normal estándar**, que corresponde a una distribución de media 0 y varianza 1

PROPIEDADES DE LA DISTRIBUCIÓN NORMAL ESTANDAR

1. El área total bajo la curva normal es igual a 1.
2. La distribución tiene forma de montículo y es simétrica; se extiende indefinidamente en ambas direcciones, tendiendo al eje horizontal pero sin tocarlo.
3. La distribución tiene una media de 0 y una desviación estándar de 1.
4. La media divide el área a la mitad; 0,50 a cada lado.
5. Casi toda el área esta entre $z=-3,00$ y $z=3,00$



b) Estandarización "z" o calificaciones estándar

La estandarización de la variable x permite la transformación hacia las calificaciones "z". La estandarización transforma cualquier variable x en una nueva variable "z", que tiene media cero y desviación uno.

$$z = \frac{x - \mu}{\sigma}$$

Propiedades de las calificaciones z

- La suma de las calificaciones z es cero: $\sum z = 0$
- La media de las calificaciones z es cero: $\bar{z} = 0$
- La suma de los cuadrados de las calificaciones "z" es igual a N: $\sum z^2 = N$
- La desviación estándar y la varianza de las calificaciones z es uno: $\sigma^2 = \sigma = 1$

Notación:

$P(a < z < b)$ denota la probabilidad de que la puntuación z este entre a y b.

$P(z > a)$ denota la probabilidad de que la puntuación z sea mayor que a.

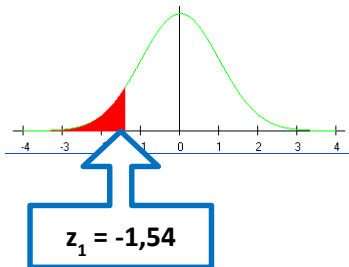
$P(z < a)$ denota la probabilidad de que la puntuación z este entre a y b.

c) Calculo de áreas bajo la curva normal estándar

Primer caso: A la izquierda de un "z" determinado: Se calcula **directo de la tabla** de áreas bajo la curva normal tipificada.

Calcule el área bajo la curva normal a la izquierda de

$z_1 = -1,54$



RESPUESTA:

0,0618

LECTURA DE LA TABLA

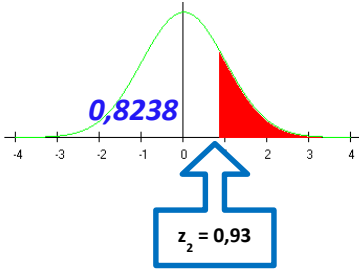
z	.00	.01	.02	.03	.04
-3.50 y menores	.0001				
-3.4	.0003	.0003	.0003	.0003	.0003
-3.3	.0005	.0005	.0005	.0004	.0004
-3.2	.0007	.0007	.0006	.0006	.0006
-3.1	.0010	.0009	.0009	.0009	.0008
-3.0	.0013	.0013	.0013	.0012	.0012
-2.9	.0019	.0018	.0018	.0017	.0016
-2.8	.0026	.0025	.0024	.0023	.0023
-2.7	.0035	.0034	.0033	.0032	.0031
-2.6	.0047	.0045	.0044	.0043	.0041
-2.5	.0062	.0060	.0059	.0057	.0055
-2.4	.0082	.0080	.0078	.0075	.0073
-2.3	.0107	.0104	.0102	.0099	.0096
-2.2	.0139	.0136	.0132	.0129	.0125
-2.1	.0179	.0174	.0170	.0166	.0162
-2.0	.0228	.0222	.0217	.0212	.0207
-1.9	.0287	.0281	.0274	.0268	.0262
-1.8	.0359	.0351	.0344	.0336	.0329
-1.7	.0446	.0436	.0427	.0418	.0409
-1.6	.0548	.0537	.0526	.0516	.0505
-1.5	.0668	.0655	.0643	.0630	.0618
-1.4	.0808	.0793	.0778	.0764	.0749



Fuente: Baldeón Crisóstomo José

Segundo caso: A la derecha de un "z" determinado: Se calcula **por su complementario**.

Calcule el área bajo la curva normal a la derecha de $z_2 = 0,93$



LECTURA DE LA TABLA

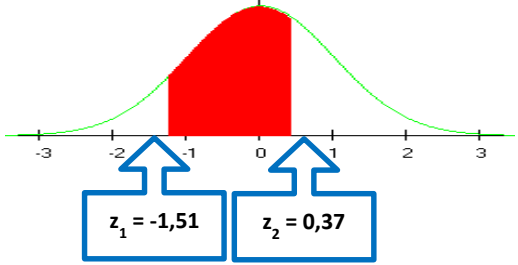
z	.03	.04	.05
0.0	.5120	.5160	.5199
0.1	.5517	.5557	.5596
0.2	.5910	.5948	.5987
0.3	.6293	.6331	.6368
0.4	.6664	.6700	.6736
0.5	.7019	.7054	.7088
0.6	.7357	.7389	.7422
0.7	.7673	.7704	.7734
0.8	.7967	.7995	.8023
0.9	.8238	.8264	.8289
1.0	.8485	.8508	.8531
1.1	.8708	.8729	.8749

RESPUESTA: $1 - 0,8238 = 0,1762$

Fuente: Baldeón Crisóstomo José

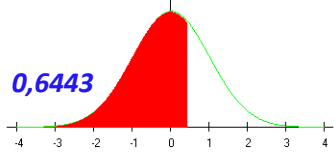
Tercer caso: Entre dos valores "z" determinados. Se calcula **por diferencia de áreas (Área pedida = Área mayor - Área menor)**.

Calcule el área bajo la curva normal comprendida entre $z_1 = -1,51$ y $z_2 = 0,37$

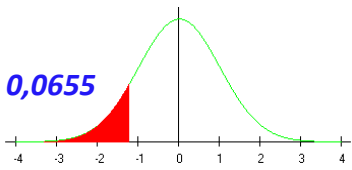


CALCULO DE AREAS

Área mayor: A la izquierda de 0,37



Área menor: A la izquierda de -1,51



RESPUESTA:
 $0,6443 - 0,0655 = 0,5788$

Fuente: Baldeón Crisóstomo José

EJERCICIOS RESUELTOS

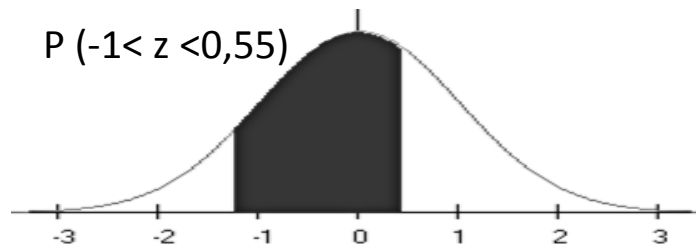
- Consideremos que la variable z sigue una distribución $N(0; 1)$. Halle las siguientes probabilidades:
 - $P(z < 2,78)$
 - $P(z > 0,50)$
 - $P(z < -1,5)$
 - $P(0,83 < z < 1,64)$
 - Valor V tal que $P(z < V) = 0,6480$
 - Valor V tal que $P(z < V) = 0,0028$



Solución:

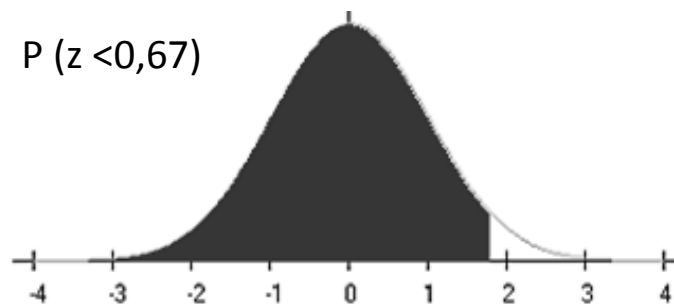
- a) Se busca en la tabla de las puntuaciones z positivas: $F(2,78)=0,9973$
Entonces se concluye que: $P(z < 2,78) = 0,9973$
- b) Buscamos en la tabla: $F(0,50)=0,6915$, entonces: $P(z > 0,50) = 1 - P(z < 0,50)$ ya que nos preguntan lo que corresponde al lado derecho:
 $P(z > 0,50) = 1 - 0,6915 = 0,3085$
- c) Se busca en la tabla de las puntuaciones z negativas: $F(-1,5)=0,0668$
Entonces se concluye que: $P(z < -1,50) = 0,0668$
- d) Al buscar en la tabla de las puntuaciones z tenemos:
 $F(0,83)=0,7967$ y $F(1,64)=0,9495$
 $P(0,83 < z < 1,64) = 0,9495 - 0,7967 = 0,1528$
- e) Valor V tal que $P(z < V) = 0,6480$; buscamos en la tabla de puntuaciones el valor "0,6480" y su correspondiente valor "z" y tenemos que para 0,6480 le corresponde $z = 0,38$. Entonces:
 $P(z < 0,38) = 0,6480$
- f) Valor V tal que $P(z < V) = 0,0028$; se busca en la tabla de puntuaciones z negativas el valor "0,0028", ya que la probabilidad es menor a "0,5000" y se observa que el valor que le corresponde a "z" es: -2,77; entonces: $P(z < -2,77) = 0,0028$

2. Encontrar el área sombreada:



Solución:

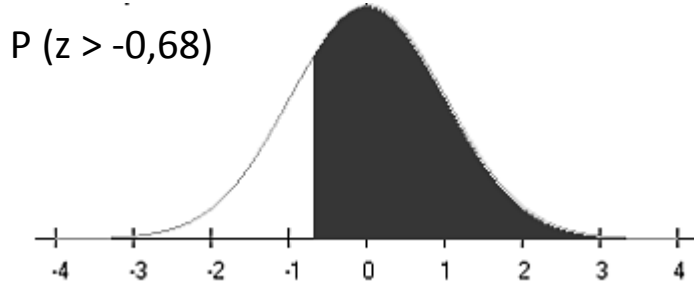
- Primero, para $z = 0,55$, buscamos en la tabla **Puntuaciones z POSITIVAS** y obtenemos en la intersección 0,7088
 - Luego, para $z = -1$, buscamos en la tabla **Puntuaciones z NEGATIVAS** y obtenemos en la intersección 0,1587
 - Finalmente, restamos ambas áreas, la diferencia es el área sombreada:
 $P(-1 < z < 0,55) = 0,7088 - 0,1587 = 0,5501 \sim 55,01\%$
3. Encontrar el área sombreada:





Solución:

- Primero, para $z=0,67$, buscamos en la tabla **Puntuaciones z POSITIVAS** y obtenemos en la intersección 0,7486
- Luego:
 $P(-1 < z < 0,55) = 0,7486 \sim 74,86\%$
- 4. Encontrar el área sombreada:



Solución:

- Primero, para $z=-0,68$, buscamos en la tabla **Puntuaciones z NEGATIVAS** y obtenemos en la intersección 0,2483
- Luego, buscamos su complementario pues la parte sombreada esta hacia la derecha y no hacia la izquierda como en casos anteriores. Así, su complementario es $1-0,2483=0,7517$
- Finalmente, tenemos:
 $P(z > -0,67) = 0,7517 \sim 75,17\%$

EJERCICIOS PROPUESTOS

1. Determine el valor Verdad (V) o Falsedad (F) de las siguientes proposiciones:
 - a) Existe una familia de distribuciones, diferenciadas por los valores de su media y su varianza. ()
 - b) La distribución de las estaturas de los estudiantes varones de la IE "Santa Isabel" de Huancayo se aproxima a una distribución normal ()
 - c) En la distribución normal $N(0; 1)$, la $P(x=0,40) = 0,000$ ()
 - d) El área bajo la curva comprendida entre los valores situados aproximadamente a dos desviaciones estándar de la media es igual a 0,90 ()
 - e) La media aritmética de las calificaciones z es positivo ()
 - f) Las notas obtenidas por los estudiantes de la Universidad Continental en el curso de Estadística I corresponden con una distribución normal. ()
2. Calcule el valor "z" que corresponda en cada caso:

$x = 36$ Kg. Media: 25 Kg. Desv. Est. : 1,4 z = _____ =	$x = 15$ años. Media: 18 años Desv. Est. : 0,83 años z = _____ =	$x = S/.$ 1200 Media: S/. 1200 Desv. Est. : S/. 645 z = _____ =
---	--	---



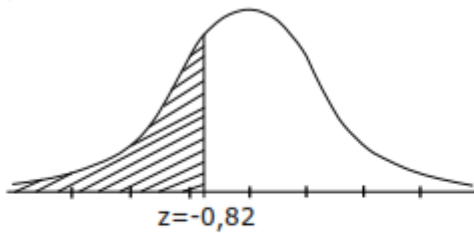
$x = 12$ ton. Media: 16,4 ton. Desv. Est. : 0,25 ton. $z = \underline{\hspace{2cm}} =$	$x = 356$ Km. Media: 385 Km. Desv. Est. : 18,34 Km. $z = \underline{\hspace{2cm}} =$	$x = 1,78$ m Media: S/. 1,68 m Desv. Est. : S/. 0,2 m $z = \underline{\hspace{2cm}} =$
---	---	---

3. Encuentre la probabilidad de que una porción de datos elegidos al azar de una población distribuida normalmente tenga un puntaje normal (z) que sea:

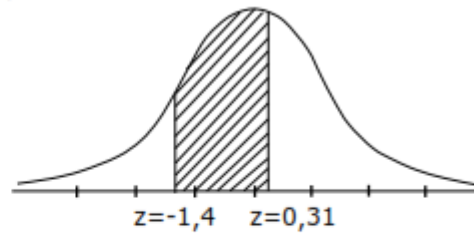
- a) Menor que 3,00
- b) Menor que -0,75
- c) Mayor que -1,25
- d) Entre -0,84 y 0,84
- e) Entre -3,5 y 3,5

4. En las siguientes curvas normales determine el porcentaje del área sombreada:

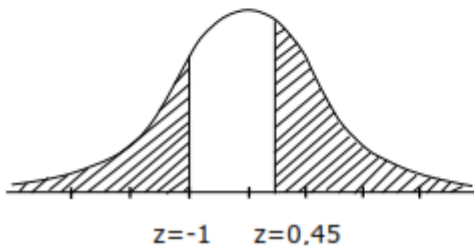
a)



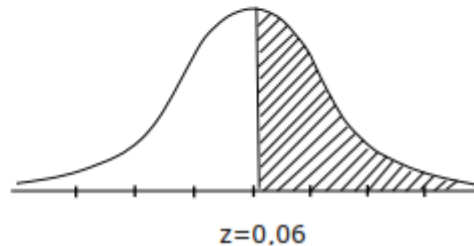
b)



c)



d)



5. Suponga que puntuaciones z se distribuyen normalmente con una media de 0 y una desviación estándar 1.

- a) Si $P(0 < z < a) = 0,3907$, halle el valor de a .
- b) Si $P(-b < z < b) = 0,8664$, determine el valor b .
- c) Si $P(z > c) = 0,0643$, halle el valor de c .
- d) Si $P(z < e) = 0,4500$ determine el valor de e

6. En la empresa DURACELL, hicieron un análisis del tiempo de duración de las pilas que desean vender. Una pila elegida al hacer dura en promedio 2.9 años, con una desviación estándar



de 0.5 años. Asumiendo que la vida de las pilas DURACELL está normalmente distribuida, encuentre la probabilidad de que una pila dada dure menos de 2.3 años.

REFERENCIAS BIBLIOGRÁFICAS Y ENLACES

- Triola Mario F. ESTADÍSTICA. Pearson Educación. México 2009
- R. Johnson, P.Kuby. ESTADISTICA ELEMENTAL. Thomson. 2da Edición. México 2002



Semana N°16: APLICACIONES DE LA DISTRIBUCIÓN NORMAL

1. TEMA: Aplicaciones de la distribución normal

2. PROPÓSITO:

- Reconoce la probabilidad en una variable aleatoria continua.
- Aplica la distribución normal estándar en la solución de problemas de su especialidad.

En la clase anterior se estudiaron las aplicaciones de la distribución normal, pero todos los ejemplos y ejercicios estuvieron basados en la distribución normal estándar con $\mu=0$ y $\sigma=1$. En esta semana, resolveremos ejercicios con interpolación y se presentarán métodos para trabajar con distribuciones normales que no son estándar, es decir, cuya media no es 0 o cuya desviación no es 1. Por lo tanto, este proceso de estandarización nos permite hacer aplicaciones mucho más realistas y significativas.

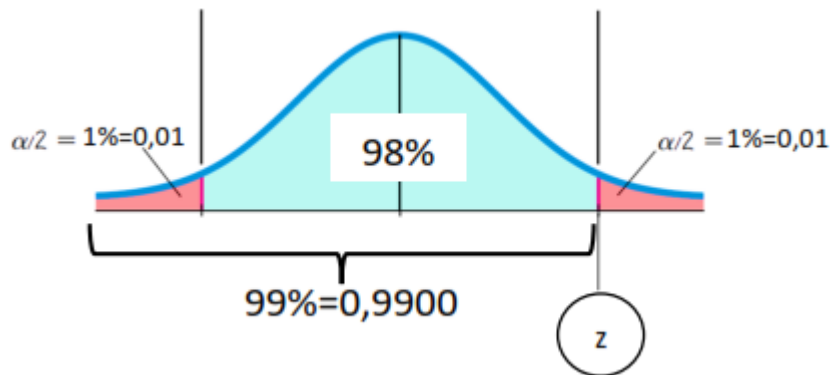
1. INTERPOLACIÓN EN LA DISTRIBUCION NORMAL

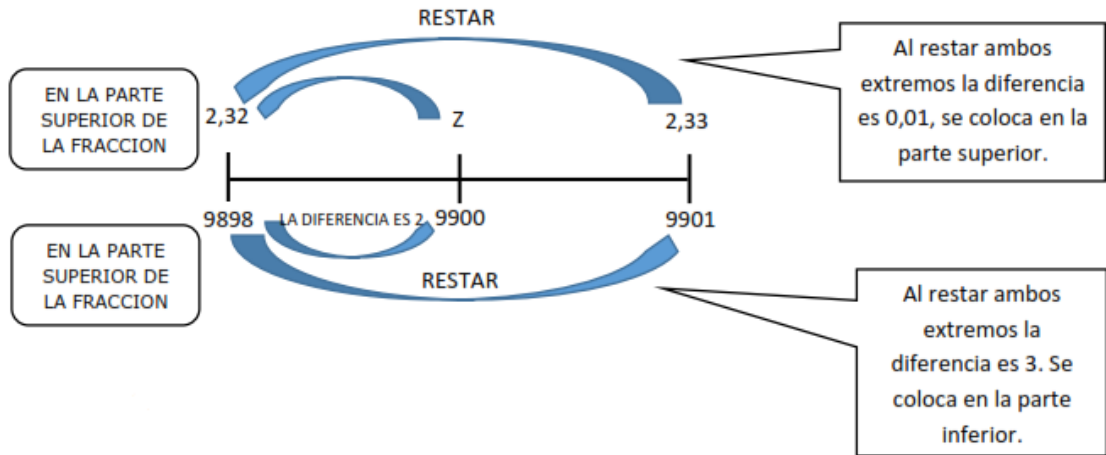
Se utiliza la interpolación cuando no encontramos algunos valores z para ciertas áreas en la tabla de PUNTUACIONES Z POSITIVAS/NEGATIVAS:

1.1. EN UNA DISTRIBUCION NORMAL CON DOS COLAS:

Por ejemplo, con un nivel de confianza del 98%, halle el valor z :

$$1 - \alpha = 98\%; \alpha = 2\%$$





i) $\frac{0,01}{3} = \frac{x}{2}$
X=0,0067

ii) Entonces, hallando z:

$$Z=2,32+x$$

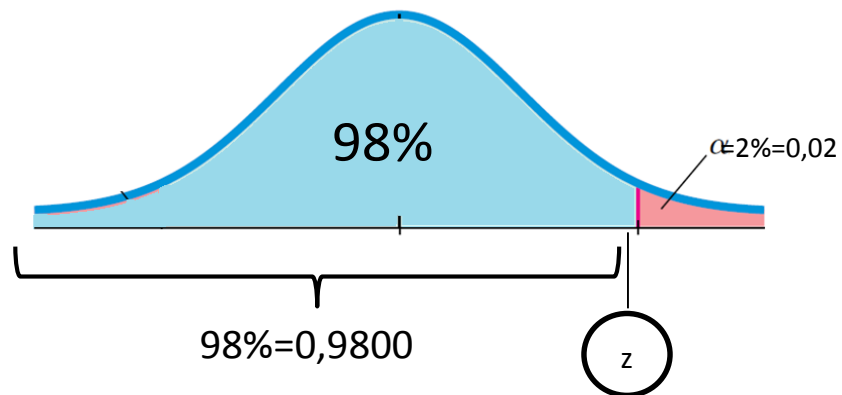
$$Z=2,32+0,0067$$

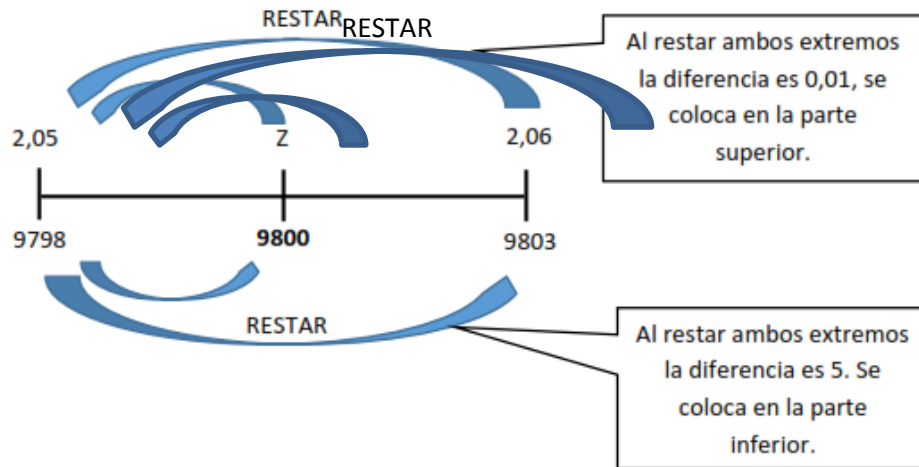
$$\mathbf{Z=2,3267}$$

1.2. EN UNA DISTRIBUCION NORMAL CON UNA COLA:

Por ejemplo, con un nivel de confianza del 98%, halle el valor z: α

$$1 - \alpha = 98\% ; \alpha = 2\%$$





a) $\frac{0,01}{5} = \frac{x}{2}$
 $x=0,004$

b) Entonces, hallando z:

$$Z=2,05+x$$

$$Z=2,05+0,004$$

$$\mathbf{Z=2,054}$$

2. APLICACIÓN DE LA DISTRIBUCION NORMAL

2.1. Estandarización "z" o calificaciones estándar

Recuerde que el puntaje normal "z", se definió la anterior semana:

$$z = \frac{x - \mu}{\sigma}$$

La estandarización de la variable x permite la transformación hacia las calificaciones "z". La estandarización transforma cualquier variable x en una nueva variable "z", que tiene media cero y desviación uno.

2.2. Distribución normal no estándar

Cuando se calcula áreas en una distribución normal no estándar debe utilizar el siguiente procedimiento:

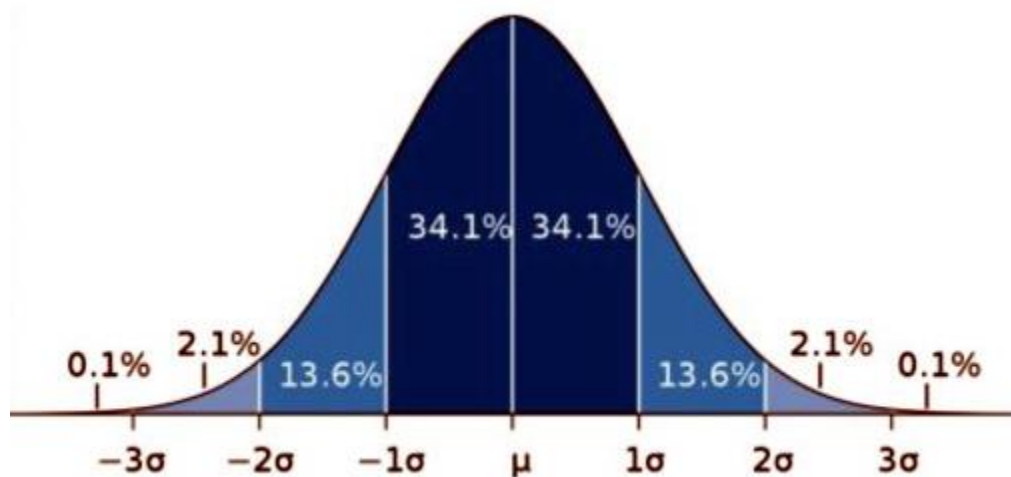
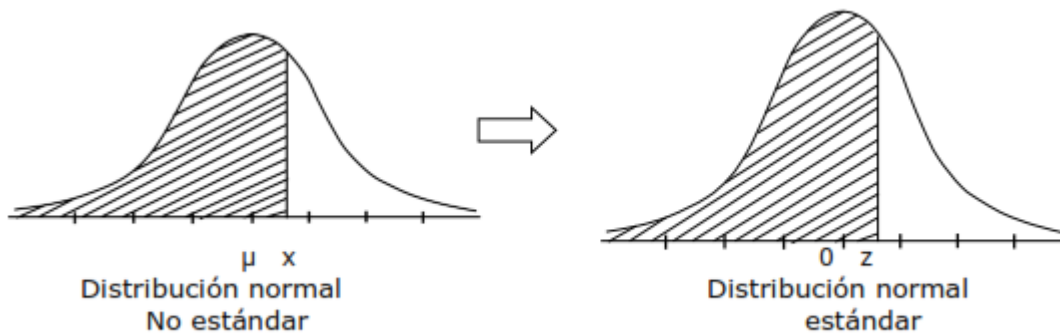
- Dibuje una curva normal, indique los valores específicos, después sombree la región que representa la probabilidad deseada.
- Para cada valor relevante de x que sea un límite de la región sombreada, utilice la formula anterior para convertir cada valor a la puntuación z equivalente.



- c) Utilice la tabla de PUNTUACIONES Z para encontrar el área de la región sombreada, que constituye la probabilidad deseada.

Cuando se calcula áreas en una distribución normal no estándar debe utilizar el siguiente procedimiento:

- Dibuje una curva normal, indique la media y los valores específicos de x , después sombree la región que representa la probabilidad deseada.
- Para cada valor relevante de x que sea un límite de la región sombreada, utilice la fórmula anterior para convertir cada el valor a la puntuación z equivalente.
- Utilice la tabla de la distribución normal para encontrar el área de la región sombreada, que constituye la probabilidad deseada.



Interpretación:

- El 68,2% de las observaciones están comprendidas en el intervalo $\langle \mu - \sigma; \mu + \sigma \rangle$
- El 95,4% de las observaciones están comprendidas en el intervalo $\langle \mu - 2\sigma; \mu + 2\sigma \rangle$
- El 99,6% de las observaciones están comprendidas en el intervalo $\langle \mu - 3\sigma; \mu + 3\sigma \rangle$

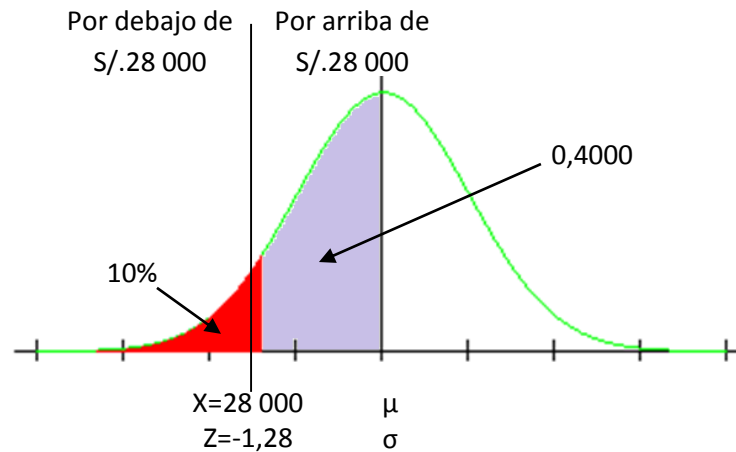
EJERCICIOS RESUELTOS



1. Los ingresos de los ingenieros en una empresa están distribuidos normalmente con una desviación estándar de S/. 1 200. Se piensa hacer un recorte de personal, por lo que los empleados que ganan menos de S/. 28 000 serán despedidos. Si el despido representa al 10% de tales ingenieros. ¿Cuál es el salario medio actual del grupo de ingenieros?

Solución:

Si el 10% de los salarios están por debajo de S/. 28 000, entonces 40% (o 0,4000) están entre S/.28 000 y la media μ . La tabla de la distribución normal indica que $z=-1,28$ que es el puntaje normal que ocurre cuando $x=28000$. Al aplicar la formula puede encontrarse el valor de μ :



$$-1,28 = \frac{28000 - \mu}{1,200}$$

$$-1,536 = 28000 - \mu$$

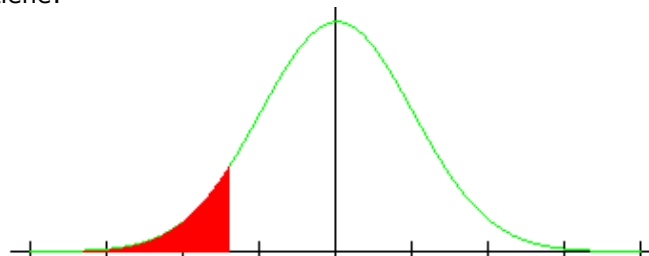
$$\mu = S/. 29536$$

Es decir, el salario actual de los ingenieros es de S/. 29 536

2. ¿Cuál es la probabilidad de elegir aleatoriamente a un trabajador de "PLAZA VEA" que tenga menos de 10% del total de tardanzas?, si de esta población se sabe que esta normalmente distribuida y su media es de 12,40% y la desviación estándar es de 0,60%.

Solución:

Gráficamente se tiene:





3. Tenemos los datos:
 $\mu=12,40\%$; $\sigma=1,40\%$ y $x=10\%$

Utilizamos la fórmula para hallar la calificación estándar:

$$z = \frac{x - \mu}{\sigma}$$

$$z = \frac{10 - 12,40}{1,40} = -1,7142 \sim -1,71 \text{ luego } z = -1,71$$

4. Con el valor que le corresponde en la tabla podemos afirmar que:
 $P(z < -1,71) = 0,0436$
5. Finalmente, se concluye que la probabilidad de elegir aleatoriamente a un trabajador que tenga menos de 10% del total de tardanzas es de 0,0436

EJERCICIOS PROPUESTOS

1. Un psicólogo está diseñando un experimento para probar la eficacia de un nuevo programa de capacitación para vigilantes de seguridad de "REAL PLAZA". Desea comenzar con un grupo homogéneo de sujetos con puntuaciones de CI comprendidas entre 85 y 125. Dado que las puntuaciones de CI se distribuyen normalmente, con una media de 100 y una desviación estándar de 15, ¿qué porcentaje de la gente tiene una puntuación de CI entre 85 y 125?
2. El tiempo de espera x del banco "INTERBANK" tiene una distribución normal con una media de 3,7 minutos y una desviación estándar de 1,4. Encuentre la probabilidad de que un cliente elegido de forma aleatoria haya tenido que esperar menos de 2,0 minutos.
3. Dada una distribución normal basada en 2100 casos sobre el peso de un electrodoméstico en TIENDAS EFE, con una media de 51 kilogramos y una varianza de 60,84 kg^2 . Determinar la proporción del área y el número de casos entre la media y la calificación 69Kg.
4. Se ha determinado que la vida útil de la marca de llantas GOODYEAR PERU tiene una distribución normal con media de 35 000 Km y una desviación estándar de 2 500 Km. Si un distribuidor hace un pedido de 400 llantas, ¿aproximadamente cuántas llantas durarán más de 41 000 Km?
5. Las calificaciones de 10 aspirantes presentados a un examen para contratación laboral, se distribuye normalmente con media 6,55 y varianza 4.
 - a) Calcule y grafique la probabilidad de que un aspirante obtenga más de 8 puntos.
 - b) Determine la proporción de aspirantes con calificaciones inferiores a 5 puntos. (Graficar)
 - c) ¿Cuál es la probabilidad de los aspirantes que obtuvieron calificaciones comprendidas entre 5 y 7,5 puntos? (Graficar)
6. La empresa PHILIPS instala en una ciudad 20 000 focos para su iluminación. La duración de una distribución normal con media de 300 días y desviación típica de 35 días. Calcular: ¿Cuántas duraran más de 350 días?
7. Las precipitaciones anuales en una región alcanzan, de media, los 1500 mm, con una desviación típica de 200mm. Calcula, suponiendo que siguen una distribución normal, la probabilidad de que en un año determinado la lluvia:
 - a. No supere los 1200 mm
 - b. Supere los 1500 mm.



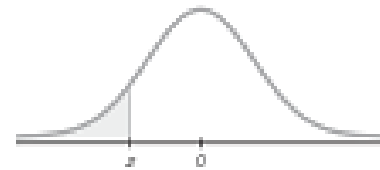
- c. Esté entre 1700 y 2300 mm.
- d. Deseamos seleccionar el 25% de los años más lluviosos, ¿a partir de qué cantidad de agua hemos de escogerlos? ¿Y si deseáramos seleccionar los menos lluviosos?

REFERENCIAS BIBLIOGRÁFICAS Y ENLACES

- Triola Mario F. ESTADÍSTICA. Pearson Educación. México 2009
- R. Johnson, P.Kuby. ESTADISTICA ELEMENTAL. Thomson. 2da Edición. México 2002



ANEXOS



Puntuaciones z NEGATIVAS

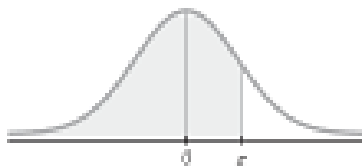
TABLA A-2 Distribución normal estándar (z): Área acumulativa desde la IZQUIERDA

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.50 y menores	.0001									
-3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
-3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
-3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
-3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
-3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
-2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
-2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
-2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
-2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
-2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
-2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
-2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
-2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
-2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
-2.0	.0238	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
-1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
-1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
-1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
-1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
-1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
-1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
-1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
-1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
-1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
-1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
-0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
-0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
-0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
-0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
-0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
-0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
-0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
-0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
-0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
-0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

NOTA: Para valores de z por debajo de -3.49, utilice 0.0001 para el área.

*Utilice estos valores comunes que resultan por interpolación:

Puntuación	Área
-1.645	0.0500
-2.375	0.0090



Puntuaciones z POSITIVAS

TABLA A-2 (continuación) Área acumulativa desde la IZQUIERDA

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.50	.9999									
v mayores										

NOTA: Para valores de z por encima de 3.49, utilice 0.9999 para el área.

*Utilice estos valores comunes que resultan por interpolación:

Puntuación	Área
z	Área
1.645	0.9500
2.575	0.9950

Valores críticos comunes

Nivel de confianza	Valor crítico
0.90	1.645
0.95	1.96
0.99	2.575



MODELOS DE ENCUESTAS

ENCUESTA SOBRE LA DEMANDA DE LOS ESTUDIANTES EN LAS DISCOTECAS

Instrucciones: Te agradecemos anticipadamente tu gentil colaboración y te solicitamos que a continuación contestes cada uno de los ítems de esta encuesta de manera sincera y anónima, ya que la veracidad de tus respuestas contribuirá significativamente en los resultados de nuestra investigación.

I) Datos generales:

- 1.1. Género: Masculino () Femenino () 1.2 Edad: años.
- 1.3. Distrito de residencia: Huancayo () El Tambo () Chilca ()
Chupaca () Otros () Especifique:
- 1.4. Tienes enamorada: Si () No ()
- 1.5. ¿Cuál es tu pasatiempo favorito?: Hacer deporte() Ir al cine o ver TV()
Leer() Pasear() Redes sociales() Otro()
Especifique:
- 1.6. ¿Qué deporte practicas? Fútbol() Vóley() Basquetbol()
Natación() Otros() Especifique:
- 1.7. ¿Con quien o quienes vives? Solo(a)() Con mis padres()
Con familiares() Otros() Especifique:
- 1.8. ¿Tus padres te apoyan económicamente en tus estudios? Si() No()
- 1.9. ¿Tienes algún tipo de problema? Si() No()
Si tu respuesta es afirmativa indica: Familiar() Sentimental() Salud() Económico() Laboral()
Otros() Especifica:

II) Datos académicos:

- 2.1. ¿Qué carrera profesional estudias?
- 2.2. ¿Alguna vez desaprobase algún curso en la Universidad? Si () No ()
- 2.3. ¿Tienes hábitos de estudio? Si() No()
- 2.4. ¿Cómo caracterizas tus hábitos de estudio? Excelente() Bueno()
Regular() Malo()
- 2.5. ¿Cuántas horas diarias accedes a internet con fines de estudio? 0 horas()
Media hora() 1 hora() 2 horas() Más de 2 horas()

III. Datos laborales

- 3.1. ¿Trabajas? Si() No()
- 3.2. ¿Cuál es tu ingreso económico mensual: S/.
- 3.3. ¿Te agrada el tipo de trabajo que haces?: Si () No()
- 3.4. ¿Cómo calificas tu trabajo actual? Bueno() Regular() Malo()
- 3.5. Turno de trabajo: Mañanas() Tardes() Noches()
Otros() Especifica:

IV. Datos sobre la asistencia a las discotecas

- 4.1. ¿Vas a las discotecas que están alrededor de la Universidad? Si() No()
- 4.2. ¿Con qué frecuencia vas a las discotecas? Nunca() 1 vez a la semana()
2 veces a la semana() 3 veces a la semana() Diario()
- 4.3. ¿Cuál es el motivo de tu asistencia a las discotecas? Por diversión()
Para olvidar mis problemas() Para conocer chicas/chicos()
Otros() Especifique:
- 4.4. ¿Cuánto gastas aproximadamente cuando vas a las discotecas? S/.
- 4.5. ¿Qué es lo que más te agrada cuando vas a las discotecas?
- 4.6. ¿Con quienes vas a las discotecas? Con amigos() Con mi enamorada/o()
Otros() Especifique:

¡Muchas gracias por tu colaboración!



ENCUESTA SOBRE LA ACEPTACIÓN DE UN RESTAURANT DE COMIDA SALUDABLE

Estimado joven/señorita, la presente encuesta tiene por finalidad conocer su opinión **sobre la aceptación de un restaurant de comida saludable**. Marque con una "X" en la viñeta que represente mejor su parecer o rellene en los espacios correspondientes de ser el caso. No está permitido marcar más de una respuesta, salvo que se indique lo contrario.

DATOS PERSONALES:

1. Edad: _____
2. Género: ① Masculino
 ② Femenino
3. Distrito de: ① Huancayo ② Tambo ③ Chilca ④ Otros

DATOS SOBRE LA ACEPTACIÓN DE UN RESTAURANT DE COMIDA SALUDABLE

4. ¿Considera que usted tiene una alimentación saludable?
① Si ② No ③ Algunas veces
5. ¿Cuántas veces usted come al día en la universidad?
① 1 ② 2 ③ 3 ④ 4
6. ¿Con que regularidad come fuera de la universidad?
① La mayoría del tiempo
② Algunas veces ③ Nunca
7. ¿Qué elementos influyen en la elección de un restaurant?
① Precio ② Buena atención
③ Buena sazón ④ Ambiente agradable ⑤ Ubicación
8. ¿Cuánto gastas a la semana en comida fuera de la universidad?
① 5 a 10 ② 10 a 15 ③ 15 a 20
9. ¿Qué comida más prefiere?
① Frituras ② Lácteos ③ Carnes ④ Harinas ⑤ Verduras
10. ¿A qué hora asiste preferentemente?
① Desayuno ② Almuerzo ③ Cena
11. ¿Cuántos productos lácteos tomas al día?
① Ninguno ② 1 ③ Entre 2 y 3 tomas
12. ¿Cuál de estos grupos de alimentos consideras más rico en fibra?
① Lentejas, cereales para el desayuno, guisantes
② Naranja, lechuga, huevos
③ Carne, zumo de naranja, pescado
13. ¿Cuál de las tres opciones tiene menos calorías?
① La mermelada ② El aceite de oliva ③ Los frutos
14. ¿Cada cuánto comes frutas a la semana los restaurants que se encuentran alrededor de la universidad?
① Siempre ② Nunca ③ Algunas veces
15. ¿Cada cuánto comes verduras a la semana en los restaurants que se encuentran alrededor de la universidad?
① 1 o 2 veces ② 3 o 4 veces ③ 5 o 6 veces ④ Todos los días ⑤ Nunca

MUCHAS GRACIAS POR SU COLABORACIÓN



ENCUESTA DE LA PERCEPCIÓN DEL PESO CORPORAL

Buen día, la presente encuesta tiene por finalidad conocer su opinión **sobre la percepción del peso corporal**. Subraye la pregunta que represente mejor su parecer o rellene en los espacios correspondientes de ser el caso. No está permitido marcar más de una respuesta, recuerde que esta encuesta es anónima y los datos obtenidos nos servirán para un trabajo de investigación

1

- DATOS PERSONALES:

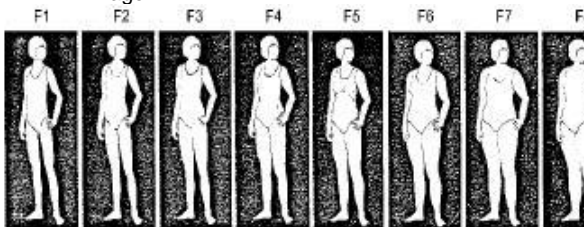
16. Edad: _____
17. Género: ① Masculino ② Femenino
18. Distrito de residencia:
① Huancayo ② Tambo ③ Chilca
④ Otros
19. Grado de instrucción:
① Primarias ② Secundaria
③ Superior Técnico ④ Universitario
20. Condición laboral ¿Trabajas?
① Sí ② No.....
21. Ingreso promedio aproximado: _____

2.- DATOS SOBRE LA PERCEPCION DEL PESO CORPOAL

1. ¿Cuál es su peso en kilogramos sin zapatos?

2. ¿Cuál es su estatura estatura sin zapatos?

3. Usted se considera:
① De peso bajo ② Normal
③ Con sobrepeso ④ obeso (a)
4. Según el siguiente gráfico, ¿Cuál de los siguientes gráficos cree usted representa su imagen?



5. ¿Considera que su peso es adecuado en relación a talla?
① Sí ② No
6. Usted desearía:
① Bajar de peso ② Mantener su peso
③ Subir de peso

7. ¿Para usted es importante la percepción que las demás personas tienen acerca de su imagen?
① Bastante ② Más o menos
③ No
8. ¿Ha realizado algún tipo de dieta?
① Sí ② No
9. Si su respuesta es sí, responda lo siguiente
① Para modificar su peso
② Por motivo de salud
③ Debido algún régimen deportivo
10. ¿Cuántas veces al día come?
① 1 vez
② 2 veces
③ 3 veces
④ 4 veces
⑤ 5 veces a más
11. ¿Cuántas veces a la semana ingiere comida rápida?
① Nunca
② 1 a 2 veces
③ 3 a 4 veces
④ 5 a 6 veces
⑤ Todos los días.
12. ¿Dentro de su casa, que comidas la familia se sienta en la mesa reunida?
① Ninguna
② Solo uno de ellos
 • Desayunos
 • Almuerzo
 • Cena
③ Solo dos de ellos:
 • Desayunos y almuerzo
 • Almuerzo y cena
 • Desayuno y Cena
④ Todos



13. Al elegir su comida, usted prefiere:
① Alimentos bajos en grasa y bajos en calorías

- ② Usted no se fija en el aporte calórico de los alimentos
③ Evita las comidas y reuniones que involucren comer

ENCUESTA DE PREFERENCIAS DE JEAN

Buenos días/tardes, estamos realizando una encuesta para evaluar el lanzamiento de una nueva marca de jeans. Le agradeceremos brindarnos un minuto de su tiempo y responder las siguientes preguntas:

I. DATOS DE IDENTIFICACIÓN

- 1.1 Género: 1. Mujer 2. Hombre
1.2 Edad _____ años cumplidos
1.3. Ingreso promedio: _____
1.4. Grado de Instrucción: _____

2.6. ¿Cuáles de estos tipos de aplicaciones prefiere en un jean?

1. Bordados en bolsillos delanteros
2. Bordados en bolsillos traseros
3. Greviches (piedras de colores) en bolsillos delanteros
4. Greviches en bolsillos traseros
5. Tachas en bolsillos traseros
6. Tachas en bolsillos delanteros
7. Sin muchas aplicaciones
8. Otros (mencione) _____

II. SOBRE USO DE JEANS

2.1. ¿Usa usted la prenda del jean?

1. Si 2. No (fin de la encuesta)

2.2. ¿Qué marcas de jeans usa o prefiere?

1. Fiorucci
2. Sibilla
3. Pioner
4. Milk
5. Tayssir
6. Otras (mencione) _____
7. No tengo marca preferida

2.7. ¿A cuál de estos lugares suele acudir para comprar sus jeans?

1. Centros comerciales
2. Tiendas por departamento
3. Bazares o galerías
4. Supermercados
5. Mercados mayoristas
6. Otros

2.3. ¿Al momento de comprar una prenda de jean, qué es lo primero que toma en cuenta?

1. La marca
2. La calidad de la tela
3. El modelo
4. Los acabados
5. El precio

2.8. ¿Con qué frecuencia compra usted la prenda del jean?

1. Semanal
2. Quincenal
3. Mensual
4. Cada 3 meses
5. Cada 6 meses



6. Lo bien que le pueda quedar

6. Mayor a un año

7. Otros

7. Otros (mencione) _____

2.4. ¿Qué modelos de jeans prefiere?

2.9. ¿Cuánto suele gastar en promedio al comprar un jean?

1. Clásicos

_____ (nuevos soles)

2. Pitillos

3. Acampanados

2.10. ¿Estaría dispuesta a probar una nueva marca de jeans?

4. Capri

1. Sí

5. De pretina ancha

2. No

6. El que esté de moda

3. No sabe / no opina

7. Otros

2.5. ¿Cuáles de estos adornos prefiere en la confección de un jean?

Muchas gracias

1. Doble bolsillo

2. Dobleces en piernas

3. Bolsillos en piernas

4. Bolsillos delanteros con cierre

5. Bolsillos traseros con cierre

6. Sin muchos adornos

7. Otros (mencione)



ENCUESTA PARA EVALUACION EDUCATIVA Y ESTADISTICA

La presente encuesta tiene como objetivo medir el grado de conocimiento que tienen los encuestados, acerca de Educación Vial.

Edad: _____ Sexo: ___M___F

Fecha:

1. ¿Cree usted que todo vehículo que circule sin prestar las seguridades necesarias será...?
 - a) Multado y detenido
 - b) Retirado de la circulación
 - c) Llamado la atención
2. ¿Cree usted que el límite máximo de velocidad establecido para los vehículos livianos en el sector urbano es de...?
 - a) 50 km/h
 - b) 90 km/h
 - c) 100 km/h
3. ¿Según usted cual es el límite máximo de velocidad establecido para los vehículos de pasajeros en el sector urbano es de...?
 - a) 40 km/h
 - b) 70 km/h
 - c) 90 km/h
4. ¿Piensa usted que el límite máximo de velocidad establecido para los vehículos de transporte de carga en carrera es de...?
 - a) 50 km/h
 - b) 70 km/h
 - c) 90 km/h
5. ¿Cree usted Cuando dos vehículos lleguen simultáneamente a una intersección, tendrá derecho de vía el que se aproxime por su...?
 - a) Derecha
 - b) Izquierda
 - c) Ninguno
6. ¿Según su criterio la pérdida de adherencia de los neumáticos del vehículo sobre la calzada se conoce como...?
 - a) Derrape
 - b) Estrellamiento
 - c) Volcamiento
7. ¿Mientras conduce, usted debe mantener una distancia de seguridad lateral mínima de...? a) 1.5 metros
 - b) 2.5 metros
 - c) 3.5 metros
8. ¿Qué sanción le impondría a usted el agente de tránsito si estaciona su vehículo en una parada de bus...?
 - a) Retiro del vehículo y multa
 - b) Llamado de atención
 - c) Ninguna
9. ¿Si su vehículo sufre un desperfecto en la carretera, a qué distancia colocaría los triángulos de seguridad...?
 - a) Entre 200 m. adelante y 40 m. en la parte posterior
 - b) Entre 50 y 150 metros, uno adelante y otro en la parte posterior.
 - c) Entre 100 m. adelante y 100 m. en la parte posterior
10. ¿Considera usted abandonado un vehículo, por el hecho de dejarlo en la vía pública sin conductor después de...?
 - a) 5 horas
 - b) 24 horas
 - c) 48 horas

Muchas gracias



Referencias bibliográficas

BÁSICA

Triola Mario F. ESTADÍSTICA. Pearson Educación. México 2010

COMPLEMENTARIA

- Bejarano, M. (1995) Estadística descriptiva, probabilidades y lineamientos para la elaboración del protocolo de investigación en ciencias de la Salud. Universidad Peruana Cayetano Heredia.
- Berenson, M. y Levine, D. (2010). *Estadística básica en administración, Conceptos y aplicaciones*. México: Prentice Hall.
- Delgado, R. (2008). *Probabilidad y estadística para ciencias e ingeniería* (1ª ed.). México publicaciones delta.
- Jorge Chue, Emma Barreno, Carlos Castillo, Rosa Millones, Félix Vásquez (2008) Estadística Descriptiva y Probabilidades. Universidad de Lima. Fondo Editorial
- Martínez, B. (2012). *Estadística y muestreo*. ECOE Ediciones.
- Pérez, C. (2012) Estadística Aplicada –IBM SPSS. España: Garceta grupo editorial.
- Robert Johnson, Patricia Kuby. Estadística Elemental. Thomson Editorial. México 2002
- Ross, S. (2011). *Probabilidad y Estadística para Ingenieros*. 3a. ed. México: Mc Graw Hill, 2001.
- Toma, J. y Rubio, J. Estadística Aplicada primera parte. Universidad del Pacífico. Centro de Investigación, 2011. ISBN:978-9972-57-109-1

ENLACES Y DIRECCIONES ELECTRÓNICAS

http://dme.ufro.cl/clinicamatematica/images/Libros/Estadistica_y_Probabilidad/Estadistica%20y%20Probabilidad.pdf

http://www.vitutor.com/estadistica/descriptiva/a_10.html

http://www.virtual.unal.edu.co/cursos/ciencias/2001065/html/un1/cont_130_30.html

http://pendientedemigracion.ucm.es/info/Astrof/users/jaz/ESTADISTICA/libro_GCZ2009.pdf