

EDUCACIÓN a DISTANCIA

MATEMÁTICA 4



Buenos Aires Provincia

Dirección General de Cultura y Educación
Dirección de Educación de Adultos

MANUAL DE EDUCACIÓN A DISTANCIA

GOBERNADORA DE LA PROVINCIA DE BUENOS AIRES

Lic. María Eugenia Vidal

DIRECTOR GENERAL DE CULTURA Y EDUCACIÓN

Lic. Gabriel Sánchez Zinny

SUBSECRETARIO DE EDUCACIÓN

Lic. Sergio Siciliano

DIRECTOR DE EDUCACIÓN DE ADULTOS

Prof. Ing. Pedro Schiuma

SUBDIRECTOR DE EDUCACIÓN DE ADULTOS

Prof. Juan Carlos Latini

RESOLUCIÓN DE CREACIÓN 106/18

Adecuación de la estructura curricular modular del Programa Educación a Distancia

Año de impresión
2019 - 2^{da} Edición

PRESENTACIÓN

Este material que hoy llega a sus manos forma parte de una serie de módulos del Programa de Educación a Distancia (Res. 106/18) de la Dirección de Educación de Adultos de la Provincia de Buenos Aires. El mismo busca ampliar el acceso a la educación secundaria de aquellos jóvenes y adultos mayores de 18 años que se encuentren imposibilitados de concurrir a nuestras escuelas.

La evolución de las tecnologías de la información y de la comunicación nos permite repensar el modelo educativo de enseñanza-aprendizaje. El objetivo de la modalidad a distancia es superar las limitaciones de tiempo y espacio de todos aquellos bonaerenses que quieran terminar sus estudios secundarios. Este Programa tiene como propósito que los estudiantes puedan ingresar y egresar en cualquier momento del año, avanzando según su propio ritmo y con la posibilidad de organizar su trayecto formativo.

La Educación a Distancia es una herramienta que se suma a las ofertas de terminalidad secundaria que ofrece la provincia de Buenos Aires en pos de alcanzar a aquellos que el sistema educativo no les proponía una alternativa de estudio que no requiera concurrir a los servicios educativos presenciales de tiempo completo y con desplazamiento diario.

Esta modalidad se caracteriza por la mediatización de la relación entre el docente y el estudiante, a través de recursos de aprendizaje específicos que permiten la actividad autónoma de éstos.

Los estudiantes contarán así con el acompañamiento permanente de un profesor tutor a través de los distintos recursos que ofrece el Campus Virtual (campusvirtualadultos.com.ar), y también en instancias presenciales de encuentros individuales e intercambios abiertos grupales para compartir intereses, preocupaciones, dudas, opiniones, explicaciones, materiales, etc.

Este material estará disponible tanto en formato digital como impreso, para que sin importar sus posibilidades, los estudiantes tengan acceso al mismo. Completar sus estudios secundarios es, fundamentalmente, dar un paso más en la construcción de su ciudadanía.

Director de Educación de Adultos
Prof. Ing. Pedro Schiuma

■ Introducción

■ Unidad 1: Estadística

Apunte de clase: Estadística

- 1. Estadística
- 2. Algunos conceptos importantes
 - 2.1. Población
 - 2.2. Muestra
 - 2.3. Variable
 - 2.4. Clasificación de Variables
- 3. Frecuencias
- 4. Intervalos de frecuencia
- 5. Gráficos
 - 5.1. Diagrama de rectángulos
 - 5.2. Diagrama circular o Pictograma
 - 5.3. Histograma
 - 5.4. Polígono de Frecuencias
 - 5.5. Perfil ortogonal

■ Unidad 2: Estadísticos muestrales

Apuntes de clase: Estadísticos muestrales

- 1. Estadísticos muestrales
- 2. Medidas de Posición y de Tendencia Central
 - 2.1. Medidas de posición: Deciles, cuartiles y percentiles
 - 2.2. Medidas de centralización: Media, mediana y moda
- 3. Medidas de dispersión
 - 3.1. Desviación estándar o desviación típica

■ Unidad 3: Probabilidad

Apuntes de clase: Probabilidad

- 1. Probabilidad
 - 1.1. Los primeros pasos de la probabilidad
 - 1.2. Fenómenos y experimentos
 - 1.3. Espacio muestral y sucesos o eventos
 - 1.4. Probabilidad de un evento
 - 1.5. Algunos conceptos importantes
 - 1.5.1. Conjunto vacío
 - 1.5.2. Intersección de dos eventos
 - 1.5.3. Unión de dos eventos
 - 1.5.4. Sucesos o eventos mutuamente excluyentes
 - 1.5.5. Complemento
 - 1.5.6. Sucesos o eventos independientes





MATEMÁTICA 4

Introducción



¡Bienvenido a este nuevo módulo!

Aquí nos encontraremos con las herramientas necesarias de la estadística para analizar e interpretar el mundo que nos rodea. Podrán ver como se trabaja con los datos estadísticos, como organizarlos por tablas. A su vez aprenderán a analizar diferentes tipos de gráficos estadísticos y su importancia. También podrán comprender cuáles son las medidas más utilizadas de tendencia central y dispersión, verán que nos describe cada una.

El módulo cuenta con muchos ejemplos de estadísticas reales, con el objetivo de contextualizar de la mejor manera nuestra interpretación.

Por otro lado estudiarán la probabilidad, ese conocimiento que nos ayuda a razonar sobre sucesos atravesados por la incertidumbre, analizarán sus principales propiedades y lograrán resolver problemas.

¡Éxitos en su recorrido!



¡ Empezamos a estudiar!

Conozcamos un poco más sobre este tema leyendo el apunte de clase que se encuentra a continuación.

Apunte de clase: Estadística



1. Estadística

Muchas veces escuchamos frases tales como:

“Según las estadísticas, los niños que recibieron todas las dosis de las vacunas del calendario oficial tienen menos tendencia a enfermarse que quienes no las recibieron”

“El índice de precios aumentó 1.5% en el mes de enero, según un análisis estadístico de la empresa XX”.

“38,2% de la población no tiene gas natural, según los últimos datos recolectados en la encuesta del mes pasado”.

O vemos ciertos gráficos como los siguientes:

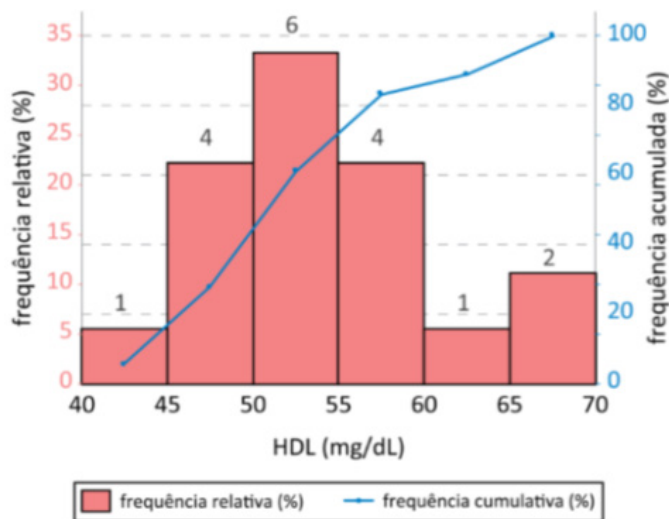


Figura 1. Histograma con polígono de frecuencia

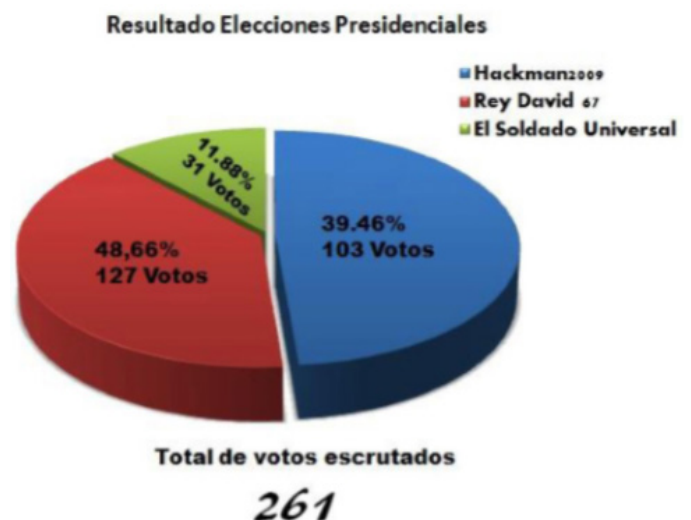


Figura 2. Gráfico circular 3d

¿Por qué aprender estadística?

Entre todas las causas que puedan existir la principal es que la estadística **nos sirve para comprender el mundo de nuestro entorno de una forma crítica.**

Actualmente los medios de comunicación nos abruman con informaciones de todo tipo con diferentes finalidades, cabe en este momento preguntarse, ¿siempre comprendemos claramente la información que nos brindan los medios?, veamos algunos ejemplos:

1. Con las cubiertas N, el auto logra frenar un 60% más rápido. (¿más rápido que qué?).

2. Siete de cada diez odontólogos prefieren la marca N de dentífricos. (¿Cuántos odontólogos fueron interrogados? ¿Cómo los escogieron? ¿Con qué otras marcas se realizaron las comparaciones?).

3. El complemento vitamínico X es recomendado por personas que más conocen del tema. (¿Quiénes son los que más conocen del tema? ¿Que tema?).

4. El 90 % de las motos de marca X vendidas en los últimos 15 años están todavía en circulación. (al presentar la información de esta forma llevaría a pensar que las motos tienen 15 años, sin embargo pudieron ser adquiridas hace 3 años).

5. Si hay 2 empanadas y dos personas, Juan y Ana. Juan se comió las dos empanadas, por lo tanto el consumo per cápita es de dos empanadas por persona. (Pero!!!, si Ana no comió).

Como vemos, en muchas ocasiones la información puede ser presentada en forma poco precisa o con intención de persuadir al consumidor para que este tome una decisión que en muchos casos es equivocada. En una nota a infobae, Sandra González, la presidenta de la Asociación de Defensa de los Consumidores y Usuarios de Argentina (ADECUA) ,explica lo siguiente:

“En los últimos años se incrementó de manera notable la cantidad de denuncias de los consumidores. Lo más raro de todo es que esa realidad nos deja dos sensaciones opuestas: por un lado, nos da la felicidad de saber que los ciudadanos se animan cada vez más a hacer reclamos; pero por el otro, nos hace dar cuenta que los casos de publicidad engañosa por parte de las compañías también es mayor”

Observemos la siguiente información que nos brinda un enunciado acompañado de un gráfico:

“Como se puede observar en la figura 3, las ventas a lo largo de un año de las empresa A y B siempre fueron parejas” (ver Figura 3).

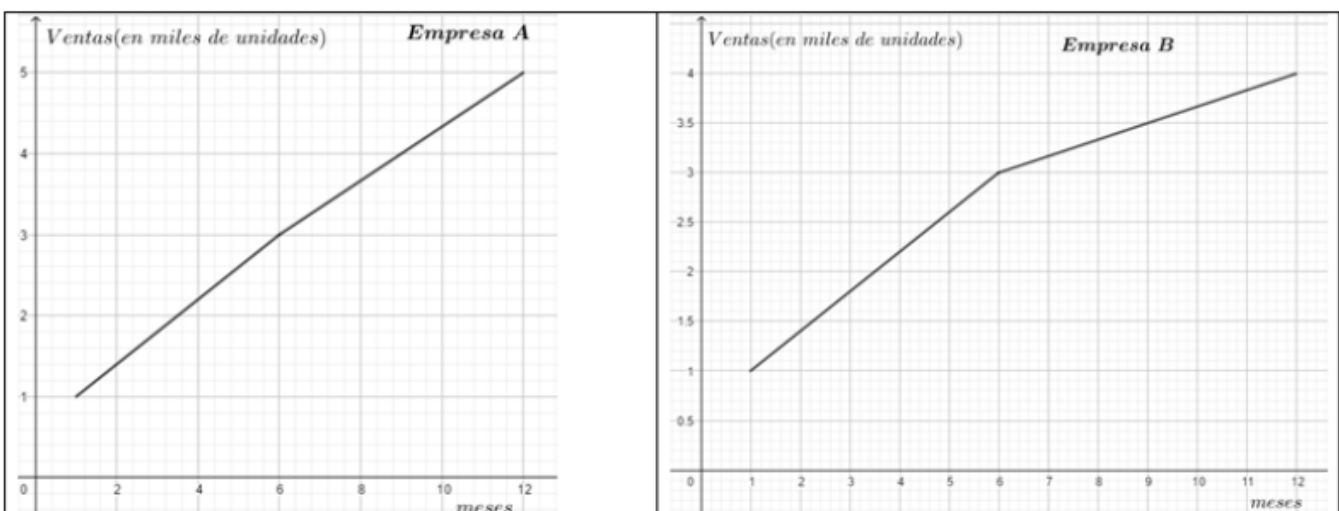


Figura 3. Ventas en un año de las empresas A y B

Si bien las ventas de la empresa A y B fueron las mismas durante los primeros seis meses, al cumplirse el año la empresa A vendió 1000 unidades más, ese efecto visual sencillamente se realiza cambiando las escalas del eje correspondiente a las ventas, es muy común al observar gráficos de este tipo quedarnos con la primera impresión sin analizar las escalas que se utilizan.

No hay dudas que el dominio de la estadística les puede brindar herramientas valiosísimas para poder interpretar críticamente la información, pero ¿Qué es la estadística?

La estadística es la ciencia que se encarga estudiar hechos y fenómenos de cualquier naturaleza con el fin de recolectar, organizar, analizar e interpretar la información para presentar o bien para realizar inferencias.

Se pueden distinguir dos ramas en la estadística, estas son la **estadística descriptiva** y la estadística inferencial.

En la estadística descriptiva, se trata de enunciar, analizar y comunicar ciertos parámetros que describen fenómenos que ocurren en las poblaciones.

En la estadística inferencial se trata de interpretar y analizar las variables estadísticas para sacar conclusiones, para ello se “apoya” en la teoría de la probabilidad. Un investigador con frecuencia desearía utilizar la información muestral para sacar algún tipo de conclusión (hacer una inferencia de alguna clase) con respecto a la población, es decir, la muestra es un medio para llegar a un fin en lugar de un fin por sí misma. Si conocemos la población en estudio podemos deducir información relevante de una muestra, esto es “ir” de lo general a lo particular, utilizamos el razonamiento deductivo. Si obtenemos una muestra representativa de la población a través de alguna técnica de muestreo podemos inferir alguna conclusión sobre la población, esto es “ir” de lo particular a lo general, utilizamos de esta forma el razonamiento inductivo. (ver Figura 4)

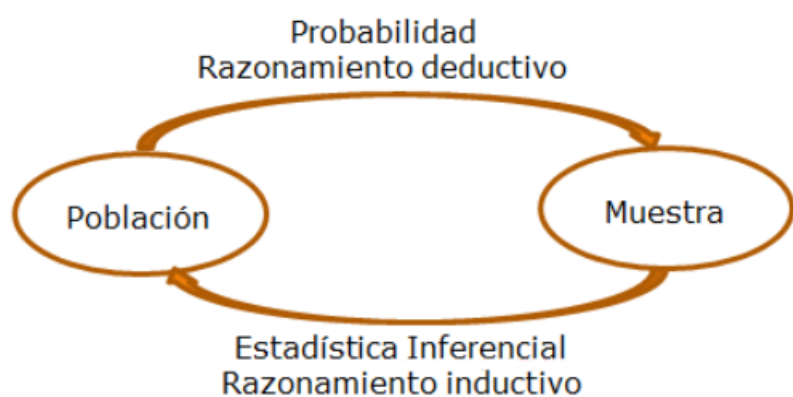


Figura 4. Relación entre la Probabilidad y la Estadística Inferencial



Pueden leer acerca de este tema en:

<https://ebevidencia.com/archivos/3568>



2. Algunos conceptos importantes de la estadística

2.1. Población

Es el grupo de todos los elementos sobre los cuales se observa una o más características que son de interés para realizar un análisis. Esta población es tomada como objeto de estudio, en razón de que sobre ella recae el interés del investigador. Por ejemplo: un conjunto de aspirantes al empleo de mozo en restaurantes de la zona costera de la provincia de Buenos Aires o conjunto de escuelas secundarias especializadas en construcciones o cantidad de peones empleados por hectárea de campo. Cuando la información deseada está disponible para todos los elementos de la población, se tiene lo que se llama **censo**.

2.2. Muestra

Es un subconjunto o subgrupo de la población representativa del total. Este subconjunto tiene que ser variado y representar a todos los individuos de la población, en el ejemplo del empleo del restaurante, una muestra podría ser: los 200 aspirantes que se presentaron a lo largo de una semana del mes de diciembre en 10 restaurantes costeros, o 50 escuelas secundarias especializadas en construcciones de diferentes partidos bonaerenses, o 10 establecimientos productores que tengan peones para el trabajo.

La obtención de la muestra no es tan sencilla, se deben seguir ciertas reglas y debe ser realizada por personal especializado. La teoría dice que una muestra, para que sea representativa y confiable, debe ser aleatoria, es decir, los elementos se deben elegir al azar, o dicho de otra forma, deben tener la misma probabilidad de ser escogidos. La elección no debe seguir ninguna norma prefijada. Cuanto más aleatoria sea, mejor y más confiable será la conclusión que podrá extenderse a toda la población.

Ejemplo de obtención de muestra:

Una fábrica quiere conocer la aceptación que tendrá un nuevo modelo de auto que están por sacar al mercado. Primero caracteriza el universo: segmento de la población que lo puede llegar a comprar por su poder adquisitivo, se analiza cuántos son hombres y cuántas son mujeres, se consideran diferentes tramos de edad a partir de los 21 años. Sobre la base de estos datos se determina a cuántas personas que reúnan determinados requisitos se entrevistarán y se elige al interior de cada caso los individuos al azar. Con estos resultados la empresa automotriz dirigirá su campaña publicitaria.

La selección representativa y confiable de la muestra es una actividad que requiere de conocimientos específicos para asegurar que los datos que se obtienen pueden aplicarse al total de la población. Esta tarea la realizan profesionales especializados no solo en estadística, sino especialmente en muestras. En general se los llama muestristas y en todos los casos indican el grado de confiabilidad del resultado obtenido, explicitando el porcentaje de error posible.

Es habitual que esta información se presente en una “ficha técnica” que acompaña los resultados de la indagación realizada, donde se describen sintéticamente las características de la muestra, los métodos utilizados, el error posible, etc.

Ejemplo de ficha técnica:

Empresa ejecutora: COP (Centro de Estudios de Opinión Pública)

- Tipo de estudio: encuesta de boca de comicios.
- Tipo de preguntas: cerradas, alternativas fijas y abiertas.
- Alcance de la muestra: nivel nacional.
- Tamaño de la muestra: submuestra especial en boca de comicio sobre un total de 2650 casos con un error de +/- 1.94% (confiabilidad del 95.5%)
- Fecha de realización del trabajo de campo: 14 de mayo de 1995.
- Procesamiento y análisis: 15 al 19 de mayo de 1995.

En nuestro país, el INDEC (Instituto Nacional de Estadísticas y Censos) es el organismo público, encargado de orientar y ejercer la dirección de todas las estadísticas oficiales que se realizan en el territorio. También coordina el Sistema Estadístico Nacional, integrado por los servicios estadísticos de los organismos nacionales, provinciales y municipales.

Es necesario hacer una distinción entre **estadístico y parámetro**. Un **estadístico** es cualquier característica numérica de la muestra mientras que un **parámetro** es cualquier característica numérica de la población.

Ejemplo: Para estudiar la población de trabajadores que fuman en una fábrica, un administrador tomó una muestra de 200 trabajadores y determinó el promedio (media aritmética) de trabajadores en la muestra que fuman. El promedio que obtuvo el administrador es un estadístico. Si hubiera tomado la totalidad de los trabajadores de la empresa y hubiera calculado el promedio de los que fuman, ese promedio representaría un parámetro poblacional.

2. 3. Variable

Es la característica observable en los individuos acerca de la cual se quiere estudiar algún aspecto. Por ejemplo: edades promedio de los ingresantes a un instituto de formación docente, inserción laboral que tienen los egresados de las escuelas secundarias especializadas en construcciones, estado civil de los empleados de una empresa menores de 40 años.

Cada variable tendrá distintos valores. Para el caso de las edades serán números, para la inserción laboral podrán ser categorías que indiquen orden: muy buena, buena, regular, mala; para el caso de estado civil serán clasificaciones: soltero, divorciado, viudo, etc.

Los valores de las variables también son determinados por el investigador en cuestión:

2. 4. Clasificación de variables

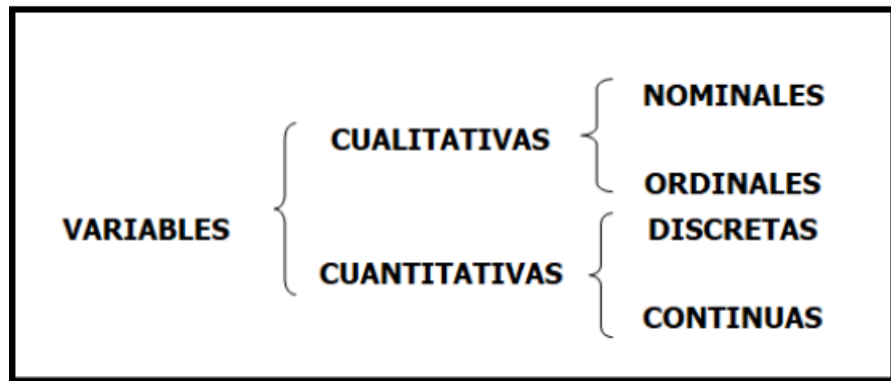


Figura 5. Clasificación de variables estadísticas.

Variables cualitativas: Son las que registran la presencia de un atributo, cualidad o categoría, estas deben ser mutuamente excluyentes y exhaustivas. Pueden ser cuantificadas de alguna manera.(ver figura 6, a cada categoría se le asignó un número, en ocasiones sirve para una mejor organización).

Mutuamente excluyente: Si una variable toma el valor de una cualidad no puede tomar el valor de otra cualidad diferente a la primera.

Exhaustivas: Deben abarcar todas las posibilidades.

Ejemplos:

Sexo				
	Hombre	1		
	Mujer	2		
Estudio formal				
	Primaria	1	Universitario	4
	Secundaria	2	Post universitario	5
	Terciario	3		

Figura 6. Ejemplos de variables cualitativas

Cualitativas nominales: No hay un orden natural. Entonces asignamos un “nombre”.

Ejemplos:

- 1) Fuma - No fuma
- 2) Grupo sanguíneo: A, B, AB, O

Cualitativas ordinales: Existe un orden natural.

Ejemplo:

- 1) Tipo de herida: Leve = 0, Moderada = 1, Grave = 2, Fatal = 3
- 2) Muy bueno, bueno, regular, malo. Estipula un orden, podemos decir que “muy bueno” es mejor que el resto de las categorías y “malo” es la peor de ellas.

Variables cuantitativas: Son las que expresan con cantidades numéricas, pueden ser discretas y continuas.

Cuantitativas discretas: Los números que asume son naturales o están en un conjunto finito; frecuentemente surgen por conteo.

Ejemplos:

- 1) Cantidad de hijos
- 2) Cantidad de terremotos en América del Sur.

Cuantitativas continuas: Es una medición que se expresa en unidades y que puede tomar, al menos teóricamente, un número infinito de valores dentro de un rango. En la práctica estos valores posibles dependen de la precisión del instrumento de medición o por la manera de medirla.

Ejemplos:

- Altura de jugadores de básquet
- Peso de boxeadores.
- Tiempo empleado en cierto experimento.

Antes de continuar los invitamos a mirar el siguiente video sobre los conceptos básicos de estadística.



VIDEO

"Conceptos básicos de estadística"

https://www.youtube.com/watch?time_continue=3&v=IYCjWcDTelQ



ACTIVIDAD 1

1) Indique y justifique cuáles de las dos muestras siguientes considera usted que son representativas del conjunto total observado (población), en este caso del total de alumnos de un CENS.

- a. De la población de alumnos del CENS se eligió un grupo de 10 alumnos de 2° año.
- b. En la misma escuela, una profesora propuso poner todos los nombres de los alumnos en una urna y luego sacar 10 al azar, como en los sorteos de los concursos.

2) Una empresa realiza una encuesta sobre intención de voto en una esquina de Buenos Aires muy transitada, los datos que obtiene son los siguientes de los 100 encuestados 40 aseguran votar al candidato A. El dato obtenido: ¿es un parámetro o es un estadístico?. Justifique.

3) Clasifique las siguientes variables en cualitativa ordinal o nominal y cuantitativa discreta o continua.

- a. Cantidad de veces que salió cara luego de revolver la moneda 10 veces.
- b. Temperatura promedio, en un día de verano elegido al azar, en la ciudad de Ushuaia.
- c. Número de bolillas blancas que se pueden extraer de un bolillero que contiene 3 rojas y 5 blancas cuando se extraen 4 bolillas.
- d. Número de hijos de una familia.
- e. Nivel de estudios alcanzado por la madre en una familia: primaria incompleta, primaria completa, secundaria incompleta, secundaria completa, terciaria incompleta, terciaria completa, universitaria incompleta, universitaria completa.
- f. La edad de los empleados en un establecimiento avícola.
- g. Preferencias de tres marcas de automóviles: ford, peugeot, fiat.
- h. Tiempo que demora una persona para llegar a su destino de vacaciones.
- i. Tipos de quemaduras: menores, moderadas, graves.
- j. Volumen de agua que hay en una piscina.
- k. Notas de un exámen: excelente, aprobado, regular, desaprobado.



3. Frecuencias

Una vez que están recolectados los datos, debemos organizarlos para inferir y sacar conclusiones posibles acerca de ellos.

Un instrumento para conseguir esa ordenación es la llamada **distribución de frecuencias**.

La frecuencia es la cantidad de veces que se repite un suceso, por ejemplo: cantidad de hijos de una familia en una muestra de 20 familias de San Pedro.

Una distribución de frecuencias debe cumplir tres funciones:

1. Organizar y ordenar racionalmente los datos reunidos.
2. Ofrecer información suficiente para representarla en un gráfico.
3. Facilitar los cálculos necesarios para realizar las inferencias y extraer conclusiones.

La frecuencia puede ser:

- **Frecuencia absoluta:** Es el número de veces que se repite un valor en la muestra. Generalmente la simbolizamos con f_a .
- **Frecuencia relativa:** Es el cociente entre la frecuencia absoluta y el tamaño total de la muestra. La suma total de ellas da 1. Generalmente la simbolizamos como f_r .

Veremos más adelante que el concepto de frecuencia relativa se encuentra muy ligado al concepto de probabilidad.

En ocasiones interesa observar las frecuencias acumuladas hasta cierto valor de la variable, les sirve también para chequear si la tabla está correctamente construida, generalmente se utilizan dos:

Frecuencia absoluta acumulada: Indica cuántos datos se van contando hasta ese momento, el último valor de la columna debe dar la totalidad de los datos. La simbolizamos con F_a .

Frecuencia relativa acumulada: Es la proporción de datos respecto al total calculada hasta cierto valor de la variable, el último valor de la columna debe dar 1. La simbolizamos como F_r .

Otra información relevante que nos proporciona una tabla de frecuencia son los porcentajes, nos sirven para facilitar la comprensión y detallar mejor la información, estos porcentajes los obtenemos al multiplicar las frecuencias relativas por 100 ($f_r \times 100$), las denominamos frecuencias porcentuales y el resultado de sumarlas es 100.

Veamos un ejemplo de cómo se organiza una TABLA DE FRECUENCIAS según el caso de la variable “cantidad de hijos” en una muestra de 20 familias.

X (cantidad de hijos)	f_a (frecuencia absoluta)	F_a (frecuencia absoluta acumulada)	f_r (frecuencia relativa)	F_r (frecuencia relativa acumulada)	$f_{\%}$ (frecuencia porcentual)
4	1	1	0,05	0,05	5
3	3	4	0,15	0,20	15
2	7	11	0,35	0,55	35
1	6	17	0,30	0,85	30
0	3	20	0,15	1	15
TOTAL	20		1		100

De la observación de la tabla podemos observar que 7 familias tienen 2 hijos, es decir que la mayor cantidad de familias de la muestra tienen 2 hijos, esto es equivalente a decir que el 35% de las familias entrevistadas tienen 2 hijos, seguidamente se observa que 6 familias tienen 1 hijo, esto es equivalente al 30%. Generalmente en las tablas se buscan datos destacados como frecuencias máximas o mínimas, notemos que solo una familia tiene 4 hijos, esto representa el 5% de la muestra y en el otro extremo observamos que hay 15 familias que no tienen hijos, esto equivale al 15% de la muestra.

Pensemos un momento sobre qué ventajas aporta la tabla en la organización de los datos, realmente se complicaría buscar información relevante si organizar los datos previamente.

¿Cuáles son los pasos a seguir para construir una tabla de frecuencias?

1) En la primera columna colocar los valores que pueden tomar la variable, si existe orden, conviene realizarla en orden ascendente o descendente

2) Realizar el recuento para completar la columna de las frecuencias absolutas. Como estrategia podemos ir tachando los datos a medida que los contamos.

3) Dividir cada frecuencia absoluta por el total de datos que tenemos para completar la columna de frecuencias relativas.

4) Multiplicar por 100 cada valor de la frecuencia relativa y completar la columna de frecuencias porcentuales.

Vemos el siguiente ejemplo:

Los siguientes valores recolectados corresponden a cantidad de horas trabajadas por un grupo de 25 operarios de una fábrica en un día determinado:

7 - 8 - 10 - 6 - 8 - 6 - 10 - 8 - 8 - 7
6 - 9 - 8 - 7 - 6 - 8 - 10 - 9 - 8 - 7
7 - 8 - 10 - 8 - 7

• Procedemos a ordenar los datos: descubrimos que el menor es (6) y el mayor es (10), entonces los valores de la variable irán desde 6 hasta 10.

- Contamos cuántas veces se repite cada uno:

6: //// (4 veces)

7: ////// (6 veces)

8: ////////// (9 veces)

9: // (2 veces)

10: //// (4 veces)

- Armamos la tabla:

X (Cantidad de horas trabajadas)	f_a (frecuencia absoluta)	F_a (frecuencia absoluta acumulada)	f_r (frecuencia relativa)	F_r (frecuencia relativa acumulada)	$f_{\%}$ (frecuencia porcentual)
6	4	4	0,16	0,16	16
7	6	10	0,24	0,40	24
8	9	19	0,36	0,76	36
9	2	21	0,08	0,84	8
10	4	25	0,16	1	16
TOTAL	25		1		100

De la tabla de frecuencias podemos observar que de la muestra obtenida, el mayor porcentaje de los operarios trabaja 8 horas diarias, esto equivale al 36% del total de operarios, seguido del 24% de los operarios que trabajan 7 horas diarias. La mayor y menor cantidad de horas trabajadas se divide equitativamente cada una con un 16% del total de los operarios. Si se observa la columna de frecuencias relativas acumuladas nos indica que el 76% de los operarios no exceden las 8 horas diarias trabajadas.

Si hasta acá se va comprendiendo, le sugerimos que observen el siguiente video:



VIDEO

"Tablas de Frecuencias"

www.youtube.com/watch?time_continue=1&v=cyXenZEBGz4





ACTIVIDAD 2

Obligatoria

Los siguientes datos corresponden a la nota de 20 alumnos de una evaluación de matemática tomada en un CENS de la provincia de Bs As:

9 - 7 - 8 - 6 - 9 - 10 - 7 - 8 - 9 - 9
7 - 9 - 9 - 10 - 8 - 9 - 9 - 10 - 9 - 7

Se pide:

- 1) Armar una tabla de frecuencias (absolutas, relativas, porcentuales y acumuladas)
- 2) ¿Cuál es porcentaje de la calificación con mayor frecuencia?
- 3) ¿Cuál es el porcentaje de la calificación con menor frecuencia?
- 4) Los alumnos que saquen una nota menor a 8 deberán presentar un trabajo práctico, ¿qué porcentaje de alumnos deberán realizar el trabajo práctico?

» » » » »
Continuamos con la lectura del apunte

4. Intervalo de Frecuencias

Al trabajar con muestras más extensas tendríamos más valores para la variable, por ejemplo si quisiéramos hacer un análisis de los tiempos empleados por 100 maratonistas del Municipio de Lomas de Zamora en una prueba de atletismo. Sería muy difícil trabajar con todas las posibilidades (aún con ayuda de medios informáticos).

Para poder trabajar con valores de este tipo se establecen intervalos de frecuencias, que son agrupaciones de valores entre dos números, teniendo un valor inferior (el menor de dos números escritos). Estos valores se eligen dependiendo del análisis que se quiera hacer de los datos.

Por ejemplo: Supongamos que se quiere analizar la edad de los alumnos de un grupo de Educación de Adultos a Distancia. Se obtienen las siguientes edades de 100 alumnos ordenados alfabéticamente por apellido:

51	35	36	41	33	28	57	62	43	69
42	33	62	53	37	36	48	53	39	41
19	47	18	62	33	54	29	35	61	60
27	31	36	44	45	50	21	52	59	52
43	49	23	37	28	27	41	35	37	28
50	34	22	31	34	43	32	36	25	30
33	25	28	31	36	25	41	44	38	51
50	46	26	40	53	36	31	34	51	65
32	34	43	63	49	61	48	38	16	41

Observando los datos notamos que el menor valor es 16 y el mayor es 69. Si quisiéramos realizar una tabla con todas las edades desde los 16 años hasta los 69 años con un incremento de un año por vez, nos quedaría muy extensa con 54 filas, este hecho no contribuiría a facilitar la observación ni el estudio de los datos. En casos como este, con numerosa cantidad de datos, se los agrupa en intervalos y se los presenta en una tabla abreviada.

Hay diferentes formas de realizar los intervalos, en este módulo explicaremos una de ellas:



1° Calcular el rango o amplitud de los datos, a este valor lo designaremos con la R, consiste en la diferencia entre el mayor y el menor valor de los datos:

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

$$R = 69 - 16 = 54$$

2° Buscamos la cantidad de intervalos dividiendo convenientemente el rango por un valor que denominamos **amplitud de intervalo (I)**, en nuestro caso, el valor adecuado es **I = 6** ya que es divisor de **54**. Esto nos dá:

$$R/I = 54/6 = 9 \text{ intervalos}$$

Si bien no hay una regla para decidir la cantidad de intervalos, y como la amplitud 6 la decidimos nosotros hay que tener en cuenta que no es prudente tomar muchos intervalos porque nos encontraríamos con el mismo problema del comienzo.

3° Armamos los intervalos comenzando desde el valor más pequeño y tomando los valores extremos de 6 en 6. (**Importante: Tener en cuenta que los valores no se deben repetir de un intervalo a otro, es decir, que son mutuamente excluyentes entre dos intervalos**).

X (intervalo de edades)	f_a (frecuencia absoluta)	F_a (frecuencia absoluta acumulada)	f_r (frecuencia relativa)	F_r (frecuencia relativa acumulada)	$f_{\%}$ (frecuencia porcentual)
16 - 21	4	4	0,04	0,04	4
22 - 27	9	13	0,09	0,13	9
28 - 33	21	34	0,21	0,34	21
34 - 39	22	56	0,22	0,56	22
40 - 45	16	72	0,16	0,72	16
46 - 51	11	83	0,11	0,83	11
52 - 57	7	90	0,07	0,90	7
58 - 63	8	98	0,08	0,98	8
64 - 69	2	100	0,02	1	2
TOTAL	100		1		100



Continuamos con la lectura del apunte

De la tabla podemos observar como primer dato sobresaliente que 59% de los alumnos tienen entre 28 y 45 años, la mayor cantidad se encuentra en el rango de edades de 28 a 39 años, quizás podamos sacar algunas conclusiones de por qué se dan estos datos, pero no sería prudente con solo mirar esta muestra. También observamos que para edades más bajas y más altas los porcentajes de alumnos disminuyen.

Antes de continuar les recomendamos mirar el siguiente video



"Intervalos de Frecuencias"

<https://www.youtube.com/watch?v=CuKr7Gzohbl>

ACTIVIDAD 3 Obligatoria

Un grupo de 20 ciclistas se está preparando para competir en una competencia y como parte de su entrenamiento se someten a una dieta estricta, a continuación se muestran los pesos en kilogramos que lograron bajar los cada uno de los ciclistas:

0,4 - 8 - 7,2 - 6,9 - 3,2 - 4,8 - 8,9 - 4,5 - 3,2 - 1,8 - 6,1 - 4,5 - 15,8 - 5,7 - 19,2 - 6,7 - 9,2 - 9,1 - 12,8 - 3

1) Realizar una tabla de frecuencias de 5 intervalos de amplitud 4. (Sugerencia el primer intervalo puede ser $[0, 4)$, siempre considerar el valor inferior del intervalo dentro del mismo y el superior fuera)

2) Responder:

- ¿En qué intervalos de pesos se registra la mayor cantidad de ciclistas?
- ¿Qué porcentaje de ciclistas bajo más de 12 kilos?
- ¿Qué porcentaje de ciclistas bajo hasta 12 kilos?
- ¿Que cantidad de ciclistas bajo entre 0 y 4 kilos?



Continuamos con la lectura del apunte

5. Gráficos

Un recurso muy utilizado para la presentación de la información son los gráficos estadísticos, mediante estos podremos visualizar de forma más clara lo que se pretende describir o analizar. Existen muchas variedades de gráficos, antes de decidirse por un gráfico es necesario pensar en cual es mejor para esa tarea ya que cada uno tiene características específicas. Un buen gráfico estadístico nos brinda la información más relevante con solo una mirada, sin necesidad de tener las dimensiones del problema que está tratando, por supuesto si la persona que lo mira es especialista en lo que se pretende informar podrá realizar un análisis mucho más profundo.

Entre los gráficos podemos encontrar:



5. 1. Diagrama de rectángulos

Se utilizan para representar variables nominales. En el eje x (de abscisas) se colocan los valores de la variable y en el eje y (de ordenadas) las frecuencias. En la Figura 7 podemos observar de una muestra de 13 personas, grupo de solteros y de casados.

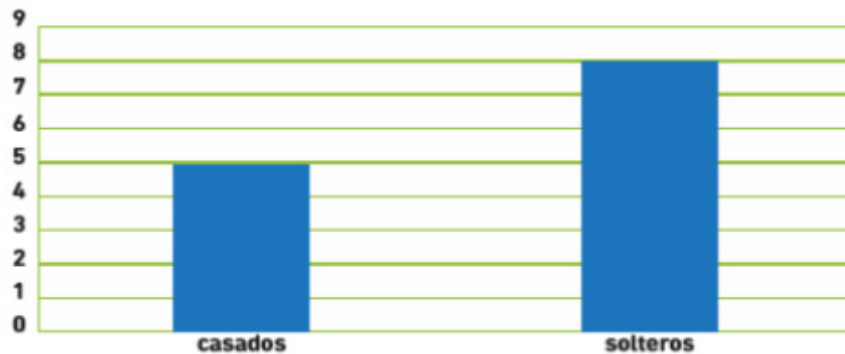
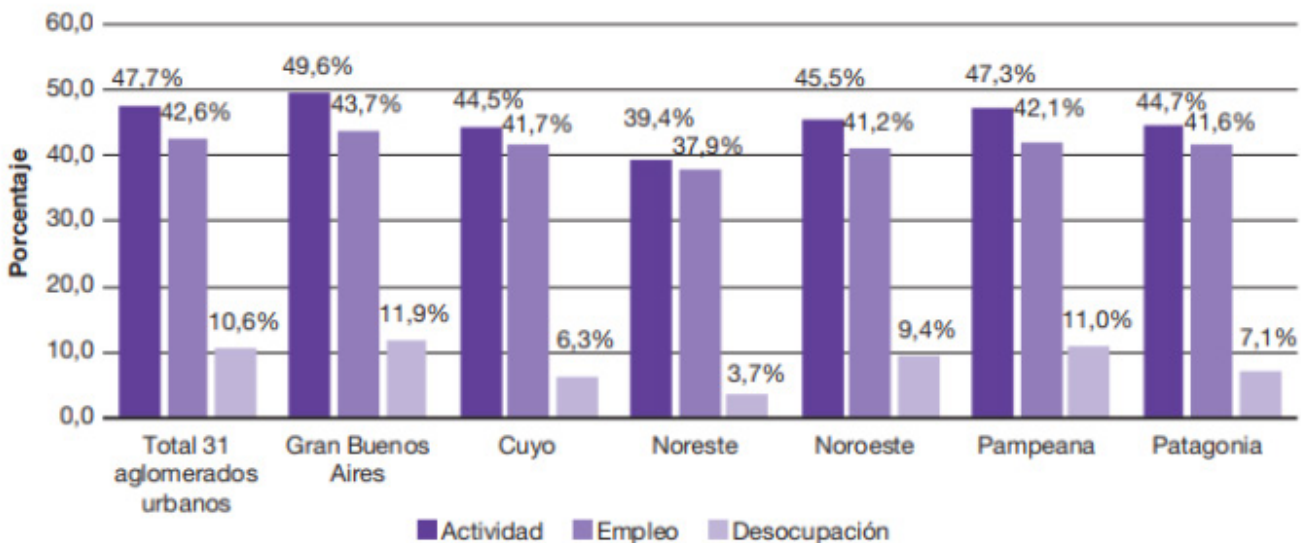


Figura 7. Diagrama de rectángulos.

En ocasiones los diagramas de rectángulos se los utiliza para efectuar comparaciones, para eso se los construye con diferentes colores por variables que se quieran mostrar, en la figura 8, se comparan el nivel de actividad, de empleo y desocupación por regiones.

Principales tasas en el total de 31 aglomerados urbanos y regiones. Segundo trimestre de 2019



Fuente: INDEC. Encuesta Permanente de Hogares.

Figura 8. Diagrama de rectángulos comparativos.

5. 2. Diagrama circular, de torta o pictograma

Sirve para representar cualquier tipo de variable. Se hace el cálculo de la amplitud del ángulo correspondiente a cada variable teniendo en cuenta que el 100% es igual a 360° .

Ejemplo: Si queremos representar la variable nivel de estudios (sin hacer distinción de estudios completos o incompletos), podemos considerar las siguientes variables: primario, secundario y terciario. Luego armamos la tabla siguiente:

X (Nivel de estudios)	f_a (frecuencia absoluta)	f_r (frecuencia relativa)	Amplitud del ángulo del sector: $360^\circ \times f_r$
Sin estudios	1	0,1	$0,1 \times 360^\circ = 36^\circ$
Primario completo	5	0,5	$0,5 \times 360^\circ = 180^\circ$
Secundario completo	3	0,3	$0,3 \times 360^\circ = 108^\circ$
Terciario completo	1	0,1	$0,1 \times 360^\circ = 36^\circ$
Total	10		360°

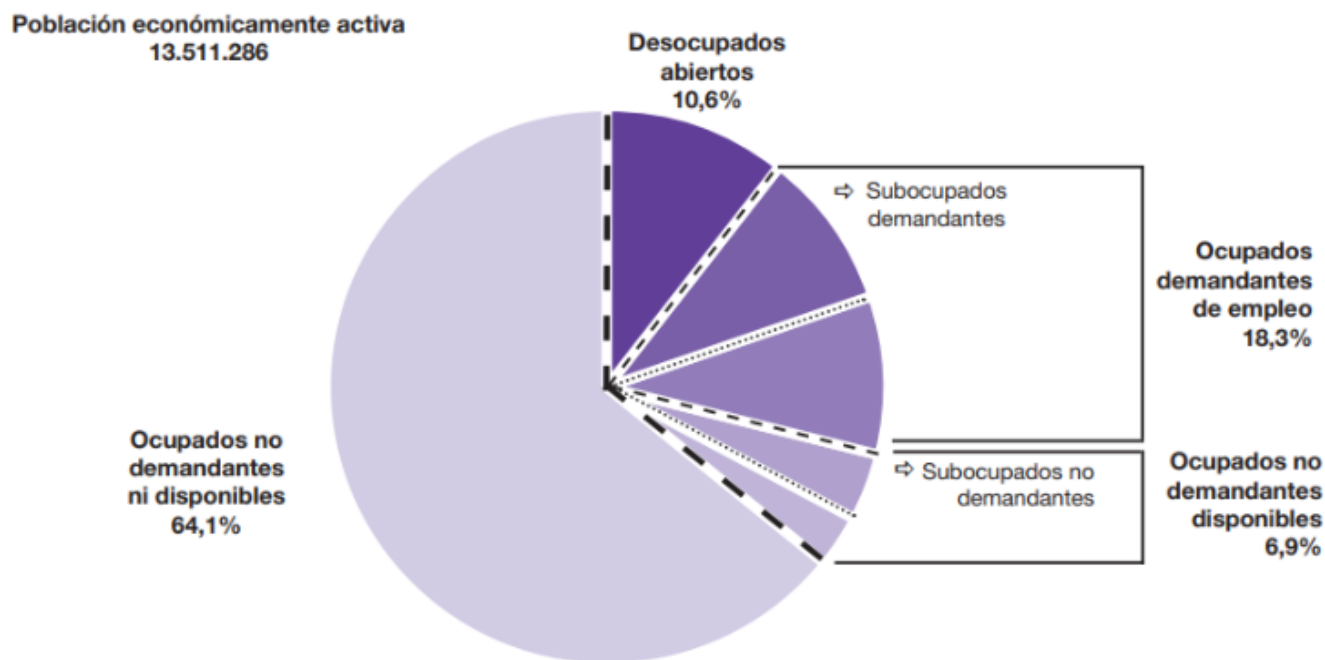
Para calcular el ángulo de amplitud de cada sector multiplicamos la frecuencia relativa por 360° y establecemos el ángulo aproximado. En la figura 9 podemos observar el gráfico terminado.



Figura 9 Diagrama circular: “Nivel de estudios”

Observemos a continuación (Figura 10) un diagrama circular un poco más complejo que nos proporciona el INDEC referente a la desocupación en el 2° trimestre del año 2019.

Grupos de población económicamente activa según tipo de presión sobre el mercado de trabajo.
Total 31 aglomerados urbanos. Segundo trimestre de 2019



Fuente: INDEC. Encuesta Permanente de Hogares.

Figura 10. Ocupación y desocupación, 2° trimestre de 2019

“La población económicamente activa (PEA) está compuesta por los ocupados y los desocupados. Este último grupo presiona activamente sobre el mercado laboral en búsqueda de una ocupación. A fin de tener un indicador sobre la presión global, se suman los ocupados que, si bien tienen un empleo, buscan activamente otro (ocupados demandantes). En una gradación menor de presión laboral, se encuentran aquellos ocupados que no demandan activamente otro empleo, pero están dispuestos a extender su jornada de trabajo. Ejemplo de estos son los subocupados no demandantes y otros ocupados no demandantes disponibles. Los grupos de población que aparecen en la figura 8 están representados según el tipo de presión sobre el mercado laboral”.

INDEC. Encuesta permanente de Hogares.

Si observamos la figura 8 observamos que de la población económicamente activa el 10,6% representan los desocupados, esto representaría 13511286×0.106 , aproximadamente 1432196 personas desocupadas.



VIDEO

Para explorar este tema, pueden visitar:

“Diagrama de sectores - Probabilidad y estadística - Educatina”

<https://www.youtube.com/watch?v=p7PpVZAQxPA>

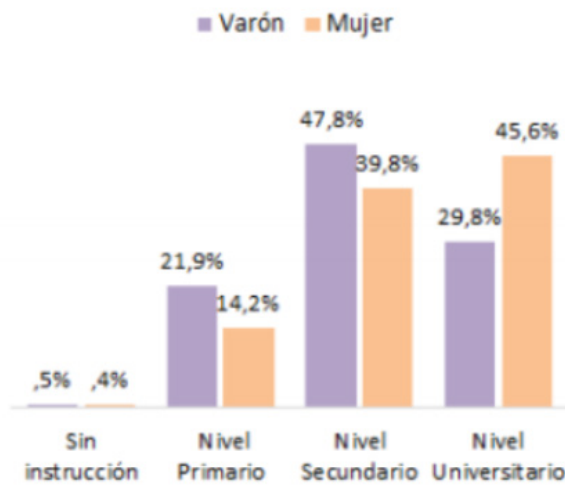




ACTIVIDAD 4

Dados los gráficos:

Nivel educativo de la población ocupada por género.



Fuente: Elaboración propia en base a la Encuesta Permanente de Hogares (EPH): 3er trimestre de 2017

Las mujeres realizan la mayor parte del trabajo doméstico no remunerado



Fuente: Encuesta sobre uso del tiempo y trabajo no remunerado, Argentina, 2013.

Observar cada gráfico y determinar para cada uno:

- 1) Tipo de gráfico, variable graficada, valores para la variable.
- 2) Valor mayor según lo representado en cada gráfico.
- 3) Relacionar ambos gráficos y extraer conclusiones estadísticas, teniendo en cuenta los siguientes puntos:
 - a. ¿Cuál es la relación entre el trabajo doméstico no remunerado y el nivel educativo en las mujeres? ¿Qué dato o datos estadísticos agregaría para complementar dicho análisis?
 - b. ¿Qué características encuentran en el caso de las mujeres que lograron completar el nivel universitario?
 - c. ¿Cómo relaciona el nivel primario y secundario con el nivel de ocupación en varones y mujeres?

Continuamos con la lectura del apunte

5. 3. Histogramas

Se usan para variables cuantitativas continuas pueden estar agrupadas en intervalos o sin agrupar. En el eje de abscisas colocamos los valores (agrupados o no) y en el de ordenadas las frecuencias. En la figura 11 se observa el nivel de colesterol en mg/dl de una muestra de personas en un rango de edades desde los 40 años hasta los 70 años, se puede observar que la amplitud de los intervalos es de 5 años, los niveles de colesterol con mayor frecuencia es la correspondiente al intervalo de 50 a 55 mg/dl.



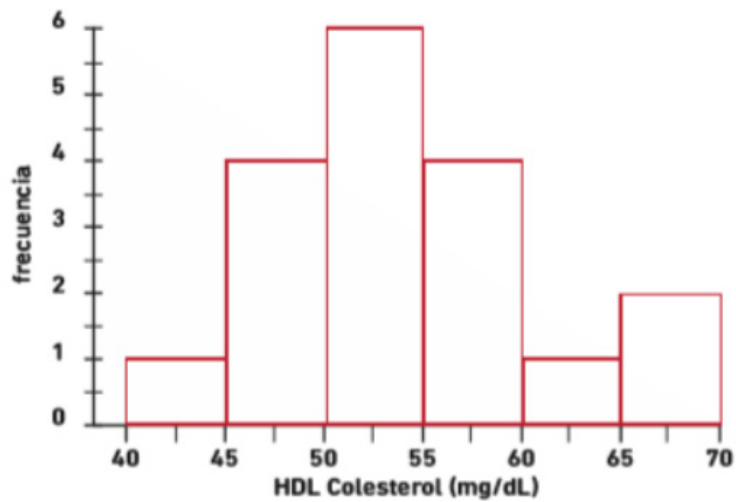


Figura 11. Histograma: "HDL Colesterol"

5. 4. Polígono de Frecuencias

Se usa para variables discretas o continuas, resulta de la unión de los puntos correspondientes a la frecuencia de la variable, en el caso de datos agrupados se selecciona el punto medio de los intervalos de clase.

X (número de hijos)	f_a (frecuencia absoluta)
0	3
1	6
2	7
3	2
4	2
TOTAL	20

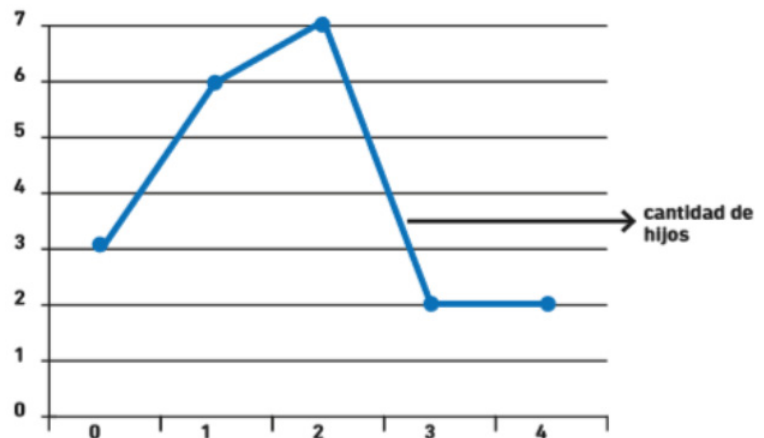


Figura 12. Polígono de Frecuencias: "Cantidad de hijos"

Si bien se trata de una variable aleatoria discreta el objetivo de unir los puntos es para dar una mejor visualización de la evolución de la variable, en la figura 13 se observa que a medida que la variable cantidad de hijos aumenta ("eje de abscisas") la frecuencia disminuye ("eje de ordenadas").



Fuente: elaboración propia en base a Ministerio de Hacienda e IPC-DGEyC CABA.

Figura 13. Polígono de Frecuencia: "Presupuesto destinado a deuda"

En la figura 11 se observa una variante de un polígono de frecuencias marcando frecuencias relativas en forma de porcentajes, se busca de esta forma indicar como desde el año 2015 al 2019 se incrementó de manera considerable parte del dinero del presupuesto a pagar intereses de la deuda.

5. 5. Perfil ortogonal

Se utilizan para comparar la evolución de una o más variables, son muy utilizadas para comparar la evolución de variables relacionadas por alguna característica.

Ejemplo: La figura 14 muestra la evolución de las calificaciones de un estudiantes a lo largo de los bimestres de un año.

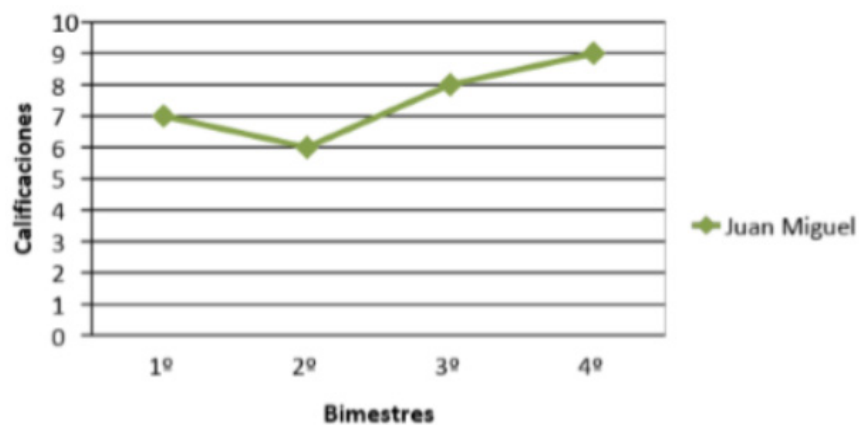


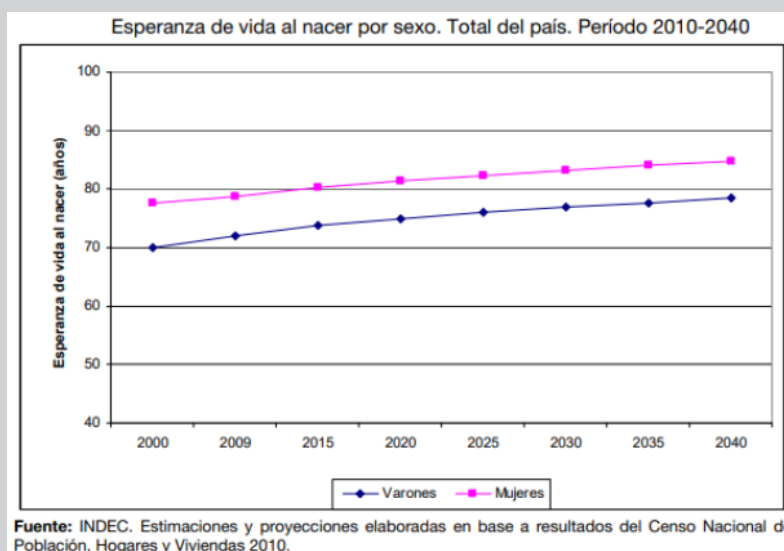
Figura 14. Perfil ortogonal: “Calificaciones de Juan Miguel en los bimestres”

En muchas ocasiones interesa realizar proyecciones de variables importantes para las poblaciones, los invitamos a realizar la siguiente actividad donde se compara la esperanza de vida por sexo en Argentina.

ACTIVIDAD 5

De acuerdo a los siguientes gráficos, observarlos y responder las consignas en cada uno de los casos propuestos El primer gráfico nos indica la esperanza de vida al nacer por sexo.

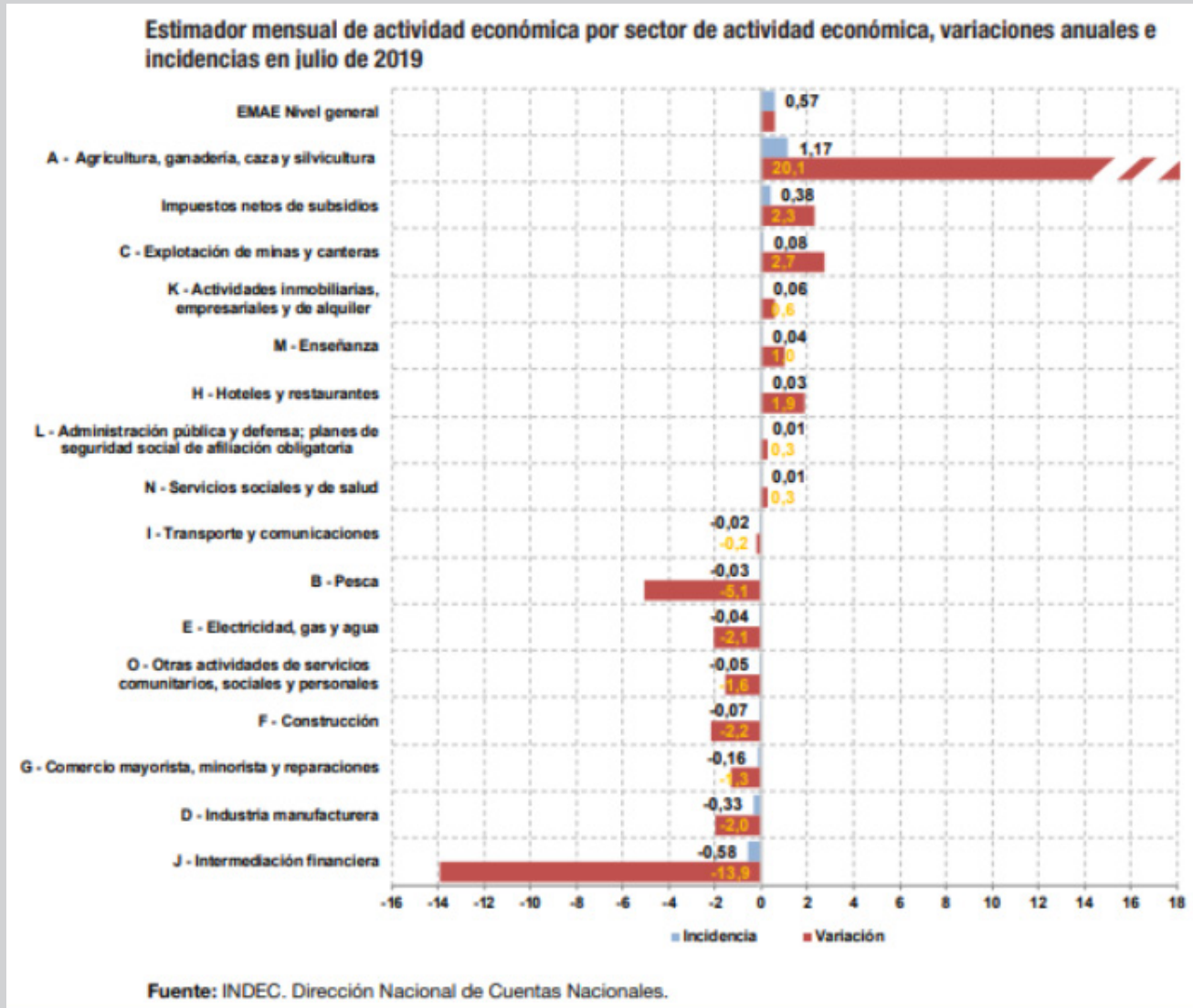
- ¿Qué sexo posee mayor esperanza de vida?
- ¿A partir de qué año la esperanza de vida de las mujeres superó los 80 años?
- ¿La esperanza de vida de los hombres supera los 80 años?



El siguiente gráfico se trata del EMAE (Estimador mensual de la actividad económica). Este indicador muestra el comportamiento de 16 sectores económicos. Dichos datos corresponden al período Julio 2019 comparado con el mismo mes del año 2018. Observar el gráfico y contestar las siguientes preguntas:

Sugerencia: En el gráfico la línea “imaginaria” del centro representa al cero, a la derecha se observan los niveles de crecimiento (los cuales se encuentran en la mitad superior del gráfico), mientras que a la izquierda se observan los valores de decrecimiento de las variables (los mismos se encuentran en la mitad inferior del gráfico).

- ¿Qué sector tuvo la mayor variación positiva?
- ¿Qué sector tuvo la mayor variación negativa?
- ¿Cuál fue la variación de la industria manufacturera?
- ¿Cuál fue el cambio del sector comercio mayorista, minorista y reparaciones?



¿Alguna actividad que no sabes cómo resolverla? Los espera el tutor en el Campus Virtual o en el encuentro presencial para acompañarlos y ayudarlos.





¡ Empezamos a estudiar!

Conozcamos un poco más sobre este tema leyendo el apunte de clase que se encuentra a continuación.

Apunte de clase: Estadísticos Muestrales



1. Estadísticos Muestrales

Existen valores que describen características importantes de la población que se está estudiando, los denominamos parámetros poblacionales, estos pueden ser: la media poblacional (μ), la dispersión poblacional (σ), la proporción poblacional (p) etc . Estos valores generalmente son desconocidos ya que para obtenerlos necesitamos recolectar la totalidad de los datos, es lo que conocemos como censo, esta acción es en muchos casos muy difícil de realizar ya sea por factores operativos o económicos, por ello en los estudios estadísticos se trabaja con muestras, es donde aparecen los **estadísticos muestrales, los utilizamos para estimar puntualmente a los parámetros poblacionales.**

Estadístico muestral: Se llama estadístico muestral a cualquier función de la muestra, si tenemos una muestra aleatoria de la variable aleatoria X , un estadístico es una función de ella: $f(X_1, X_2, \dots, X_n)$. En otras palabras, es una medida cuantitativa que deriva de un conjunto de datos de una muestra con el objetivo de estimar un parámetro poblacional, que nos brinda información importante de la población en estudio.

Los invitamos a ver el siguiente video donde se explica el concepto de estadístico muestral:



VIDEO

"¿Qué es un estadístico?"

<https://www.youtube.com/watch?v=mCGnj16wvFA>

2. Medidas de Posición y de Tendencia Central

2.1. Medidas de Posición: Deciles, cuartiles y centiles

Estas medidas indican la posición dentro del grupo, es decir, qué lugar ocupa en el conjunto.

Supongamos que nos dicen que un aspirante a obtener un empleo obtuvo en una prueba de desempeño la puntuación 35. Ese número, por sí mismo, no nos dice nada. Si el máximo fuera 35, entonces el aspirante habría obtenido la puntuación máxima, pero si el máximo era de 100, la puntuación obtenida evidentemente no es la mejor.... Necesitamos referencias para que el 35 nos brinde información adicional. **Las medidas de posición más usuales reciben el nombre de cuantiles. Los cuantiles más usados son:**



- **Centiles:** Los centiles también llamados **percentiles** son los 99 valores de la variable que dividen la muestra en 100 secciones, cada una conteniendo a la centésima parte de las observaciones. Una vez ordenados los datos de menor a mayor, el centil indica el valor de la variable por debajo del cual se encuentra un porcentaje dado de observaciones del total, por ejemplo si 12 es centil 20, esto indica que el 20 % de los datos totales son valores menores o iguales a 12

En nuestro ejemplo, suponiendo que el número 35 corresponde al centil 90, sabremos que es una puntuación que supera al 90 de las 100 observaciones realizadas y solo es superada por 10 de las 100, en este caso sería una buena puntuación.

$$C_{90} = 35$$

- **Deciles:** Similar definición a los centiles, en este caso son los 9 valores que dividen a la muestra en 10 secciones con un décimo de la cantidad total cada una.

El decil se representa como D. Existe una equivalencia directa entre centiles y deciles. Al C_{90} equivale el D_9 .

$$D_1 = C_{10}$$

$$D_2 = C_{20}$$

- **Cuartiles:** Son valores que dividen a la muestra en 4 partes iguales, conservando cada una el 25% del total de la muestra.

Se designa como Q y también tiene equivalencias con deciles y centiles.

$$Q_1 = C_{25}$$

$$Q_2 = C_{50} = D_5$$

$$Q_3 = C_{75}$$

Ejemplo: Quizás alguno haya vivido la experiencia de acompañar a sus hijos la pediatra y observar en alguna ocasión que luego de medir y pesar al niño utilice unas tablas de crecimiento, pero: ¿que son esos gráficos, que miraba la pediatra y que marcaba?, justamente son tablas de crecimiento (peso, altura) de niños y niñas desde que nacen hasta los 36 meses, están realizadas por especialistas, en ellas aparecen los percentiles (figura 13).

Hay muchos y variados factores que influyen en el crecimiento de un niño, la genética, el sexo, la nutrición, la actividad física, problemas de salud, el ambiente y las hormonas. Desde ya es una medida de control y prevención y de ninguna manera indican que un niño que esté en un percentil bajo tenga alguna enfermedad.

¿Qué puede indicar un problema?

- Cuando el percentil del peso o la altura de un niño cambia de un patrón que ha estado siguiendo a otro.
- Cuando los niños no crecen con la misma velocidad a la que engordan.

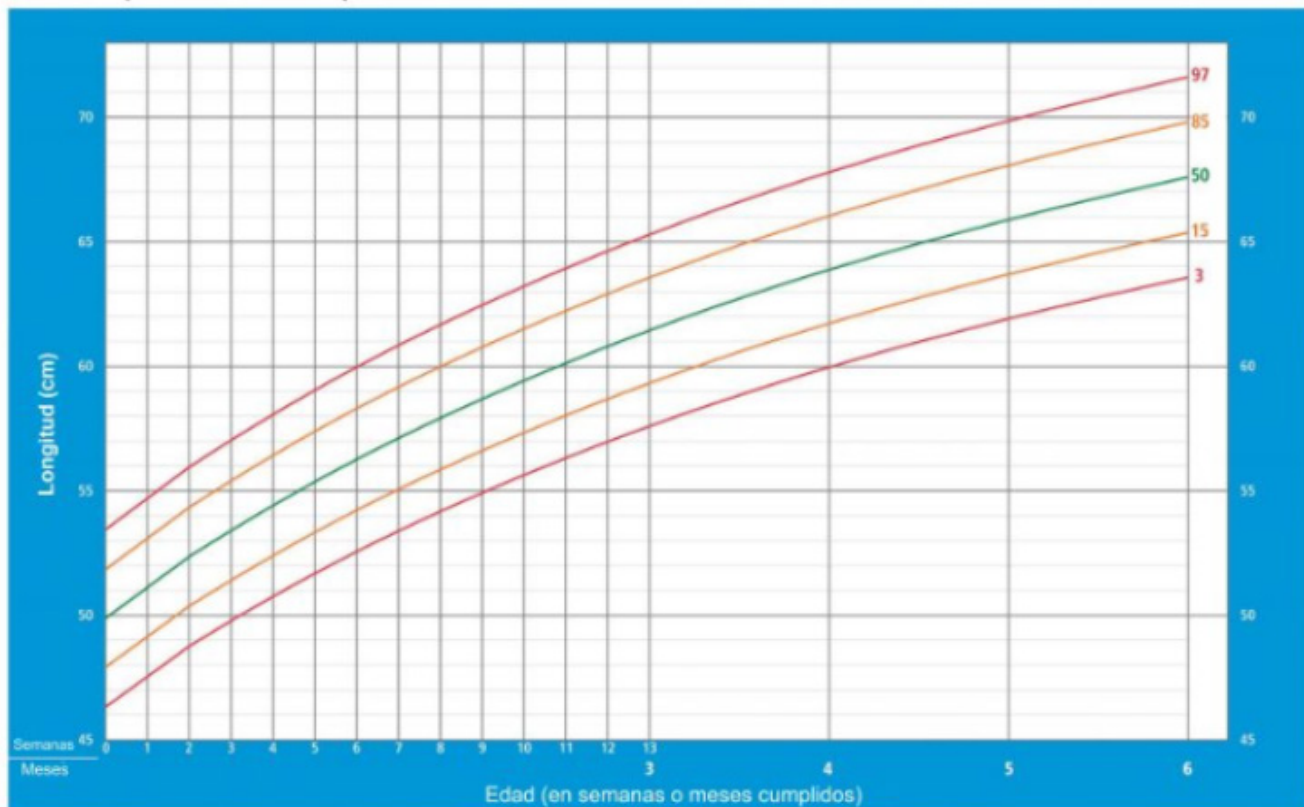
Veamos algún ejemplo:

Estadísticamente una niño que tiene una estatura que coincide con el percentil 30, nos dice que de la población de referencia, es más alto que el 30% de su edad y más bajo que el 70% de los niños de su edad.

En la figura 15 se muestra la tabla para las estaturas de niños hasta los 6 meses, notemos que en el eje de las abscisas ("x") hay una escala hasta 13 semanas (3 meses ya cumplidos) y luego se marcan los meses. Los percentiles están marcados en color, el de color verde es el percentil 50, aparece en todos los gráficos ya que se trata del valor medio. El pediatra toma la talla del niño e ingresa a la tabla utilizando esos dos valores (edad, talla).

Longitud para la edad Niños

Percentiles (Nacimiento a 6 meses)



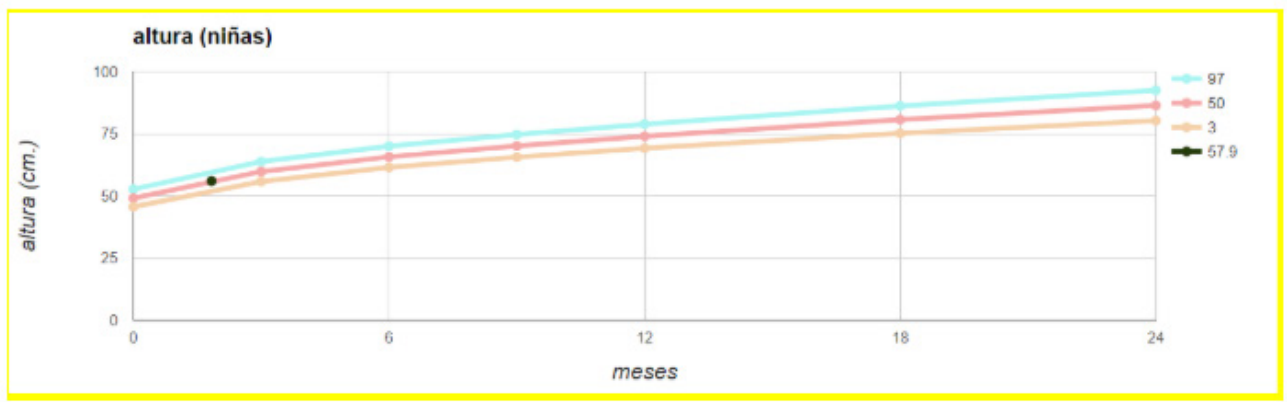
Patrones de crecimiento infantil de la OMS

Figura 15. Tablas de longitud para niños de OMS

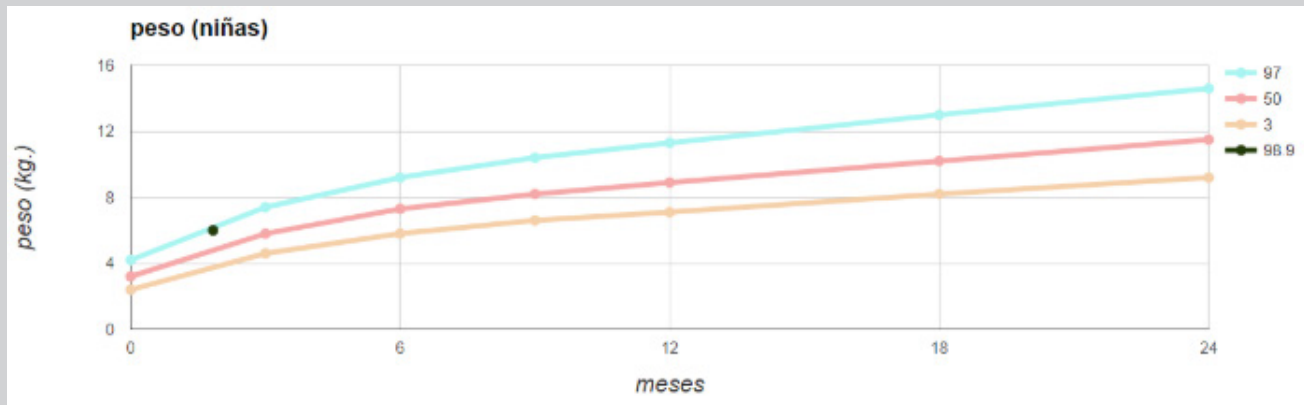


María Eva es una beba que tiene 8 semanas de edad, mide 56 cm, pesa 6 kg y su perímetro craneal es de 35 cm. Se adjuntan los percentiles de edad, peso, índice de masa corporal y perímetro craneal. En cada uno de los gráficos el punto negro es donde se encuentra María Eva, para ello debemos observarlos y contestar las consignas.

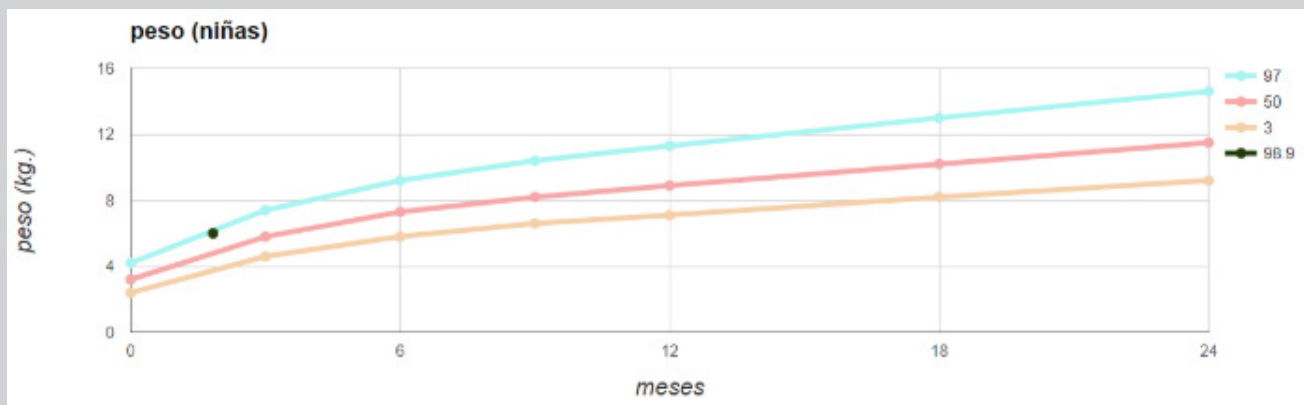
1) ¿Cómo es su altura comparada con la mayoría de los niños de sus edad?



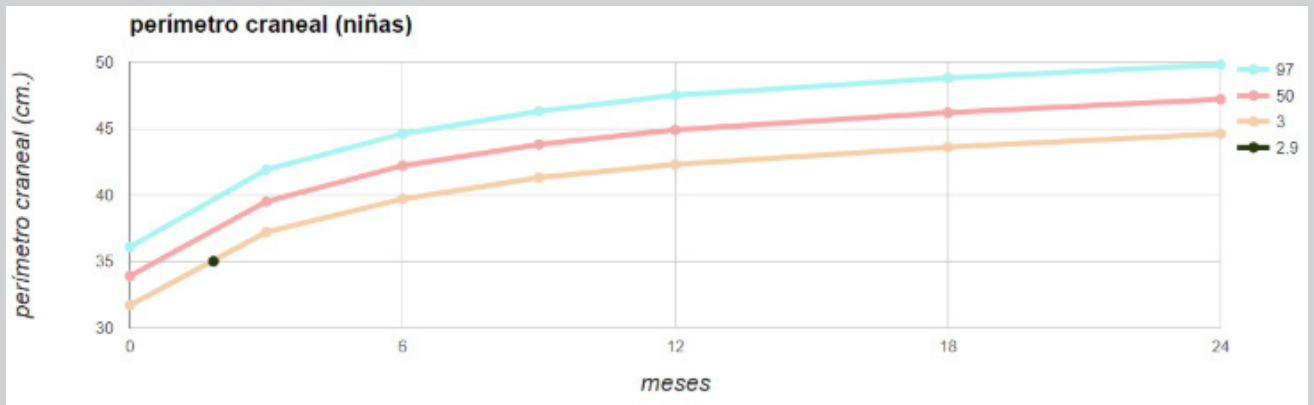
2) ¿Cómo es su peso comparado con la mayoría de los niños de sus edad?



3) ¿Cómo se encuentra su índice de masa corporal comparado con la mayoría de los niños de sus edad?



4) ¿Cómo es su perímetro craneal comparado con la mayoría de los niños de sus edad?



Te sugerimos que pruebes la calculadora de niveles de crecimiento en la siguiente página:

<https://www.percentilesinfantiles.es/>

» » » » »
Continuamos con la lectura del apunte

2. 2. Medidas de centralización

Son aquellos valores que permiten representar un valor de referencia central en la muestra. No ocupan cualquier posición sino la central.

Una medida de tendencia central es un estadístico representativo de una magnitud general observada en la muestra. Entre ellas, las medidas de tendencia central como la media muestral, la moda y la mediana, ayudan a conocer de forma aproximada el comportamiento de una distribución estadística.

• **Media muestral o promedio muestral:** Dados un conjunto de n datos x_1, x_2, \dots, x_n se llama media muestral a la suma de todos los datos dividido el total de observaciones.

La calculamos a través de la siguiente fórmula:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

Recordar! que el símbolo Σ lo que conocemos como sumatoria e indica que debemos sumar todos los valores de la variable desde el primero ($i = 1$) hasta el último ($i = n$).

x_i indica el i -ésimo elemento de la muestra (el que se encuentra en la posición i)

Ejemplo 1: Supongamos que queremos calcular el promedio de calificaciones en diferentes materias de Julia. Las notas que obtuvo son las siguientes: Biología 9 (nueve), Matemática 8 (ocho), Construcción de la Ciudadanía 10 (diez), Inglés 7 (siete), Educación Física 7 (siete), Ciencia Social 10 (diez).

$$\text{Media} = \bar{x} = \frac{9+8+10+7+7+10}{6} = 8.5$$

Podemos tomar al 8.5 como el nivel de desempeño de Julia, como un valor que representa el desempeño que tuvo en todas las asignaturas.

Ejemplo 2: Los siguientes datos representan la estatura en centímetros (cm) de los jugadores de basquet de un equipo:

207; 206; 203; 204; 202; 193; 196; 199.

$$\text{Media} = \bar{x} = \frac{207 + 206 + 203 + 204 + 202 + 193 + 196 + 199}{8} = 201.25$$

Para calcular la media de un conjunto de intervalos utilizamos la siguiente fórmula:

$$\bar{x} = \frac{\sum_{i=1}^n f_i \cdot x_m}{n}$$

f_i = Frecuencia absoluta de datos en el intervalo i

x_m = Punto medio del intervalo

n = Cantidad de datos de la muestra

Como nos indica la fórmula anterior, para calcular la media en un conjunto de intervalos, se calcula el punto medio de cada intervalo y se multiplica por la frecuencia absoluta observada.

Ejemplo: La siguiente tabla muestra a la variable kilómetros recorridos en entrenamientos de 70 ciclistas.

Km recorridos por	Punto medio x_m	f_a (frecuencia absoluta)	F_a (frecuencia acumulada)	$x_m \cdot f_a$
ciclistas				
[5 , 9]	7	38	38	7 . 38 = 266
[10 , 14]	12	20	58	12 . 20 = 240
[15 , 19]	17	12	70	17.12 = 204

$$\bar{x} = \frac{266 + 240 + 204}{70} = 10,14 \text{ km}$$



La media muestral o promedio muestral en los km recorridos en los entrenamientos es de 10,74 km.

• **Mediana:** Es el valor de la variable que se encuentra justo en la mitad, en otras palabras es el valor que solo es superada por la mitad de las observaciones y supera a la otra mitad. Es equivalente al percentil 50 (C_{50}). La simbolizamos como \tilde{x}

Forma de calcularla si la variable es cuantitativa discreta:

1. Si la cantidad de datos es impar, se ordenan de menor a mayor y se elige justo valor central, por ejemplo, si tenemos los siguientes datos:

9; 12; 11; 7; 8; 17; 15; 8; 13

Los ordenamos de izquierda a derecha (de menor a mayor)

7; 8; 8; 9; 11; 12; 13; 15; 17



La mediana es el valor central: $\tilde{x} = 11$, notemos que a la derecha y a la izquierda de 11 tenemos la misma cantidad de datos.

2. Si la cantidad de datos es par, se eligen los dos del centro y se calcula el promedio entre ellos.

Ejemplo: Supongamos que tenemos el siguiente conjunto de datos:

40; 23; 33; 43; 12; 32; 20; 22

Procedemos de la misma forma los ordenamos de menor a mayor:

12; 20; 22; 23; 32; 33; 40; 43

Seleccionamos los dos centrales,

12; 20; 22; 23; 32; 33; 40; 43

Notemos que a la derecha y a la izquierda de los mismos tenemos la misma cantidad de datos, luego calculamos la media entre ellos,

$$\bar{x} = \frac{23+32}{2} = 27,5$$

De esta forma la mediana es: $\tilde{x} = 27,5$

En el caso de tener datos agrupados para calcular la mediana lo primero que hacemos es ubicar el intervalo que contiene a la misma, para ello utilizamos:

$\frac{n}{2}$ en el ejemplo del problema de los ciclistas este valor es $\frac{70}{2}$, es decir que la mediana se encuentra en el primer intervalo ya que nos da 35. Luego utilizamos la siguiente fórmula:

$$\tilde{x} = L + \frac{(\frac{n}{2} - f_{ac}) \cdot I}{f_a}$$

L = límite inferior del intervalo

f_{ac} = frecuencia acumulada del intervalo anterior (si la mediana está en el primer intervalo es cero)

f_a = frecuencia acumulada del intervalo que contiene a la mediana

I = longitud del intervalo que contiene a la mediana

Por lo tanto, en el ejemplo de los ciclistas la mediana es igual a 8,68 km

$$\tilde{x} = 5 + \frac{(35-0) \cdot 4}{38} = 8,68$$

• **Moda:** La moda es el valor que más se repite en una muestra, es decir, el valor con mayor frecuencia. Puede ocurrir que haya más de una moda, por ejemplo una distribución con dos modas la denominamos bimodal, y con varias modas multimodal. Si los datos tienen todos la misma frecuencia decimos que no hay moda.

¿Puedes relacionar este concepto con lo que pensabas que era la moda?

Ejemplo 1: En el ejemplo de los ciclistas el intervalo modal es el correspondiente al punto medio del intervalo de mayor frecuencia, en este caso sería el intervalo que va desde 5 a 9 km, y la moda es:

$$\text{Moda} = \frac{5+9}{2} = 7\text{km}$$

Ejemplo 2: Dados los siguientes datos:

12; 22; 23; 16; 17; 22; 40; 32; 12; 22

Podemos observar que el dato que más se repite es 22, de esta forma. La moda es 22.

Como toda estadístico las medidas de centralización son estimadores puntuales y para utilizarlas y que sean representativas de la muestra se deberían tener en cuenta ciertos aspectos de la muestra. Observemos el ejemplo siguiente:

Supongamos que se obtiene una muestra de los sueldos de 11 personas de una empresa:

\$32400; \$55000; \$52000; \$47000; \$31460; \$55000; \$37200;
\$49100; \$55000; \$230000; \$26000

Calculamos las medidas de centralización:

Media

$$\bar{x} = \frac{32400+55000+52000+47000+31146+55000+37200+49100+55000+230000+26000}{11} = 60895$$

Mediana

Ordenamos los datos de menor a mayor

\$26000; \$31146; \$32400; \$37200; \$47000; \$49100; \$52000;
\$55000; \$55000; \$55000; \$230000

El dato correspondiente a la mediana es \$49100 $\tilde{x} = \$49100$

Moda

La moda es \$55000, como podemos observar es el dato que más se repite.

El objetivo de las medidas de tendencia central en este ejemplo sería obtener un valor que represente lo que ganan la mayoría de las personas que trabajan en la empresa. Observemos que mientras la media o promedio muestral da \$60895, la mediana da \$49100 y la moda da \$55000.

La media nos estaría informando que el sueldo promedio de los empleados de la empresa es 60895, y como vemos en la muestra casi la mitad de los empleados ganan mucho menos que ese valor, lo que ocurre es que la media o promedio se vio afectada por un dato atípico (en inglés, outlier), estos son datos que se encuentran muy distantes del resto y la media es muy sensible a este tipo de datos, en la muestra este dato es \$230000. En cambio, la moda no se vio tan afectada, y la medida más representativa en este ejemplo sería la mediana, decimos que este estadístico es una medida “robusta”, es una medida que no es sensible a estos datos.



ACTIVIDAD 7

Obligatoria (El punto 3 es opcional, no hace falta que lo entreguen)

1) Los siguientes datos corresponden a las notas obtenidas por 24 estudiantes de una universidad:

43; 56; 72; 87; 46; 57; 75; 91; 52; 68; 81; 93; 50; 57; 83; 97; 53; 57; 68; 80; 65; 77; 59; 79

a. Completar las siguiente tabla:

Intervalos de resultados (Notas)	Punto medio	f_a (frecuencia absoluta)	F_a (frecuencia acumulada)
40 - 54			
55 - 69			
70 - 84			
85 - 99			

b. Hallar la media, la mediana y la moda.

2) La media de los pesos de cinco hombres es de 76 kg. Los pesos de cuatro de ellos son: 72 Kg, 74 kg, 75 kg y 81 kg. ¿Cuál es el peso del quinto hombre?

3) Desafío (opcional). Lee atentamente y trata de responder la pregunta del problema: “El peso total de 6 jugadores de rugby es de 540 Kg y el peso medio de 14 bailarines es de 40 kg. ¿Cuál es el peso medio de un grupo de 20 personas formado por los 6 jugadores de rugby y los 14 bailarines”



3. Medidas de Dispersión

Otras medidas muy importantes en estadística son las medidas de dispersión, en ocasiones nos interesa observar cómo están distribuidos los datos de la muestra, hay casos en que la variabilidad es alta y otras en que es baja, en este último caso los datos están más agrupados con respecto a los valores de centralización y decimos que la muestra es homogénea. Las medidas de dispersión más utilizadas son el rango, la varianza y la desviación estándar ó desviación típica y la desviación media. La medida más simple de variabilidad en una muestra es el rango, el cual es la diferencia entre mayor valor de la muestra y el menor, el punto débil de esta medida es que solo depende de dos valores.

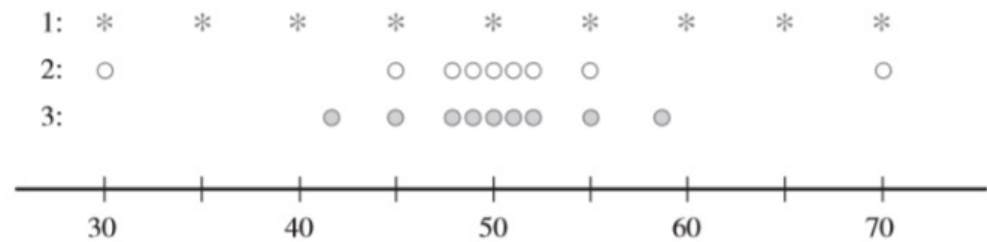


Figura 16. Muestras con igual medidas de centralización y diferente dispersión

3. 1. Desviación estándar o desviación típica

La desviación estándar da una visión detallada de la forma en que los datos están dispersos alrededor de la media como centro de distribución. Su principal uso es la comparación de dos conjuntos de datos, **en determinados trabajos de investigación se busca que la dispersión sea menor**. Está medida es la que más se utiliza para hallar la variabilidad de los datos ya que se encuentra en las mismas unidades que la variable de estudio, se la simboliza con la letra griega σ , su fórmula es:

$$\sigma = \sqrt{\frac{\sum(x - \bar{x})^2}{n}}$$

Ejemplo: Hallar la media y la desviación típica del conjunto de números: 3, 4, 7, 9, 12

Primero calculamos la media:

$$\bar{x} = \frac{3+4+7+9+12}{5} = 7$$

Realizamos una tabla:

x	$(x - \bar{x})$	$(x - \bar{x})^2$
3	-4	16
4	-3	9
7	0	0
9	2	4
12	5	25
		$\sum (x - \bar{x})^2 = 54$

De esta forma obtenemos la desviación estándar:

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n}} = \sqrt{\frac{54}{5}} = 3,2863$$

Ejemplo 2: Hallar la media y la desviación estándar de la siguiente distribución de frecuencias absolutas:

x	0	2	3	6	8	10
f_a (frecuencia absoluta)	2	2	1	1	3	1

Calculamos la media:

$$\bar{x} = \frac{\sum f_a \cdot x_i}{n} = \frac{2x0 + 2x2 + 1x3 + 1x6 + 3x8 + 1x10}{10} = 4,7$$

Realizamos la siguiente tabla:

x	f_a	$(x - \bar{x})$	$(x - \bar{x})^2$	$f_a \cdot (x - \bar{x})^2$
0	2	-4,7	22,09	44,18
2	2	-2,7	7,29	14,58
3	1	-1,7	2,89	2,89
6	1	1,3	1,69	1,69
8	3	3,3	10,89	32,67
10	1	5,3	28,09	28,09
	$n = 10$			$\sum f_a \cdot (x - \bar{x})^2 = 124,1$

Para la distribución de frecuencias utilizamos la fórmula siguiente, notemos que en el numerador del radicando aparece la sumatoria de la última columna de la tabla, de esta forma obtenemos la desviación estándar:

$$\sigma = \sqrt{\frac{\sum f_a \cdot (x - \bar{x})^2}{n}} = \sqrt{\frac{124,1}{10}} = 3,2863$$

**ACTIVIDAD 8****Obligatoria**

- 1) Hallar la media y la desviación estándar de los siguientes conjuntos de número:
- 23, 28, 12, 33, 56, 44, 19, 32
 - 108, 106, 112, 113, 120, 117
- 2) Hallar la media y la desviación estándar de la siguiente distribución de frecuencia:

X	1	3	4	7	8
f_a (frecuencia absoluta)	2	4	1	2	3

- 3) Un ingeniero de producción puede fabricar baterías para autos mediante dos métodos. Una vez fabricadas unas cuantas baterías por los dos métodos toma 8 baterías de cada uno y realiza un control de calidad para medir las respectivas vidas medias. La tabla siguiente muestra las vidas medias obtenidas:

<i>Fabricadas por el método A</i>	21	24	22	23	25	23	26	20
<i>Fabricadas por el método B</i>	23	27	20	22	26	20	27	19

Hallar la media y la desviación estándar para cada muestra. ¿Cuál es el método de fabricación que tú recomendarías al ingeniero para utilizar en el futuro? Razona tu respuesta (Recuerda que en la comparación nos quedaremos con la variable que tenga menor desviación).



¿Alguna actividad que no sabes cómo resolverla?
Los espera el tutor en el Campus Virtual o en el encuentro presencial para acompañarlos y ayudarlos.





¡ Empezamos a estudiar!

Conozcamos un poco más sobre este tema leyendo el apunte de clase que se encuentra a continuación.

Apunte de clase: Probabilidad



1. Probabilidad

Continuemos estudiando este concepto tan interesante.

El término probabilidad se refiere al estudio de azar y la incertidumbre en cualquier situación en la cual varios posibles sucesos pueden ocurrir; la disciplina de la probabilidad proporciona métodos de cuantificar las oportunidades y probabilidades asociadas con varios sucesos.

1. 1. Los primeros pasos de la probabilidad

UN POCO DE HISTORIA

- La prehistoria de la probabilidad data 300 años a.C en la India (Jainies). El impulso surgió por el interés en los juegos de azar.
- Esto finaliza con el libro de Girolamo Cardano "Libro de los juegos de dados" (1663).
- El estudio de la probabilidad como lo concebimos hoy empezó en el siglo XVII asociado a los juegos de azar. Con la correspondencia entre Blaise Pascal y Pierre de Fermat en el año 1654.

1. 2. Fenómenos y Experimentos

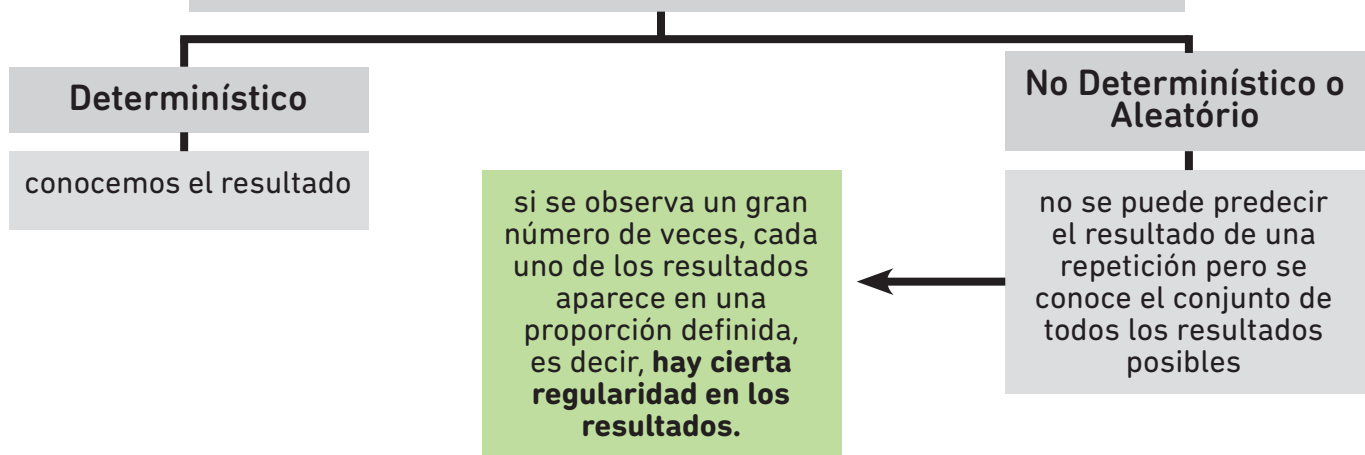
Distinguiremos dos tipos de episodios, los fenómenos y los experimentos:

- **Experimento:** Es toda acción o procedimiento repetible indefinidas veces en condiciones similares mediante el cual se generan resultados. Ejemplo: Se tiene una aula con alumnos y se selecciona uno al azar y se observa una característica en ese alumno.

- **Fenómeno:** Su ocurrencia no requiere de la intervención humana y puede ser observado sólo si ocurre. Ejemplo: En un puerto observo la cantidad de buques que arriban por día durante un mes.

Veamos los tipos de experimentos o fenómenos que hay:

TIPOS DE EXPERIMENTOS O FENÓMENOS



Particularmente nos referiremos a partir de ahora a experimentos o fenómenos aleatorios, es decir, aquellos cuyo resultado está sujeto a incertidumbre.

Muy posiblemente pensaron alguna vez en los juegos de azar: Quini 6, loto, quiniela, ruleta. Los llamados “juegos de azar” no responden a reglas generales, son aleatorios, no se puede predecir el resultado pero si se conocen todos los resultados posibles que pueden tener, pero no sería “negocio” apostar a todas las posibilidades, seguramente invertiríamos mucho más dinero que el que ganaremos.

Como vimos en un poco de historia, hace muchos años importantes matemáticos comenzaron a desafiarse presentándose mutuamente situaciones en las cuales se pretendían analizar todas las “probabilidades” de ganar que tenían. Fue el punto donde los matemáticos comenzaron a preocuparse no solo por los resultados exactos sino también por la resolución de problemas en los cuales interviene el azar.

1. 3. Espacio muestral y sucesos o eventos

Espacio muestral: Realizado un experimento aleatorio, el espacio muestral es el conjunto de todos los resultados posibles, generalmente lo simbolizamos con la letra S.

Ejemplos:

Experimento 1: *Arrojo una moneda 2 veces e indico que salió*

Espacio muestral del experimento 1: $S_1 = \{(c,c) (c,s) (s,c) (s,s)\}$

Experimento 2: *Arrojar una moneda dos veces y registrar el número de caras.*

Espacio muestral del experimento 2: $S_2 = \{0,1,2\}$

Experimento 3: *En una central telefónica, número de llamadas que entran en un período de tiempo.*

Espacio muestral del experimento 3: $S_3 = \{0,1,2,3,\dots\}$ En este caso el espacio muestral no tiene número final y coincide con el conjunto de los números naturales más el cero.

Experimento 4: *El tiempo que tarda cierta reacción química.*
Espacio muestral del experimento 4: $S_4 = \{t \geq 0\}$ En este caso el espacio muestral coincide con el conjunto de los números reales positivos.

Suceso o evento aleatorio: Es un subconjunto cualquiera del espacio muestral S .

Ejemplo: Volvamos al ejemplo del experimento 1.

Experimento 1: *Arrojo una moneda 2 veces e indico que salió*

Espacio muestral del experimento 1: $S_1 = \{(c,c) (c,s) (s,c) (s,s)\}$

Defino dos eventos:

Evento A = "Salen dos caras"

Los resultados posibles de A son: $A = \{(c,c)\}$

¿Evento B = "Salió una cara"

Los resultados posibles de B son: $B = \{(c,s) (s,c)\}$

Otros conceptos útiles:

Suceso o evento elemental: Es el que solo tiene un resultado, por ejemplo el evento.

A: "Salen dos caras"

Suceso o evento imposible: Son los que nunca ocurren, por ejemplo el evento.

B: "Salen tres caras al arrojar una moneda dos veces"

Suceso o evento cierto: Son los que siempre ocurren, por ejemplo el evento.

C: "Sale cara o seca al arrojar una moneda"

Y así llegamos a una definición de la probabilidad.

1. 4. Probabilidad de un suceso o evento

La primera definición de probabilidad que aparece en la historia es la referida a la teoría clásica de Laplace:

Dado un experimento o fenómeno aleatorio, con espacio muestral asociado S finito, y un evento A , de este espacio muestral; se llama **probabilidad de que ocurra el suceso A al cociente entre el número de puntos muestrales de A (resultados favorables) y el total de puntos muestrales de S (resultados posibles), siempre que todos sean equiprobables.**

$$P(A) = \frac{\text{Número de resultados favorables}}{\text{Número de resultados posibles}}$$

Esta definición tiene ciertas limitaciones ya que solo se aplica a resultados equiprobables, esto significa que cada uno de los posibles resultados tienen que tener la misma posibilidad de salir o suceder como por ejemplo el experimento de arrojar una moneda no cargada y anotar qué sale, sabemos que tanto cara como seca tienen la misma posibilidad de salir, 0,5 pero, ¿cómo llegamos a ese número?. Podemos afirmar que ese número surge de repetir muchas veces el experimento, es donde surge una definición posterior, la definición frecuentista, ya que justamente depende de aumentar la frecuencia de realizar el experimento. Actualmente la definición más general que se utiliza es la axiomática formulada por un matemático ruso llamado Andréi Nicoláyevich Kolmogórov.



Andréi Nicoláyevich Kolmogórov fué un matemático ruso que hizo progresos importantes en los campos de la teoría de la probabilidad y de la topología.

Fecha de nacimiento: 25 de abril de 1903, Tambov, Rusia.

Fecha de muerte: 20 de octubre de 1987, Moscú, Rusia.

Educación: Universidad Estatal de Moscú (1929)

Premios: Premio Estatal de la URSS, Premio Lenin, Premio Wolf en Matemática.

En 1933, Kolmogórov publicó su libro, "Fundamentos de la Teoría de Probabilidad", detallando la Definición Axiomática de la Probabilidad.



Como recién estamos conociendo estos significados utilizaremos la definición clásica para comenzar a calcular algunas probabilidades de experimentos equiprobables.

Ejemplos:

Experimento: Se lanza un dado una vez

Evento A: "Sale un número par"

$$\text{Probabilidad de } A = P(A) = \frac{3 \text{ (casos favorables)}}{6 \text{ (casos posibles)}} = 0,5$$

Evento B: "Sale el número 5"

$$\text{Probabilidad de } B = P(B) = \frac{1 \text{ (casos favorables)}}{6 \text{ (casos posibles)}} = 0,166$$

Evento C: "Sale un número mayor que 4"

$$\text{Probabilidad de } C = P(C) = \frac{2 \text{ (casos favorables)}}{6 \text{ (casos posibles)}} = 0,33$$

Evento D: "Sale un número mayor que 6"

$$\text{Probabilidad de } D = P(D) = \frac{0 \text{ (casos favorables)}}{6 \text{ (casos posibles)}} = 0$$

Notemos que este último es un evento imposible, su probabilidad de ocurrencia es cero.

Evento E: "Sale un número entre 1 y 6"

$$\text{Probabilidad de } E = P(E) = \frac{6 \text{ (casos favorables)}}{6 \text{ (casos posibles)}} = 1$$

En este caso se trata de un suceso o evento cierto, su probabilidad de ocurrencia es 1.

RECORDAR!! LA PROBABILIDAD ES SIEMPRE UN NÚMERO ENTRE 0 (SUCESO IMPOSIBLE) Y 1 (SUCESO CIERTO)

Dados un experimento y un espacio muestral S , el objetivo de la probabilidad es asignar a cada evento A un número $P(A)$, llamado la probabilidad del evento A , el cual dará una medida precisa de la oportunidad de que A ocurra

Dado un evento A asociado a un experimento la probabilidad es: $0 \leq P(A) \leq 1$

En ocasiones la probabilidad viene relacionada con un porcentaje de ocurrencia, por ejemplo, mañana hay 70% de probabilidad de chaparrones, el número que expresa esa probabilidad es 0,7. Dado un valor de probabilidad le asignamos el porcentaje al multiplicarlo por 100.



ACTIVIDAD 9

Obligatoria

1) En los siguientes experimentos describir el espacio muestral:

- Las personas A, B y C corren una carrera, se debe anotar el primero, segundo y tercer puesto. (Sugerencia (A, B, C), indica que A salió primero, B salió segundo y C tercero)
- Se lanzan dos monedas normales y se anotan los resultados posibles (cara y seca).
- Una caja contiene 3 bolas blancas y 2 negras que son exactamente iguales excepto por el color. Se extraen dos bolas y se observan sus colores.
- Se lanzan dos dados normales y se anotan los dos números que salen.

2) Dados los siguientes eventos, indicar en cada caso si se trata de un evento probable, un evento cierto o un evento imposible.

- Mañana esté nublado
- Que un águila aprenda a nadar
- Vivir hasta los 200 años
- Vivir hasta los 76 años
- Que nuestro perro nos haga la comida
- Tardar 5 horas en auto para llegar a Mar del Plata desde CABA
- Encender el microondas y que se corte la luz
- Que tire una piedra y caiga al piso
- Que me tire a una piscina llena de agua y me moje
- Que juegue al loto y lo gane

3) Calcular la probabilidad en los siguientes casos.

- Sacar un una "sota" en una baraja de cartas españolas.
- Sacar un número impar al tirar un dado normal.
- Sacar una carta de espadas en una baraja de cartas españolas.
- Acertar un color negro en una jugada de ruleta
- Sacar un número mayor que 7 en una baraja de cartas españolas.
- Sacar un número mayor que 1 al tirar un dado normal
- Sacar un número menor que 5 al lanzar un dado normal.
- En una urna hay 5 bolas blancas, 8 azules y 3 amarillas; al sacar una sola bola, ¿cuál es la probabilidad de que sea azul?
- Si se lanzan dos dados y se anotan los resultados (por ejemplo (2,5) salió el dos y el cinco), se pide
 - la probabilidad que la suma de los dos dados sea 5
 - la probabilidad que la suma de los dos dados sea 3



1.5. Algunos conceptos importantes

Para facilitar la comprensión de algunos conceptos importantes la probabilidad se “nutre” de la teoría de conjuntos, explicaremos brevemente algunos de ellos:

Cuando decimos que un evento A de un espacio muestral S está asociado a un experimento, podemos interpretar esto en la figura 17, donde claramente se ve que A es un subconjunto de S.

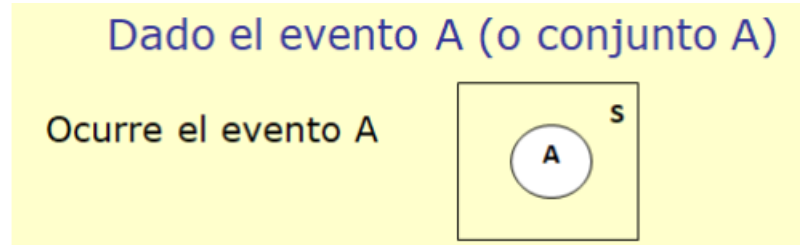


Figura 17. Evento (A) asociado a un espacio muestral (S)

Podemos decir que la probabilidad de que una vez realizado el experimento tengamos resultados es 1, no tendríamos resultados si el experimento no es realizado. De esta forma podemos afirmar siempre que la probabilidad del espacio muestral es 1:

$$P(S) = 1$$

Mirando el gráfico deducimos que si la probabilidad del espacio muestral es igual a 1, la probabilidad del evento A debe ser:

$$0 \leq P(A) \leq 1$$

1.5.1. Conjunto vacío

El conjunto vacío es el que no tiene elementos, lo simbolizamos con el símbolo: \emptyset

La probabilidad del conjunto vacío es cero: $P(\emptyset) = 0$

1.5.2. Intersección de dos eventos

Al hablar de la intersección de dos conjuntos estamos considerando a los elementos que están en ambos conjuntos, es necesario que estén en ambos para que los consideremos. En términos de la probabilidad deben ocurrir ambos eventos. La intersección la simbolizamos con el símbolo:

\cap

La probabilidad de A intersección B la escribimos:

$$P(A \cap B)$$

Relacionaremos este concepto a continuación con el de eventos o sucesos independientes.

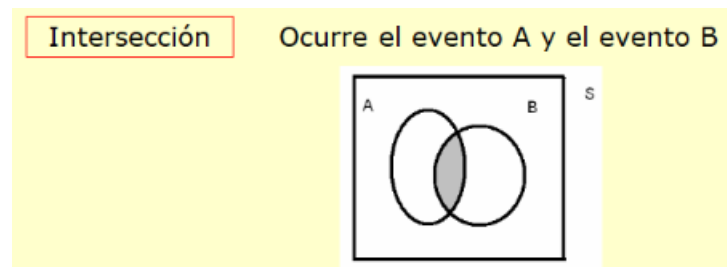


Figura 18. Intersección de dos eventos

1. 5. 3. Unión de dos eventos

Al señalar la unión de dos conjuntos estamos hablando de todos los elementos que están en un conjunto ó en el otro, otra forma de razonarlo es pensar en elementos que están al menos en un conjunto, al decir al menos podría ocurrir que estén en los dos. Lo simbolizamos con el símbolo:

U

En términos de la probabilidad hablamos de ocurrencia, probabilidad de que ocurra al menos uno (ver figura 19).

La probabilidad de A unión B la escribimos: $P(A \cup B)$

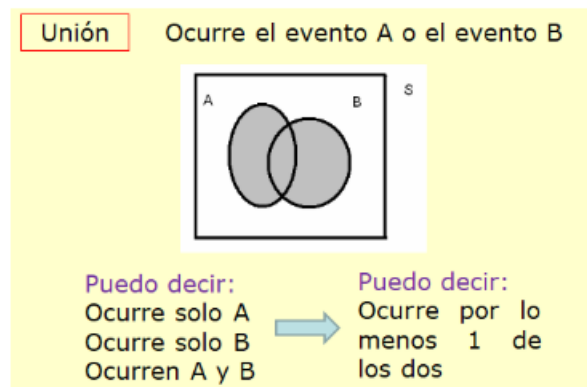


Figura 19. Unión de dos eventos

La probabilidad de la unión dos eventos cualesquiera A y B viene dada por: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Podríamos preguntarnos: ¿Por qué restamos la probabilidad de la intersección?

Si miramos la fórmula del lado derecho (segundo miembro) y consideramos los dos primeros términos $P(A) + P(B)$ la región de la intersección se considera dos veces, una cuando escribimos $P(A)$ y otra cuando escribimos $P(B)$, por lo tanto es necesaria restarla una vez.

Ejemplo: En un pueblo hay dos periódicos locales, “Nuevas voces” y “Opinión”, el 30% de la gente lee “Nuevas Voces” y el 22% lee “Opinión”, mientras que un 7% lee ambos periódicos. ¿Cuál es la probabilidad que una persona elegida al azar lea “Opinión” o “Nuevas Voces”?

Siempre para resolver un problema de probabilidad es útil definir los eventos:

A = La persona lee “Nuevas Voces”

B = La persona lee “Opinión”

Observemos que en el problema nos pregunta la probabilidad de que lea un periódico u otro, esto es equivalente a pensar que la persona lea algún periódico, al menos uno, nos conduce a escribir la probabilidad de la unión:

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

$$P(A \cup B) = 0,3 + 0,22 - 0,07 = 0,45$$

La probabilidad de que una persona lea al menos uno de los dos periódicos es 0,45.

1. 5. 4. Sucesos o eventos mutuamente excluyentes

Dos eventos son mutuamente excluyentes o disjuntos cuando la ocurrencia de uno excluye la posibilidad de ocurrencia del otro, en otras palabras no se pueden dar al mismo tiempo, por lo tanto podemos afirmar que la intersección entre ambos es vacía, y como vimos previamente la probabilidad del conjunto vacío es cero.

DEFINICIÓN: Dos eventos (conjuntos) A y B son disjuntivos si $A \cap B = \phi$. Se dice entonces que los eventos A y B son:

MUTUAMENTE EXCLUYENTES

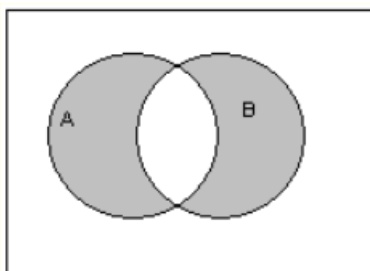


Figura 20. Mutuamente excluyentes

Si A y B son mutuamente excluyentes la probabilidad de la intersección nos queda: $P(A \cap B) = P(\phi) = 0$

Ejemplo 1: Si lanzo una moneda normal y defino los eventos:

A = “Sale cara”

B = “Sale seca”

Sabemos que $P(A) = 0,5$ y $P(B) = 0,5$

Si sale cara no puede salir seca y viceversa, la probabilidad de A intersección B es cero: $P(A \cap B) = P(\emptyset) = 0$

Si quisiera escribir la probabilidad de la unión de A y B, en palabras sería, la probabilidad que salga cara o que salga seca, es más que evidente pensar que alguna de las dos saldrá y por lo tanto esa probabilidad será 1. Utilizando la simbología nos quedará: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Pero cómo sabemos que la probabilidad de la intersección es cero obtenemos:

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

De esta forma llegamos a: $P(A \cup B) = P(A) + P(B)$
Y como la probabilidad que salga cara es 0,5 y que salga seca igual, nos queda: $P(A \cup B) = 0.5 + 0.5 = 1$

De esta forma llegamos a deducir otra propiedad importante:

Si dos eventos A y B son mutuamente excluyentes, la probabilidad de la unión de ambos es:

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

Ejemplo 2: Una automotriz produce dos autos, el color puede ser rojo ó azul.

A = "El auto es rojo"

B = "El auto es azul"

$$P(A \cap B) = P(\emptyset) = 0$$

1. 5. 5. Complemento

En la teoría conjuntista el complemento de un conjunto hace referencia a todo lo que no está en dicho conjunto, es decir todo lo que está por fuera. En términos de la probabilidad diríamos que el complemento de un evento A es equivalente a que no ocurre A, ver figura 21.

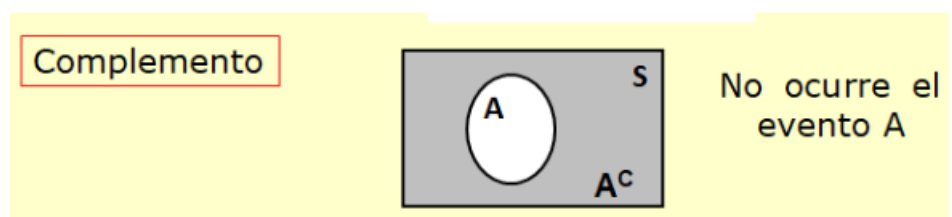


Figura 21. Complemento de un evento

Si observamos la figura 18, podemos deducir las siguientes importantes conclusiones:

Ya sabemos que la probabilidad del espacio muestral es igual a 1,

$$P(S) = 1$$

Pero del gráfico podemos ver claramente que el espacio muestral está formado por A más todo lo que está por fuera de A, es así que:

$$S = A \cup A^c$$

De esta forma podemos escribir la probabilidad del espacio muestral como:

$$P(A \cup A^c) = P(S) = 1$$

Pero como A y su complemento son mutuamente excluyentes, ya que si ocurre A no puede ocurrir su complemento (lo que está por fuera de A)

$$P(A \cup A^c) = P(A) + P(A^c) = 1$$

La expresión que acabamos de deducir es otra importante propiedad:

Dado un evento A asociado a un espacio muestral S siempre se verifica: $P(A) + P(A^c) = 1$

En forma equivalente podemos escribir:

$$P(A) = 1 - P(A^c)$$

Si conocemos la probabilidad de un evento podemos conocer la de su complemento. Si la probabilidad de que mañana llueva es de 0,4 (40%) la probabilidad de que no llueva es de 0,6 (60%), si la probabilidad de que me aumenten el sueldo en el trabajo es de 0,25 (25%) la probabilidad de que no me lo aumenten es de 0,75 (75%).

1. 5. 6. Sucesos o eventos independientes

Cuando la ocurrencia de un evento no afecta la probabilidad de ocurrencia de otro, en dos intervenciones del azar, se dice que ambos son eventos independientes. Por ejemplo que una persona elegida al azar lea Página 12 no afecta la probabilidad de que la próxima elegida al azar lea Clarín.

Vamos con una última definición importante.

Dos eventos A y B son independientes si y sólo si:

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

La probabilidad de la intersección se multiplica, veamos el siguiente ejemplo.

Ejemplo: Se sabe que 30% de las tv smart de cierta empresa requieren servicio estando en garantía, mientras que sólo 10% de las consolas de juegos de la misma empresa necesitan de dicho servicio.

Si alguien adquiere tanto un tv smart como una consola fabricadas por esta empresa, ¿cuál es la probabilidad de que ambas máquinas requieran servicio de garantía?

Definimos dos eventos:

A = “El tv smart comprado necesita servicio técnico estando en garantía”

B = “La consola de juegos necesita servicio técnico estando en garantía”

Entonces: $P(A) = 0,3$ y $P(B) = 0,1$

Por lo tanto la probabilidad de que ambas máquinas requieran servicio técnico estando en garantía la podemos escribir como la intersección, ya que deben ocurrir los dos:

$$P(A \cap B) = P(A) \cdot P(B) = 0.3 \times 0.1 = 0.03$$

Lo que nos dice que la probabilidad de tener que utilizar el servicio técnico para ambos productos es 0,03.

Es importante aclarar que es común confundir este concepto con el de mutuamente excluyente, si dos eventos son mutuamente excluyentes no son independientes, son fuertemente dependientes, ya que la ocurrencia de uno depende de la ocurrencia o no del otro. En cambio cuando son independientes, si uno ocurre nada me dice de que vaya a suceder con el otro.



ACTIVIDAD 10

Obligatoria

1) Dado los experimentos y los eventos asociados, indicar si se trata de eventos mutuamente excluyentes o sucesos independientes:

a. Experimento = Se arroja un dado y se observa que salió

A = “Sale un 2” B = “Sale un 6”

b. Experimento = Se elige al azar dos personas y se les pregunta qué comió.

A = “Tomó una sopa de verduras” B = “Comió un bife con ensalada”

c. Fenómeno = “Se pronostica el estado del tiempo de mañana.

A = “LLueve” B = “Hay sol”

d. Experimento = Se selecciona un frasco de mermelad de una gran producción y se observa si la etiqueta está bien pegada

A = “La etiqueta está bien pegada” B = “La etiqueta está mal pegada”

e. Experimento = “Se plantan semillas de un árbol frutal en dos macetas diferentes, 1 y 2, y se las cuida de igual forma”

A = “Las semillas de la maceta 1 germinan” B = “Las semillas de la maceta 2 germinan”

f. Experimento = Se observa el sexo del próximo nacimiento en una clínica.

A = “Es mujer” B = “Es varón”

2) Dadas las siguientes probabilidades: $P(A) = 0.6$, $P(B) = 0.2$ y $P(A \cap B) = 0.15$. Se pide:

a. Calcular $P(AB)$ sabiendo que A y B son mutuamente excluyentes.

b. Calcular $P(AB)$ sabiendo que A y B son independientes.

c. Calcular $P(AB)$ sabiendo que A y B no son dos eventos cualesquiera. (ni mutuamente excluyentes ni independientes).



- 3)** Se lanza una moneda y un dado, hallar la probabilidad que salga seca y un 3.
- 4)** Se supone que los nacimientos son igualmente probables en cualquiera de los días de la semana. Si se escogen dos personas al azar, ¿cuál es la probabilidad de que ambas hayan nacido un día miércoles?
- 5)** Una bolsa contiene 10 bolas rojas, 5 bolas azules y 7 bolas verdes, hallar:
- La probabilidad de extraer una bola roja.
 - La probabilidad de extraer una bola azul.
 - La probabilidad de extraer una bola azul ó roja
 - La probabilidad de extraer una bola roja ó verde



¿Alguna actividad que no sabes cómo resolverla?
Los espera el tutor en el Campus Virtual o en el encuentro presencial para acompañarlos y ayudarlos.

