

# ESTADÍSTICA DESCRIPTIVA APLICADA A LA ADMINISTRACIÓN Y FINANZAS



© TOSCANO GUERRERO FRANCISCO EDUARDO, RODRÍGUEZ ULCUANGO OLGA MARITZA,  
CEVALLOS RAMOS CARINA DEL ROCÍO, SANTILLÁN ESPINOZA DIEGO IVÁN

# **ESTADÍSTICA DESCRIPTIVA APLICADA A LA ADMINISTRACIÓN Y FINANZAS**

Toscano Guerrero, Francisco Eduardo  
Rodríguez Ulcuango, Olga Maritza  
Cevallos Ramos, Carina Del Rocío  
Santillán Espinoza, Diego Iván



**ESTADÍSTICA DESCRIPTIVA  
APLICADA A LA ADMINISTRACIÓN Y FINANZAS**

© Autores

Toscano Guerrero, Francisco Eduardo  
Docente de la Escuela Superior Politécnica de  
Chimborazo Riobamba, Ecuador

Rodríguez Ulcuango, Olga Maritza  
Docente de la Escuela Superior Politécnica de  
Chimborazo Riobamba, Ecuador

Cevallos Ramos, Carina Del Rocío  
Docente de la Escuela Superior Politécnica de  
Chimborazo Riobamba, Ecuador

Santillán Espinoza, Diego Iván  
Docente de la Escuela Superior Politécnica de  
Chimborazo Riobamba, Ecuador



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

Editado y distribuido por:

**Editorial:** Casa Editora del Polo  
**Sello Editorial:** 978-9942-816  
Manta, Manabí, Ecuador. 2019  
**Teléfono:** (05) 6051775 / 0991871420  
**Web:** www.casedelpo.com  
**ISBN:** 978-9942-621-56-6  
**DOI:** <https://doi.org/10.23857/978-9942-621-56-6>

© Primera edición  
© Septiembre - 2023  
Impreso en Ecuador

**Revisión, Ortografía y Redacción:**  
Lic. Jessica Mero Vélez

**Diseño de Portada:**  
Michael Josué Suárez-Espinar

**Diagramación:**  
Ing. Edwin Alejandro Delgado-Veliz

**Director Editorial:**  
Dra. Tibusay Milene Lamus-García

Todos los libros publicados por la Casa Editora del Polo, son sometidos previamente a un proceso de evaluación realizado por árbitros calificados. Este es un libro digital y físico, destinado únicamente al uso personal y colectivo en trabajos académicos de investigación, docencia y difusión del Conocimiento, donde se debe brindar crédito de manera adecuada a los autores.

© **Reservados todos los derechos.** Queda estrictamente prohibida, sin la autorización expresa de los autores, bajo las sanciones establecidas en las leyes, la reproducción parcial o total de este contenido, por cualquier medio o procedimiento, parcial o total de este contenido, por cualquier medio o procedimiento.

Comité Científico Académico

Dr. Lucio Noriero-Escalante  
Universidad Autónoma de Chapingo, México

Dra. Yorkanda Masó-Dominico  
Instituto Tecnológico de la Construcción, México

Dr. Juan Pedro Machado-Castillo  
Universidad de Granma, Bayamo. M.N. Cuba

Dra. Fanny Miriam Sanabria-Boudri  
Universidad Nacional Enrique Guzmán y Valle, Perú

Dra. Jennifer Quintero-Medina  
Universidad Privada Dr. Rafael Beloso Chacín, Venezuela

Dr. Félix Colina-Ysea  
Universidad SISE. Lima, Perú

Dr. Reinaldo Velasco  
Universidad Bolivariana de Venezuela, Venezuela

Dra. Lenys Piña-Ferrer  
Universidad Rafael Beloso Chacín, Maracaibo, Venezuela

Dr. José Javier Nuñez-Castillo  
Universidad Cooperativa de Colombia, Santa Marta,  
Colombia

## Constancia de Arbitraje

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

## Comité Editorial

Abg. Néstor D. Suárez-Montes  
Casa Editora del Polo (CASEDELPO)

Dra. Juana Cecilia-Ojeda  
Universidad del Zulia, Maracaibo, Venezuela

Dra. Maritza Berenguer-Gouarnaluses  
Universidad Santiago de Cuba, Santiago de Cuba, Cuba

Dr. Víctor Reinaldo Jama-Zambrano  
Universidad Laica Eloy Alfaro de Manabí, Ext. Chone

# CONTENIDO

CAPÍTULO I.....	19
Introducción a la Estadística.....	19
1. Estadística, conceptos y definiciones.....	21
1.1. Estadística descriptiva.....	21
1.2. Estadística inferencial.....	21
1.3. Población.....	21
1.4. Muestra.....	21
1.5. Muestreo.....	22
1.6. Individuo.....	22
1.7. Valor de datos.....	22
1.8. Datos.....	22
1.9. Hipótesis.....	22
1.10. Parámetro.....	23
1.11. Estadístico.....	23
1.12. Variable estadística.....	23
2. Población y muestra.....	23
2.1. Población.....	24
2.2. Muestra.....	24
2.3 Individuo.....	25
3. Variables estadísticas, ejemplos y ejercicios.....	26
3.1 Clasificación.....	27
3.2. Variables cualitativas.....	27
3.2.1. Variable cualitativa ordinal.....	28
3.2.2. Variable cualitativa nominal.....	28
3.3. Variables cuantitativas.....	28
3.3.1 Variable cuantitativa discreta.....	29
3.3.2. Variable cuantitativa continua.....	29
4. La sabiduría de las masas y el buey de Galton .....	30
CAPÍTULO II.....	31
Tablas y gráficas.....	31

5.Tablas de frecuencias.....	33	en intervalos.....	67
5.1. Tablas de frecuencias con datos no agrupados		10.1. Media.....	67
.....	33	10.2. Mediana.....	69
5.2. Tablas de frecuencias con datos agrupados....	37	10.3. Moda.....	71
5.3. Tablas de frecuencias en Excel.....	40	11. Medidas de dispersión.....	73
5.4. Tabla de frecuencias para datos no agrupados		11.1. Medidas de dispersión o variabilidad.....	73
.....	40	11.1.1. Medidas de dispersión.....	75
5.5. Tabla de frecuencias para datos agrupados		12. Rango, amplitud o recorrido.....	76
.....	40	12.1. La fórmula del rango es la siguiente:.....	77
6. Diagrama de barras, gráfico circular y polígono		12.2. Características del rango.....	78
de frecuencias.....	41	13.1. Varianza de la población ( $\sigma^2$ ).....	79
6.1. Diagrama de barras.....	41	13.2. Desviación estándar de la población ( $\sigma$ ).....	80
6.2. Polígono de frecuencias.....	43	13.3. Varianza de la muestra ( $s^2$ ).....	80
6.3. Gráfico circular .....	45	13.4. Desviación estándar de la muestra ( $s$ ).....	80
7. Histogramas, ejemplos y ejercicios.....	47	14. Varianza y desviación estándar de datos agru-	
7.1. Polígono de frecuencias.....	49	pados de variable discreta .....	85
7.2. Ojiva.....	50	15. Coeficiente de variación.....	90
7.3. Diferencia entre el polígono de frecuencias y		15.1. Introducción.....	91
ojiva.....	51	15.2. Definición y fórmula.....	91
		16. Desviación media.....	94
CAPÍTULO III.....	53	17. Desviación media de datos agrupados de va-	
Medidas de tendencia central.....	53	riable discreta.....	96
8. Media, mediana y moda, ejemplos y ejercicios		17. Desviación media para datos agrupados por	
.....	55	intervalos.....	99
8.1. Media (media aritmética).....	55	17.2. Fórmulas de la desviación media para datos	
8.2. Mediana .....	56	agrupados.....	100
8.3. Moda.....	58		
9. Media, mediana y moda para tablas sin interva-		CAPÍTULO IV.....	105
los.....	60	Técnicas de conteo.....	105
9.1. Media.....	60	18. Permutaciones y combinaciones, ejercicios re-	
9.2. Mediana.....	62	sueltos.....	107
9.3. Moda.....	66	18.1 Permutaciones.....	107
10. Media, mediana y moda para datos agrupados		18.2. Combinaciones.....	108

19. Permutación con repetición, ejercicios resueltos	109
20. Permutación circular	111
21. Variaciones, combinaciones y permutaciones	112
21.1. Definiciones	112
<b>CAPÍTULO V</b>	<b>115</b>
Probabilidades	115
22. Probabilidades	117
23. Experimento aleatorio, espacio muestral, evento y probabilidad	120
23.1. Experimento aleatorio	120
23.2. Espacio muestral (S)	120
23.4. Probabilidad	121
24. Regla del complemento en probabilidades	123
25. Regla de la suma o adición de probabilidades	124
26. Probabilidad condicional	127
27. Regla de la multiplicación o producto de probabilidades	131
28. Diagrama de árbol (probabilidades)	136
29. Teorema de la probabilidad total	142
29.1. Partición	142
29.2. Teorema de la probabilidad total	142
29.3. Solución con el teorema de probabilidad total:	144
30. Teorema de Bayes	146
30.1. Forma simple del teorema de Bayes	147
30.2. Forma extendida del teorema de Bayes	150
<b>CAPÍTULO VI</b>	<b>151</b>
Variables aleatorias y distribuciones de probabili-	

dad	151
31. Variables aleatorias discretas y continuas	153
31.1. Variable aleatoria	153
31.2. Variable aleatoria discreta	158
31.3. Variable aleatoria continua	159
32. Función de probabilidad de una variable aleatoria discreta	161
33. Función de distribución acumulativa de una variable aleatoria discreta	168
34. Media (valor esperado), varianza y desviación estándar de una variable aleatoria discreta	180
34.1. Valor esperado	181
34.2. Varianza	181
34.3. Desviación estándar	182
35. Función de densidad de probabilidad con variable continua	186
35.1. Variable aleatoria continua	186
35.2. Función de densidad de probabilidad	187
35.2. Con áreas	190
35.3. Con integrales	191
36. Función de distribución acumulativa de una variable aleatoria continua	192
36.1. Función de distribución acumulativa	194
37. Media o valor esperado, varianza y desviación estándar de una variable aleatoria continua	198
37.1. Media o valor esperado	198
<b>CAPÍTULO VII</b>	<b>203</b>
Distribución de Bernoulli	203
38.1. Distribución de Bernoulli	207
38.2. Media y varianza de una variable aleatoria de Bernoulli	207

CAPÍTULO VIII.....	211
Distribución binomial.....	211
39. Distribución binomial (Introducción).....	213
39.1. Experimento binomial.....	213
39.2. Función de probabilidad binomial (fórmula) .....	218
39.3. Media, varianza y desviación estándar de la distribución binomial.....	219
40. Media, varianza y desviación estándar de una variable aleatoria binomial.....	227
40.1. Media o valor esperado de una variable alea- toria binomial.....	227
40.2. Varianza de una variable aleatoria binomial .....	228
40.3. Desviación estándar de una variable aleatoria binomial.....	228
40.4. Media o valor esperado.....	229
40.5. La varianza de X.....	229
40.6. La desviación estándar de X.....	230
41. Función de distribución acumulativa de una variable aleatoria binomial.....	230
41.1. Tabla de probabilidades acumuladas de la distribución binomial.....	231
41.2. ¿Cómo usar la tabla?.....	231
42. Distribución binomial en Excel.....	235
43. Función de probabilidad binomial.....	235
44. Función de distribución acumulativa.....	236
45. Distribución de Poisson.....	236
45.1 Distribución de Poisson Introducción.....	237
45.2. Experimento de Poisson.....	238
 CAPÍTULO IX.....	 243
Distribución normal.....	243

46. Distribución normal.....	245
46.1. Características de la distribución normal...247	
46.2. La distribución normal estándar.....	248
46.3. Calculadora de distribución normal en Geo- Gebra.....	252
46.3. Pasos para calcular el área bajo la curva....	252
47. Distribución normal en calculadoras Casio ClassWiz 570 y 991.....	253
47.1. Función de distribución normal en Casio....	254
47.2. Función de distribución normal inversa.....	255
48. Tabla de distribución normal estandarizada (tabla z).....	256
48.1. Tabla t de Student.....	257
Referencias bibliográficas .....	259



## CAPÍTULO I

### INTRODUCCIÓN A LA ESTADÍSTICA

## 1. Estadística, conceptos y definiciones

Estadística es la ciencia que se encarga de **recolectar, describir e interpretar datos** para hacer comparaciones y establecer conclusiones. La estadística tiene dos grandes ramas: estadística descriptiva y estadística inferencial.

Veamos algunos **conceptos y definiciones** de estadística:

### 1.1. Estadística descriptiva

Es la rama de la estadística que se ocupa de la clasificación de la información, representación gráfica y resumen de la información a través de medidas estadísticas, lo que permite realizar un análisis de la información captada.

### 1.2. Estadística inferencial

Es la rama de la estadística que se ocupa de la aplicación de procedimientos que permiten realizar **generalizaciones** hacia toda la población, utilizando la información proporcionada por una **muestra aleatoria** extraída de ella.

### 1.3. Población

Es el **conjunto de todos los elementos cuyas propiedades se van a estudiar**. También es llamada **universo**.

### 1.4. Muestra

Es un **subconjunto de la población**. En muchas

ocasiones, es importante trabajar con una **muestra representativa de la población**, para ello, debemos trabajar con criterios y técnicas de muestreo. Una muestra representativa debe reflejar las características de la población.

### 1.5. Muestreo

Es el **proceso de selección de una muestra** a partir de la población estadística que se desea estudiar.

### 1.6. Individuo

Es **cada uno de los elementos que componen la población**. También se le conoce como unidad estadística.

Si quieres revisar algunos ejercicios de muestra y población, dale un vistazo a este artículo.

### 1.7. Valor de datos

Es el valor de la variable asociado con un elemento de una población o muestra. Este valor puede ser un número, una palabra o un símbolo.

### 1.8. Datos

Es cada uno de los valores recolectados de la variable que se han obtenido al realizar un estudio estadístico. Por ejemplo, si le preguntamos a 7 personas cuál es su bebida preferida, **obtenemos 7 datos**: Sprite, Coca Cola, Pepsi, Coca Cola, agua, Gatorade, Coca Cola.

### 1.9. Hipótesis

Es una declaración explícita de predicción. Describe en términos concretos lo que se espera que pase o suceda en los resultados.

### 1.10. Parámetro

Es una medida estadística o indicador que resume la información proporcionada por todas las unidades elementales que forman una población. Es un valor único o constante que usualmente es desconocido.

### 1.11. Estadístico

Es una medida estadística o indicador que permite resumir la información de las observaciones proporcionadas por los elementos que conforman una muestra. Puede tomar un valor diferente de una muestra a otra muestra, por lo tanto, es una variable.

### 1.12. Variable estadística

Es **cada una de las características o cualidades que poseen individuos de una población**.

## 2. Población y muestra

**Población** es el conjunto de todos los elementos cuyas propiedades se van a estudiar; mientras que la **muestra** es un subconjunto de casos o individuos de la población.

Veamos a detalle estos conceptos, y también, algunos ejemplos y ejercicios.

## 2.1. Población

Es el conjunto de todos los elementos cuyas propiedades se van a estudiar. También es llamada universo.

Una población puede ser finita o infinita:

- Población finita: es aquella cuya cantidad de elementos es posible de determinar. Ejemplo: conjunto de librerías de la ciudad de Lima.

- Población infinita: es aquella cuya cantidad de elementos es imposible de determinar. Ejemplo: conjunto de lápices fabricados en un proceso continuo.

## 2.2. Muestra

Es un subconjunto de la población. En muchas ocasiones, es importante trabajar con una muestra representativa de la población, para ello, debemos trabajar con criterios y técnicas de muestreo. Una muestra representativa debe reflejar las características de la población.

En la práctica, para estudiar una población grande, debemos tomar una muestra. Por ejemplo, si queremos saber cuál es el candidato preferido para las próximas elecciones presidenciales de Colombia, tomaría mucho tiempo preguntarles a todos los electores por su candidato preferido, además, sería muy caro contratar tantos encuestadores, digitadores y estadísticos. Por ello, es mejor, analizar una muestra de electores, aplicar una encuesta, y a partir de allí sacar conclusiones de la

población.

## 2.3 Individuo

Es cada uno de los elementos que componen la población. También se le conoce como unidad estadística.

### Ejemplo 1

Para estudiar cuál es el candidato presidencial por el cual votarán los peruanos en las próximas elecciones, se toma una muestra de 3500 personas de todo el país. La pregunta es la siguiente, ¿por quién votará en las próximas elecciones presidenciales? Determine la población, muestra e individuos.

- En este caso, la población sería la población electoral del país, es decir, peruanos con derecho a voto.

- La muestra sería el conjunto de 3500 peruanos que forman parte de la población.

- Un individuo sería cada uno de los peruanos con derecho a voto.

### Ejemplo 2

Un estudiante de estadística quiere conocer si los profesores de su universidad, UNAM, prefieren dictar clases con ropa formal o con ropa informal. Para ello, realiza una encuesta a 120 profesores de la UNAM elegidos de forma aleatoria. Identifique la población, muestra e individuos.

- Población: conjunto de todos los profesores de la UNAM.
- Muestra: 120 profesores de la UNAM.
- Individuo: cada uno de los profesores de la UNAM.

### Ejemplo 3

Un profesor desea realizar un análisis estadístico de las notas del examen final de matemáticas de sus alumnos de último año. Por ello, coloca todas las notas obtenidas en Excel y usa las funciones y herramientas estadísticas. La información obtenida, ¿pertenece a la muestra o a la población?

– En este caso, la población, son todos los alumnos de último año. Se estudiarán sus notas, pero todas las notas obtenidas. No se ha realizado ningún muestreo, por ello, la información obtenida, pertenece a la población.

### 3. Variables estadísticas, ejemplos y ejercicios

Una variable estadística es cada una de las características o cualidades que poseen individuos de una población. Aquí vamos a analizar la clasificación de las variables estadísticas y veremos muchos ejemplos y ejercicios resueltos en los videos que hemos preparado.

Ejemplos de variables estadísticas:

- Color de las bebidas gaseosas: rojo, amarillo, negro, naranja.

- Contenido de las bebidas gaseosas: 0.5 litros, 1 litro, 1.5 litros, 2.5 litros.

### 3.1 Clasificación

Las variables estadísticas se clasifican de la siguiente manera:

Ilustración 1 Tipos de variables estadísticas



Fuente: Matemóvil

### 3.2. Variables cualitativas

Son aquellas que expresan características o cualidades, y no pueden ser medidas con números. Pueden ser ordinales o nominales.

Ejemplos de variables cualitativas:

- El color de los ojos de tus amigos.
- El estado civil de una persona.

### 3.2.1. Variable cualitativa ordinal

Es aquella que presenta valores **no numéricos**, pero existe un orden.

Ejemplos:

- *Las medallas conseguidas en una competencia. Los valores serían: oro, plata, bronce.*
- *Grado de satisfacción laboral en una compañía. Los valores serían: muy satisfecho, satisfecho, regular, insatisfecho, muy insatisfecho (mañana mismo renuncio).*

### 3.2.2. Variable cualitativa nominal

Es aquella que presenta valores **no numéricos**, y no existe un orden.

Ejemplos:

- *El estado civil. Los valores serían: soltero, casado, divorciado, viudo.*
- *El lugar de nacimiento de tus amigos. Los valores serían: Lima, Santiago, Buenos Aires, Zagreb, entre otras ciudades.*

### 3.3. Variables cuantitativas

Son aquellas que se expresan mediante un número, por lo tanto, se puede realizar operaciones aritméticas con ellas. Puede ser discretas o continuas.

Ejemplos de variables cuantitativas:

- *Peso de una bolsa de café.*
- *El número de hijos en una familia.*

### 3.3.1 Variable cuantitativa discreta

Es aquella que puede asumir un número contable de valores.

Ejemplos:

- *El número de hijos en las familias. Puede ser 0, 1, 2, 3, 4, ...*
- *Otro ejemplo sería el número de alumnos en un aula.*

### 3.3.2. Variable cuantitativa continua

Es aquella que puede asumir un número incontable de valores.

Ejemplos:

- *La estatura de los habitantes de una ciudad. Existen infinitos valores posibles, un habitante puede medir 1,784596 metros, otro puede medir 1,589641254125 metros y otro puede medir 1,6457843120 metros. Existen infinitos valores posibles, es decir, un número incontable de valores.*
- *El ancho de las puertas producidas en una fábrica. Existen infinitos valores posibles. Una puerta medir 95,24513 cm, otra puerta medir 96,41 cm, etc.*

#### 4. La sabiduría de las masas y el buey de Galton

Francis Galton fue un estadístico muy famoso, hizo grandes aportes como:

- Inventó el uso de la línea de regresión y fue el primero en explicar la regresión de la media.
- Fue uno de los pioneros en el uso de la famosa distribución normal.
- Introdujo el concepto de la correlación en 1888, que luego fue desarrollado por Pearson y Spearman.



## CAPÍTULO II

### TABLAS Y GRÁFICAS

## 5.Tablas de frecuencias

Una tabla de frecuencias o distribución de frecuencias es una tabla que muestra cómo se distribuyen los datos de acuerdo con sus frecuencias. Elaborar una tabla de frecuencias es muy sencillo y en este artículo te mostraremos cómo hacerlo.

5.1. Tenemos dos tipos de tablas de frecuencias:

- Tablas de frecuencias con datos no agrupados.
- Tablas de frecuencias con datos agrupados.

### 5.1. Tablas de frecuencias con datos no agrupados

Usamos este tipo de tablas cuando tenemos variables cualitativas, o variables cuantitativas con pocos valores.

Esta tabla está compuesta por las siguientes columnas:

- **Valores de la variable:** son los diferentes valores que toma la variable en el estudio.
- **Frecuencia absoluta:** es la cantidad de veces que aparece el valor en el estudio. La sumatoria de las frecuencias absolutas es igual al número de datos.
- **Frecuencia acumulada:** es el acumulado o suma de las frecuencias absolutas, indica cuantos datos se van contando hasta ese momento o cuántos datos se van reportando.
- **Frecuencia relativa:** es la fracción o proporción

de elementos que pertenecen a una clase o categoría. Se calcula dividiendo la frecuencia absoluta entre el número de datos del estudio.

- **Frecuencia relativa acumulada:** es la proporción de datos respecto al total que se han reportado hasta ese momento. Es la suma de las frecuencias relativas, y se puede calcular también dividiendo la frecuencia acumulada entre el número de datos del estudio.

- **Frecuencia porcentual:** es el porcentaje de elementos que pertenecen a una clase o categoría. Se puede calcular rápidamente multiplicando la frecuencia relativa por 100%.

- **Frecuencia porcentual acumulada:** es el porcentaje de datos respecto al total que se han reportado hasta ese momento. Se puede calcular rápidamente multiplicando la frecuencia relativa acumulada por 100%.

**Ejemplo:**

Se le pidió a un grupo de personas que indiquen su color favorito, y se obtuvo los siguientes resultados:

*Tabla 1 Colores Favoritos y resultados*

negro	azul	amarillo	rojo	azul
azul	rojo	negro	amarillo	rojo
rojo	amarillo	amarillo	azul	rojo
negro	azul	rojo	negro	amarillo
Color	Frecuencia absoluta	Frecuencia acumulada	Frecuencia relativa	Frecuencia relativa acumulada
Negro	4	4	0,20	0,20
Azul	5	9	0,25	0,45
Amarillo	5	14	0,25	0,70
Rojo	6	20	0,30	1
Total	20		1	

*Fuente: Matemóvil*

Con los resultados obtenidos, **elaborar una tabla de frecuencias.**

En la primera columna, colocamos los valores de nuestra variable, en la segunda la frecuencia absoluta, luego la frecuencia acumulada, seguida por la frecuencia relativa, y finalmente la frecuencia relativa acumulada. Por ser el primer problema, no haremos uso de las frecuencias porcentuales.

**Ejemplo 2:**

En una tienda de autos, se registra la cantidad de autos Toyota vendidos en cada día del mes de Setiembre.

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

Con los datos obtenidos, **elaborar una tabla de frecuencias.**

**Solución:**

En la primera columna, colocamos los valores de nuestra variable, en la segunda la frecuencia absoluta, luego la frecuencia acumulada, seguida por la frecuencia relativa, y finalmente la frecuencia relativa acumulada. Ahora vamos a agregar la columna de frecuencia porcentual, y frecuencia porcentual acumulada.

*Tabla 2 Frecuencias*

Autos vendidos	Frecuencia absoluta	Frecuencia acumulada	Frecuencia relativa	Frec. relativa acumulada	Frecuencia porcentual	Frec. porcentual acumulada
0	8	8	0,267	0,267	26,7%	26,7%
1	7	15	0,233	0,500	23,3%	50,0%
2	7	22	0,233	0,733	23,3%	73,3%
3	5	27	0,167	0,900	16,7%	90,0%
4	3	30	0,100	1	10,0%	100%
Total	30		1		100%	

*Obtenido de: Matemovil.com*

**5.2. Tablas de frecuencias con datos agrupados**

Usamos las tablas de frecuencias con datos agrupados cuando la variable toma un gran número de valores o es una variable continua. Para ello, se agrupan los diferentes valores en intervalos de igual amplitud, a los cuáles llamamos clases.

Aparecen además algunos parámetros importantes:

- **Límites de clase:** cada clase es un intervalo que va desde el límite inferior, hasta el límite superior.
- **Marca de clase:** es el punto medio de cada intervalo, y representa a la clase para el cálculo de algunos parámetros.

- **Amplitud de clase:** es la diferencia entre el límite superior y el límite inferior.

- Los pasos para elaborar una tabla de frecuencias con datos agrupados son los siguientes:

- Hallar el rango(R):  $R = X_{\max} - X_{\min}$

- Hallar el número de intervalos (K). Si el problema no indica cuántos intervalos usar, se recomienda usar la regla de Sturges:  $K = 1 + 3,322 \cdot \log(n)$  ; siendo n el número de datos.

- Determinar la amplitud de clase (A):  $A = R/K$

- Hallar el límite inferior y superior de cada clase, así como las marcas de clase.

- Colocar los valores hallados en las columnas de la tabla de frecuencias, con el siguiente orden: clases (intervalos), marcas de clase, frecuencia absoluta, frecuencia acumulada, frecuencia relativa, frecuencia relativa acumulada. Además, se puede colocar la frecuencia porcentual y la frecuencia porcentual acumulada.

Recuerda que los intervalos no deben superponerse, es decir, deben ser mutuamente excluyentes.

**Ejemplo 3:**

Las notas de 35 alumnos en el examen final de estadística, calificado del 0 al 10, son las siguientes:

0; 0; 0; 0; 1; 1; 1; 1; 2; 2; 2; 3; 3; 3; 3; 4; 4; 4; 4; 5; 5; 5; 5; 6; 6; 6; 7; 7; 7; 8; 8; 8; 9; 10; 10.

Con los datos obtenidos, elaborar una tabla de frecuencias con 5 intervalos o clases.

**Solución:**

- Hallamos el rango:  $R = X_{\max} - X_{\min} = 10 - 0 = 10$ .

- El número de intervalos (k), me lo da el enunciado del problema:  $k = 5$ .

- Calculamos la amplitud de clase:  $A = R/k = 10/5 = 2$ .

- Ahora hallamos los límites inferiores y superiores de cada clase, y elaboramos la tabla de frecuencias.

*Tabla 3 Frecuencias*

Intervalo	Marca de clase	Frecuencia absoluta	Frecuencia acumulada	Frecuencia relativa	Frec. relativa acumulada
[0 – 2)	1	8	8	0,229	0,229
[2 – 4)	3	7	15	0,200	0,429
[4 – 6)	5	8	23	0,229	0,658
[6 – 8)	7	6	29	0,171	0,829
[8 – 10]	9	6	35	0,171	1
Total		35		1	

*Fuente: Matemóvil*

### 5.3. Tablas de frecuencias en Excel

Veamos cómo construir en Excel tablas de frecuencias para datos no agrupados, y también tablas de frecuencias para datos agrupados. Las tablas de frecuencias deben incluir las siguientes columnas:

- Valores de la variable o intervalos.
- Frecuencia absoluta.
- Frecuencia acumulada.
- Frecuencia relativa.
- Frecuencia relativa acumulada.
- Frecuencia porcentual.
- Frecuencia porcentual acumulada.

### 5.4. Tabla de frecuencias para datos no agrupados

Este tipo de tabla es muy sencillo de construir en Excel, en el siguiente video veremos los pasos para hacerlo. Haremos uso de la función “**contar. Si**”.

La [tabla de frecuencias para datos no agrupados](#), la usamos cuando trabajamos con [variables cualitativas](#) o variables cuantitativas discretas con pocos valores.

### 5.5. Tabla de frecuencias para datos agrupados

Por otro lado, si vamos a construir una tabla de frecuencias con datos agrupados, haremos uso de la función “**frecuencia**” de Excel.

## 6. Diagrama de barras, gráfico circular y polígono de frecuencias

Un **diagrama de barras** es un gráfico usado para representar datos cualitativos o datos cuantitativos discretos tomando en cuenta la frecuencia absoluta, relativa o porcentual. Un **polígono de frecuencias** es un gráfico que se forma uniendo los extremos de las barras (de un gráfico de barras) mediante segmentos. Un **gráfico circular**, es un gráfico que se utiliza para representar frecuencias, porcentajes y proporciones. Veamos como elaborar estos gráficos.

### 6.1. Diagrama de barras

Usado para representar datos cualitativos o datos cuantitativos discretos. Los valores de la variable se colocan en el eje horizontal (x); mientras que en el eje vertical (y), se coloca la frecuencia absoluta, la frecuencia relativa o la frecuencia porcentual. **La altura de cada barra es proporcional a la frecuencia.**

También es llamado gráfico de barras o diagrama de columnas.

Ejemplo:

Se le pidió a un grupo de personas que indiquen su color favorito, y se obtuvo los resultados de la siguiente tabla de frecuencias:

Tabla 4 Diagrama de columnas

Color	Frecuencia absoluta	Frecuencia relativa	Frecuencia porcentual
Negro	4	0.20	20%
Azul	5	0.25	25%
Amarillo	5	0.25	25%
Rojo	6	0.3	30%
<b>Total</b>	<b>20</b>	<b>1</b>	<b>100%</b>

Fuente: Matemóvil

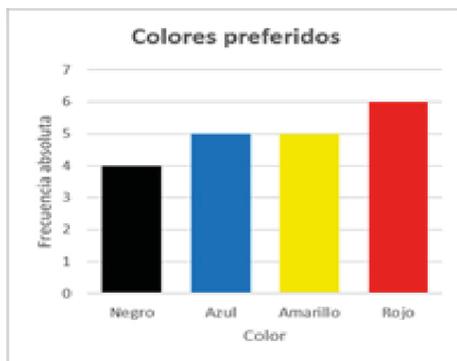
Elaborar una gráfica de barras a partir de dichos resultados.

Solución:

En el eje horizontal (x), colocamos los valores de la variable, es decir, los colores preferidos: negro, azul, amarillo y rojo. En el eje vertical (y), colocaremos la frecuencia. Dado que el problema no indica cuál frecuencia utilizar, absoluta, relativa o porcentual, realizaremos los 3 gráficos.

Veamos primero el diagrama de barras con frecuencia absoluta.

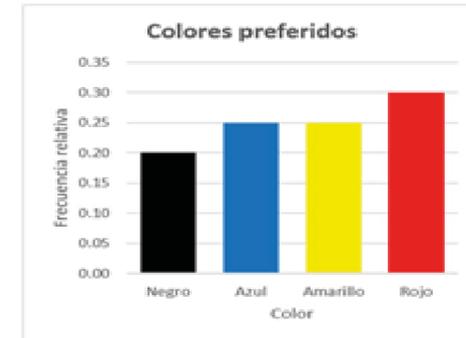
Ilustración 2 Diagrama de barras con frecuencia absoluta



Obtenido de: Matemovil.com

Ahora veamos el diagrama de barras con frecuencia relativa.

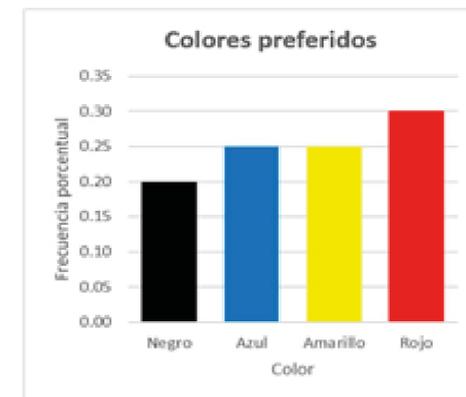
Ilustración 3 Diagrama de barras, frecuencia relativa



Fuente: Matemóvil

Finalmente, viene el diagrama de barras con frecuencia porcentual.

Ilustración 4 Diagrama de barras con frecuencia porcentual



Fuente: Matemóvil

## 6.2. Polígono de frecuencias

Es un gráfico que se forma uniendo los puntos medios de la parte superior de las barras mediante segmentos de recta. El polígono de frecuencias es de mucha utilidad cuando se representa más de una serie en una misma gráfica.

Ejemplo:

Con los datos del problema anterior, elaborar un polígono de frecuencias. Recordemos la tabla de frecuencias inicial:

Tabla 5 Frecuencias inicial

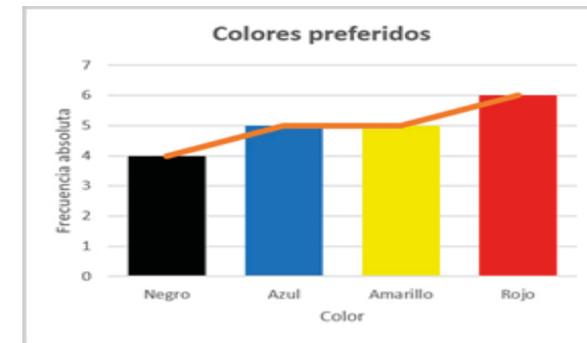
Color	Frecuencia absoluta	Frecuencia relativa	Frecuencia porcentual
Negro	4	0.20	20%
Azul	5	0.25	25%
Amarillo	5	0.25	25%
Rojo	6	0.3	30%
<b>Total</b>	<b>20</b>	<b>1</b>	<b>100%</b>

Fuente: Matemóvil

Solución:

Partimos de la gráfica de barras que realizamos en el problema anterior. Luego, en el punto medio de la parte superior de cada una de las barras, trazamos un segmento hacia el punto medio de la parte superior de la siguiente barra.

Ilustración 5 Diagrama de barras



Fuente: Matemóvil

### 6.3. Gráfico circular

Es un gráfico usado para representar frecuencias, porcentajes y proporciones. Se suele usar con variables cualitativas, ya que con variables cuantitativas puede generar confusiones.

También es llamado, gráfico de pastel, gráfico de torta o gráfica de 360°.

El ángulo central de cada sector es proporcional a la frecuencia. Se calcula de la siguiente manera, teniendo en cuenta la frecuencia a graficar:

$$\alpha = \frac{360^\circ}{n} \cdot \text{frecuencia absoluta}$$

$$\alpha = 360^\circ \cdot \text{frecuencia relativa}$$

$$\alpha = \frac{360^\circ}{100\%} \cdot \text{frecuencia porcentual}$$

Ejemplo:

Con los datos del problema anterior, elaborar un gráfico circular con las frecuencias porcentuales. Recordemos la tabla de frecuencias inicial:

Tabla 6 Frecuencias

Color	Frecuencia absoluta	Frecuencia relativa	Frecuencia porcentual
Negro	4	0.20	20%
Azul	5	0.25	25%
Amarillo	5	0.25	25%
Rojo	6	0.3	30%
<b>Total</b>	<b>20</b>	<b>1</b>	<b>100%</b>

Fuente: Matemóvil

Usaremos la frecuencia porcentual. Calculemos el ángulo central de cada sector:

$$\alpha = \frac{360^\circ}{100\%} \cdot \text{frecuencia porcentual}$$

$$\alpha_{\text{negro}} = \frac{360^\circ}{100\%} \cdot 20\% = 72^\circ$$

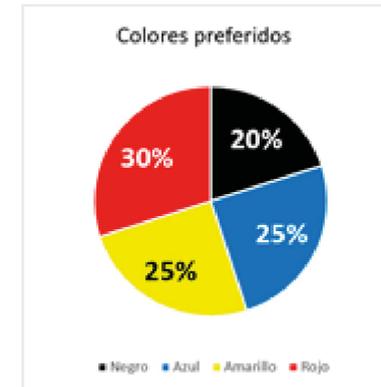
$$\alpha_{\text{azul}} = \frac{360^\circ}{100\%} \cdot 25\% = 90^\circ$$

$$\alpha_{\text{amarillo}} = \frac{360^\circ}{100\%} \cdot 25\% = 90^\circ$$

$$\alpha_{\text{rojo}} = \frac{360^\circ}{100\%} \cdot 30\% = 108^\circ$$

Usando el transportador, medimos cada uno de los ángulos centrales, y dibujamos el gráfico.

Ilustración 6 Porcentaje de colores



Obtenido de: Matemovil.com

## 7. Histogramas, ejemplos y ejercicios

Un histograma es una representación gráfica de una variable en forma de barras, teniendo en cuenta que la superficie de cada barra es proporcional a la frecuencia de los valores representados. Un histograma nos permite ver cómo se distribuyen los valores de la variable en estudio.

Usamos los histogramas cuando analizamos [variables continuas](#), o cuando trabajamos con variables discretas que toman un gran número de valores y son agrupadas en intervalos. Cuando tenemos variables cualitativas, se emplean los [diagramas de barras](#).

¿Cómo construir un histograma?

Partimos de una [tabla de frecuencias con datos agrupados](#), y seguimos los siguientes pasos:

- En el eje horizontal (X), colocamos los límites de clase. Opcionalmente, puedes colocar las marcas de clase.

- En el eje vertical (Y), colocamos las frecuencias. Se suele tomar la frecuencia absoluta, pero también se puede trabajar con la frecuencia relativa o con la frecuencia porcentual.

- Dibujamos las barras de cada clase, teniendo en cuenta que la altura de cada barra es igual a la frecuencia.

Ejemplo:

Se registran los tiempos de las llamadas recibidas en un call-center, y se obtiene la siguiente tabla de frecuencias con datos agrupados. Construir un histograma de frecuencias.

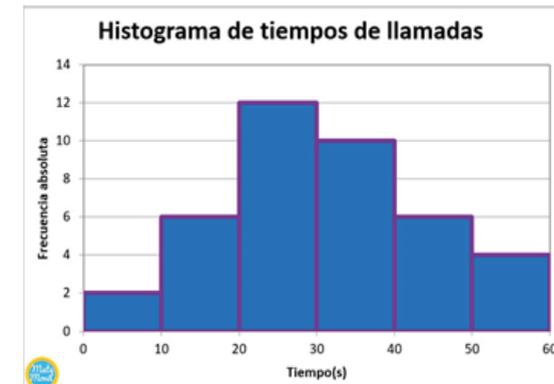
Tabla 7 Datos para un histograma

Tiempo de llamadas	Marcas de clase	Frecuencia absoluta	Frecuencia acumulada	Frecuencia porcentual
[0 - 10)	5	2	2	5%
[10 - 20)	15	6	8	15%
[20 - 30)	25	12	20	30%
[30 - 40)	35	10	30	25%
[40 - 50)	45	6	36	15%
[50 - 60)	55	4	40	10%
<b>Total</b>		<b>40</b>		<b>100%</b>

Fuente: Matemóvil

Solución:

Ilustración 7 Diagrama de barras de un histograma



Fuente: Matemóvil

Recuerda que, si vas a trabajar con una variable cualitativa o variable discreta que asume pocos valores, deberás usar un [diagrama de barras](#) y no un histograma.

### 7.1. Polígono de frecuencias

Es un gráfico que se forma uniendo los puntos medios de la parte superior de las barras mediante segmentos de recta. El polígono de frecuencias es de mucha utilidad cuando se representa más de una serie en una misma gráfica.

Los polígonos de frecuencias se trazan tomando en cuenta las **marcas de clase** de cada barra.

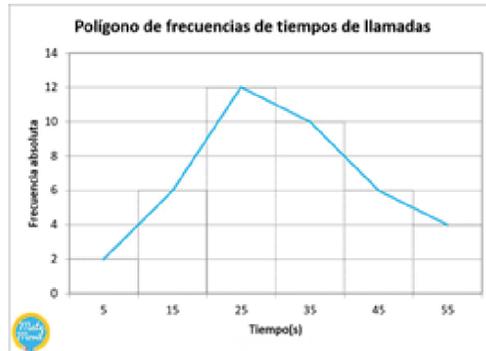
Ejemplo:

A partir del histograma del ejemplo anterior, construir

el polígono de frecuencias.

Solución:

Ilustración 8 Polígono de frecuencias



Obtenido de: Matemovil.com

Ejemplo 3

A partir del histograma del ejemplo anterior, construir la ojiva.

Ilustración 9 Gráfico de Ojiva



Fuente: Matemóvil

## 7.2. Ojiva

La ojiva es una gráfica asociada a la distribución de frecuencias acumuladas. Nos permite ver cuántos datos u observaciones se encuentran por encima o por debajo de determinado valor.

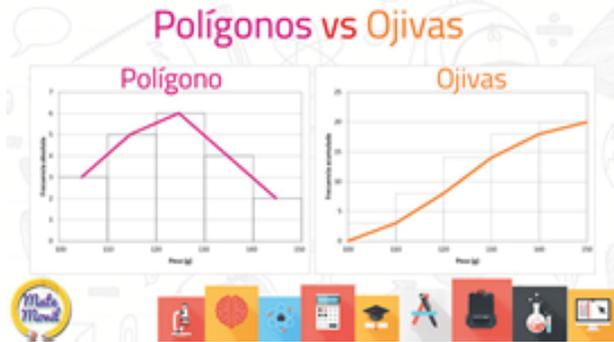
Las ojivas se trazan tomando en cuenta los límites superiores de cada clase o intervalo, es decir, tomando el extremo derecho de la parte superior de cada barra. Dibujar una ojiva es muy similar a dibujar un polígono de frecuencias acumuladas.

Una ojiva también se puede construir con las frecuencias relativas acumuladas o frecuencias porcentuales acumuladas.

## 7.3. Diferencia entre el polígono de frecuencias y ojiva

El polígono de frecuencias parte desde el histograma de frecuencias absolutas, mientras que la ojiva parte del histograma de frecuencias acumuladas. Además, el polígono de frecuencias se forma uniendo los puntos medios de la parte superior de cada barra, mientras que la ojiva se forma uniendo el extremo derecho de la parte superior de cada barra.

Ilustración 10 Polígonos vs Ojivas



Obtenido de: [Matemovil.com](http://Matemovil.com)



## CAPÍTULO III

### MEDIDAS DE TENDENCIA CENTRAL

## 8. Media, mediana y moda, ejemplos y ejercicios

La **media** es el valor que se obtiene al sumar todos los datos y dividir el resultado entre la cantidad de datos. La **mediana** es el valor que ocupa la posición central cuando todos los datos están ordenados en orden creciente o decreciente. La **moda** es valor que más se repite. Veamos cada una de ellas a detalle con ejemplos y ejercicios

Las medidas de tendencia central, como la **media**, **mediana** y **moda**, son medidas que tratan de **ubicar la parte central de un conjunto de datos**.

### 8.1. Media (media aritmética)

La media es el valor que se obtiene al sumar todos los datos y dividir el resultado entre la cantidad de datos.

Su fórmula es la siguiente:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum x_i}{n}$$

Aunque la fórmula parezca complicada, calcular el valor de la media es muy sencillo.

Ejemplo 1:

Calcular la media de los siguientes datos: 11, 6, 7, 7, 4.

$$\bar{x} = \frac{11 + 6 + 7 + 7 + 4}{5} = \frac{35}{5} = 7$$

Ejemplo 2:

Las edades de 8 niños que van a una fiesta son: 2, 2, 3, 5, 7, 7, 9, 10. Hallar la edad media:

$$\bar{x} = \frac{2+2+3+5+7+7+9+10}{8} = \frac{45}{8} = 5,625$$

Ejemplo 3:

En un examen calificado del 0 al 10, 3 personas obtuvieron 5 de nota, 5 personas obtuvieron 4 de nota, y 2 personas obtuvieron 3 de nota. Calcular la nota media:

$$\bar{x} = \frac{5+5+5+4+4+4+4+4+3+3}{10} = \frac{41}{10} = 4,1$$

## 8.2. Mediana

La mediana es el valor que ocupa la posición central cuando todos los datos están ordenados en orden creciente o decreciente.

La mediana se representa con las letras:  $M_e$ .

Ejemplo 1:

Calcular la mediana de los siguientes datos: 11, 6, 7, 7, 4.

**Solución:**

Ordenamos los datos de menor a mayor: 4, 6, 7, 7, 11.

Ahora tomamos el dato que se encuentra al centro: 4, 6, 7, 7, 11.

El valor de la mediana es:  $M_e = 7$ .

• ¿Y si la cantidad de datos es un número par?

En ese caso, **la mediana es la media entre los dos valores centrales.**

Ejemplo 2:

Calcular la mediana de los siguientes datos: 3, 6, 7, 9, 4, 4.

**Solución:**

Primero ordenamos los datos de menor a mayor: 3, 4, 4, 6, 7, 9.

La cantidad de datos es 6, es decir, un número par, así que vamos a ubicar los 2 valores centrales: 3, 4, 4, 6, 7, 9.

Entonces, la moda sería la media entre 4 y 6:

$$M_e = \frac{4+6}{2} = \frac{10}{2} = 5$$

Ejemplo 3:

En un examen calificado del 0 al 10, 3 personas obtuvieron 5 de nota, 5 personas obtuvieron 4 de nota, y 2 personas obtuvieron 3 de nota. Calcular la mediana.

**Solución:**

Primero hacemos una lista de las notas obtenidas: 5, 5, 5, 4, 4, 4, 4, 4, 3, 3.

Ahora ordenamos los datos de menor a mayor: 3, 3,

4, 4, 4, 4, 4, 5, 5, 5.

Como el número de datos es par (10), entonces nos enfocamos en los 2 valores centrales: 3, 3, 4, 4, **4, 4**, 4, 5, 5, 5.

Finalmente, encontramos la media de estos 2 valores centrales:

$$M_e = \frac{4 + 4}{2} = \frac{8}{2} = 4$$

Si al momento de calcular la mediana, ordenas los datos en forma decreciente o descendente, obtendrás el mismo resultado que al hacerlo de forma creciente o ascendente.

### 8.3. Moda

La moda es el valor que más se repite. También podemos decir que la moda es el valor con mayor frecuencia absoluta o el valor que ocurre con más frecuencia.

La moda se representa con las letras:  $M_o$ .

Ejemplo 1:

Calcular la moda de los siguientes datos: 11, 6, 7, 7, 4.

Podemos ver que el valor que más se repite es el 7, ya que tiene una frecuencia absoluta de 2, por lo tanto,  $M_o = 7$ .

Ejemplo 2:

En un examen calificado del 0 al 10, 3 personas obtuvieron 5 de nota, 5 personas obtuvieron 4 de nota, y 2 personas obtuvieron 3 de nota. Calcular la moda.

**Solución:**

Los datos son los siguientes: 5, 5, 5, **4, 4, 4, 4, 4**, 3, 3.

El valor que más se repite es el 4, que aparece 5 veces, por lo tanto,  $M_o = 4$ .

- ¿Y si hay varias modas?

Si en un grupo de datos, dos o más valores tienen la misma frecuencia, y es la frecuencia máxima, entonces la distribución tiene dos o más modas y decimos que es **bimodal (2 modas)**, o **multimodal (varias modas)**.

Ejemplo 3:

Calcular la moda de los siguientes datos: 3, 4, 4, 6, 7, 7, 9, 11.

**Solución:**

Como vemos, hay 2 valores que se repiten 2 veces, el 4 y el 7, por lo tanto, los valores de la moda son  $M_o = 4; 7$ .

- ¿Y si todos los valores tienen la misma frecuencia?

Si todos los valores tienen la misma frecuencia, entonces, **no hay moda**.

Ejemplo 4:

Encontrar la moda de los siguientes datos: 3, 3, 5, 5, 6, 6, 7, 7.

Todos los valores tienen una frecuencia de 2, por lo tanto, **no hay moda**.

### 9. Media, mediana y moda para tablas sin intervalos

La **media** es el valor que se obtiene al sumar todos los datos multiplicados por su frecuencia y dividir el resultado entre la cantidad de datos. La **mediana** es el valor que ocupa la posición central si n es impar, y es el promedio de los dos datos centrales si n es par, cuando todos los datos están ordenados. La **moda** es valor con mayor frecuencia absoluta. Cuando buscamos encontrar la media, mediana y moda para datos agrupados sin intervalos o agrupados puntualmente, necesitamos realizar algunos pasos adicionales y conocer un par de fórmulas. Veamos cómo hacerlo.

#### 9.1. Media

La **media** es el valor que se obtiene al sumar todos los datos multiplicados por su frecuencia y dividir el resultado entre la cantidad de datos. Usamos la siguiente fórmula:

$$\bar{x} = \frac{x_1 \cdot f_1 + x_2 \cdot f_2 + \dots + x_n \cdot f_n}{n} = \frac{\sum_{i=1}^n x_i \cdot f_i}{n}$$

Ejemplo 1:

Calcular la media de la siguiente distribución:

Tabla 8 Distribución

Variable	Frec. Absoluta	Frec. Acumulada
x	f	F
3	9	9
4	11	20
5	10	30
<b>Total</b>	<b>30</b>	

Fuente: Matemóvil

Solución:

Para calcular la media, vamos a agregar una columna adicional, en la que **multiplicaremos el valor de la variable (x) por la frecuencia absoluta (f)**.

Tabla 9 Multiplicación X.F

Variable	Frec. Absoluta	Frec. Acumulada	x.f
x	f	F	
3	9	9	27
4	11	20	44
5	10	30	50
<b>Total</b>	<b>30</b>		<b>121</b>

Fuente: Matemóvil

Recordamos la fórmula:

$$\bar{x} = \frac{x_1 \cdot f_1 + x_2 \cdot f_2 + x_3 \cdot f_3}{n} = \frac{121}{30} = 4,033$$

## 9.2. Mediana

La **mediana** es el valor que ocupa la posición central si n es impar, y es el promedio de los dos datos centrales si n es par, cuando todos los datos están ordenados.

Una manera rápida de encontrar la mediana es encontrar el dato que ocupa la siguiente posición:

$$posición = \frac{n + 1}{2}$$

Ejemplo (n impar):

Calcular la mediana de la siguiente distribución:

Tabla 10 Distribución ejemplo 2

Variable	Frec. Absoluta	Frec. Acumulada
x	f	F
3	9	9
4	11	20
5	13	33
<b>Total</b>	<b>33</b>	

Fuente: Matemóvil

Encontramos la posición del valor que está ubicado al centro:

$$posición = \frac{n + 1}{2} = \frac{33 + 1}{2} = \frac{34}{2} = 17$$

Ahora, buscamos la posición 17 en la columna de frecuencias acumuladas:

Tabla 11 Frecuencias acumuladas

Variable	Frec. Absoluta	Frec. Acumulada
x	f	F
3	9	9
4	11	20
5	13	33
<b>Total</b>	<b>33</b>	

Fuente: Matemóvil

El valor (x) que ocupa dicha posición es 4, por lo tanto,  $M_e = 4$ .

Ejemplo 3 (n par):

Calcular la mediana de la siguiente distribución:

Tabla 12 Calculo de mediana

Variable	Frec. Absoluta	Frec. Acumulada
x	f	F
3	10	10
4	11	21
5	13	34
<b>Total</b>	<b>34</b>	

Obtenido de: Matemovil.com

Encontramos la posición del valor que está ubicado al centro:

$$\text{posición} = \frac{n + 1}{2} = \frac{34 + 1}{2} = \frac{35}{2} = 17,5$$

Como nos ha quedado un valor con decimales, significa que la mediana será la media aritmética del valor que ocupa la posición 17, con el valor que ocupa la posición 18.

Buscamos los valores de posición 17 y 18 en la tabla de frecuencias acumuladas:

Tabla 13 Frecuencias acumuladas

Variable	Frec. Absoluta	Frec. Acumulada
x	f	F
3	10	10
4	11	21
5	13	34
<b>Total</b>	<b>34</b>	

Fuente: Matemóvil

La posición 17, pertenece al valor  $x = 4$ , y la posición 18 también pertenece al mismo valor. Por lo tanto:

$$M_e = \frac{4 + 4}{2} = 4$$

Ejemplo (n par):

Calcular la mediana de la siguiente distribución:

Tabla 14 Cálculo de la mediana

Variable	Frec. Absoluta	Frec. Acumulada
X	f	F
3	16	16
4	13	29
5	29	58
<b>Total</b>	<b>58</b>	

Fuente: Matemóvil

Encontramos la posición del valor que está ubicado al centro:

$$\text{posición} = \frac{n + 1}{2} = \frac{58 + 1}{2} = \frac{59}{2} = 29,5$$

Como nos ha quedado un valor con decimales, significa que la mediana será la media aritmética del valor que ocupa la posición 29, con el valor que ocupa la posición 30. Buscamos los valores de posición 29 y 30 en la tabla de frecuencias acumuladas:

Tabla 15 Frecuencias acumuladas

Variable	Frec. Absoluta	Frec. Acumulada
X	f	F
3	16	16
4	13	29
5	29	58
<b>Total</b>	<b>58</b>	

Fuente: Matemóvil

El valor que ocupa la posición 29, es  $x = 4$ ; mientras

que el valor que ocupa la posición 30, es  $x = 5$ . Por lo tanto, la mediana sería:

$$M_e = \frac{4 + 5}{2} = \frac{9}{2} = 4,5$$

### 9.3. Moda

La moda es el valor con mayor frecuencia absoluta.

Ejemplo:

Encontrar la moda de la siguiente distribución:

Tabla 16 Cálculo de la moda

Variable x	Frec. Absoluta f	Frec. Acumulada F
3	9	9
4	11	20
5	10	30
<b>Total</b>	<b>30</b>	

Fuente: Matemóvil

La moda es el valor con mayor frecuencia absoluta, así que ubicamos el mayor valor de frecuencia absoluta con su correspondiente valor de la variable x:

Tabla 17 Valor de variable x

Variable x	Frec. Absoluta f	Frec. Acumulada F
3	9	9
4	11	20
5	10	30
<b>Total</b>	<b>30</b>	

Fuente: Matemóvil

El valor con mayor frecuencia absoluta es el 4, por lo

tanto:  $M_o = 4$

### 10. Media, mediana y moda para datos agrupados en intervalos

Calcular la media, mediana y moda cuando trabajamos con datos agrupados en intervalos o tablas de frecuencias con intervalos es muy sencillo, y solo se necesitan algunas fórmulas.

#### 10.1. Media

La media se calcula usando la siguiente fórmula:

$$\bar{x} = \frac{x_1 \cdot f_1 + x_2 \cdot f_2 + \dots + x_n \cdot f_n}{n} = \frac{\sum_{i=1}^n x_i \cdot f_i}{n}$$

Ejemplo

Determina la media de la siguiente distribución:

Tabla 18 Determinación de la media

	Intervalos	Marca de clase x	Frecuencia absoluta f	Frecuencia acumulada F
1	[0 - 4)	2	3	3
2	[4 - 8)	6	5	8
3	[8 - 12)	10	6	14
4	[12 - 16)	14	4	18
5	[16 - 20)	18	3	21
	<b>Total</b>		<b>21</b>	

Fuente: Matemóvil

Dado que tenemos 5 intervalos, la media la calculamos

usando la fórmula:

$$\bar{x} = \frac{x_1 \cdot f_1 + x_2 \cdot f_2 + x_3 \cdot f_3 + x_4 \cdot f_4 + x_5 \cdot f_5}{n} = \frac{\sum_{i=1}^5 x_i \cdot f_i}{n}$$

En la tabla, agregamos una columna donde colocaremos todos los valores de x.f :

Tabla 19 Valores de X.F

	Intervalos	Marca de clase x	Frecuencia absoluta f	Frecuencia acumulada F	x.f
1	[0 - 4)	2	3	3	$x_1 \cdot f_1$
2	[4 - 8)	6	5	8	$x_2 \cdot f_2$
3	[8 - 12)	10	6	14	$x_3 \cdot f_3$
4	[12 - 16)	14	4	18	$x_4 \cdot f_4$
5	[16 - 20)	18	3	21	$x_5 \cdot f_5$
	Total		21		$\sum_{i=1}^n x_i \cdot f_i$

Fuente: Matemóvil

Calculamos los valores de x.f :

Tabla 20 Calculo de X.F

	Intervalos	Marca de clase x	Frecuencia absoluta f	Frecuencia acumulada F	x.f
1	[0 - 4)	2	3	3	6
2	[4 - 8)	6	5	8	30
3	[8 - 12)	10	6	14	60
4	[12 - 16)	14	4	18	56
5	[16 - 20)	18	3	21	54
	Total		21		206

Fuente: Matemóvil

Finalmente, calculamos el valor de la media, dividiendo la suma de valores de la columna x.f entre n.

$$\bar{x} = \frac{x_1 \cdot f_1 + x_2 \cdot f_2 + x_3 \cdot f_3 + x_4 \cdot f_4 + x_5 \cdot f_5}{n} = \frac{\sum_{i=1}^5 x_i \cdot f_i}{n} = \frac{206}{21} = 9,810$$

El valor de la media sería 9,810.

## 10.2. Mediana

Para estimar la mediana, hay que seguir 2 pasos:

Encontrar el intervalo en el que se encuentra la mediana usando la fórmula:

$$posición = \frac{n + 1}{2}$$

Usar la fórmula de la mediana:

$$M_e = L_i + \frac{\frac{n}{2} - F_{i-1}}{f_i} \cdot A_i$$

Donde:

- $L_i$ : límite inferior del intervalo en el cual se encuentra la mediana.
- n: número de datos del estudio. Es la sumatoria de las frecuencias absolutas.
- $F_{i-1}$ : frecuencia acumulada del intervalo anterior al que se encuentra la mediana.
- $A_i$ : amplitud del intervalo en el que se encuentra la mediana.
- $f_i$ : frecuencia absoluta del intervalo en el que se

encuentra la mediana.

Ejemplo

Encontrar la mediana de la siguiente distribución:

Tabla 21 Cálculo de la mediana

	Intervalos	Marca de clase x	Frecuencia absoluta f	Frecuencia acumulada F
1	[0 - 4)	2	3	3
2	[4 - 8)	6	5	8
3	[8 - 12)	10	6	14
4	[12 - 16)	14	4	18
5	[16 - 20)	18	3	21
	Total		21	

Fuente: Matemóvil

Para estimar el valor de la mediana, seguimos los 2 pasos.

Primero encontramos el intervalo en el cual se encuentra la mediana usando la fórmula:

$$\text{posición} = \frac{n + 1}{2} = \frac{21 + 1}{2} = \frac{22}{2} = 11$$

Este valor, lo buscamos en la columna de frecuencias acumuladas. Si no aparece, buscamos el valor que sigue. Como vemos, después del 11 sigue el 14, por lo tanto, es el bloque 3.

Tabla 22 Frecuencias acumuladas

	Intervalos	Marca de clase x	Frecuencia absoluta f	Frecuencia acumulada F
1	[0 - 4)	2	3	3
2	[4 - 8)	6	5	8
3	[8 - 12)	10	6	14
4	[12 - 16)	14	4	18
5	[16 - 20)	18	3	21
	Total		21	

Fuente: Matemóvil

Ahora, aplicamos la fórmula de la mediana:

$$M_e = L_i + \frac{\frac{n}{2} - F_{i-1}}{f_i} \cdot A_i = 8 + \frac{\frac{21}{2} - 8}{6} \cdot 4$$

$$M_e = 8 + \frac{10,5 - 8}{6} \cdot 4 = 8 + \frac{2,5}{6} \cdot 4$$

$$M_e = 9,667$$

El valor de la mediana sería:  $M_e = 9,667$ .

### 10.3. Moda

Para estimar la moda, se siguen los siguientes pasos:

Encontrar el intervalo en el cual se encuentra la moda, que es el intervalo con mayor frecuencia absoluta.

Usar la siguiente fórmula para estimar el valor de la moda:

$$M_o = L_i + \frac{f_i - f_{i-1}}{f_i - f_{i-1} + f_i - f_{i+1}} \cdot A_i$$

Donde:

$L_i$ : límite inferior del intervalo en el cual se encuentra la moda.

$f_{i-1}$ : frecuencia absoluta del intervalo anterior en el que se encuentra la moda.

$f_i$ : frecuencia absoluta del intervalo en el que se encuentra la moda.

$f_{i+1}$ : frecuencia absoluta del intervalo siguiente en el que se encuentra la moda.

$A_i$ : amplitud del intervalo en el que se encuentra la moda.

### Ejemplo 3

Encontrar la moda de la siguiente distribución:

Tabla 23 Cálculo de moda

	Intervalos	Marca de clase x	Frecuencia absoluta f	Frecuencia acumulada F
1	[0 - 4)	2	3	3
2	[4 - 8)	6	5	8
3	[8 - 12)	10	6	14
4	[12 - 16)	14	4	18
5	[16 - 20)	18	3	21
	Total		21	

Fuente: Matemóvil

Primero, encontramos el intervalo en el cual se

encuentra la moda, es decir, el intervalo con mayor frecuencia absoluta.

Tabla 24 Intervalo con mayor frecuencia absoluta

	Intervalos	Marca de clase x	Frecuencia absoluta f	Frecuencia acumulada F
1	[0 - 4)	2	3	3
2	[4 - 8)	6	5	8
3	[8 - 12)	10	6	14
4	[12 - 16)	14	4	18
5	[16 - 20)	18	3	21
	Total		21	

Obtenido de: Matemovil.com

Ahora, aplicamos la fórmula para estimar la moda:

$$M_o = L_i + \frac{f_i - f_{i-1}}{f_i - f_{i-1} + f_i - f_{i+1}} \cdot A_i = 8 + \frac{6 - 5}{6 - 5 + 6 - 4} \cdot 4$$

$$M_o = 8 + \frac{1}{1 + 2} \cdot 4 = 8 + \frac{1}{3} \cdot 4$$

$$M_o = 9,333$$

Por lo tanto, el valor de la moda sería:  $M_o = 9,333$ .

## 11. Medidas de dispersión

### 11.1. Medidas de dispersión o variabilidad

Las medidas de dispersión son medidas estadísticas que muestran la variabilidad en la distribución de los datos. Las medidas de tendencia central, como la media, la mediana y la moda, solo describen el centro de los

datos, pero no nos dicen nada acerca de la dispersión (separación) de los datos. Y en ocasiones, es muy importante conocer que tan dispersos o separados se encuentran los datos, y esto se consigue con las medidas de dispersión o variabilidad.

Veamos los siguientes ejemplos.

Ejemplo 1:

Si tu estatura es de 1,80 metros y un guía de turismo te dice que el río que estás a punto de cruzar a pie tiene una profundidad promedio de 1,50 metros, ¿lo cruzarías a pie? Antes de tomar una decisión, sería mejor saber cuál es la profundidad máxima y cuál es la mínima. Si el guía te dice que la profundidad mínima es de 1,40 metros y máxima es de 1,60 metros, seguramente te animarás a cruzar el río a pie.

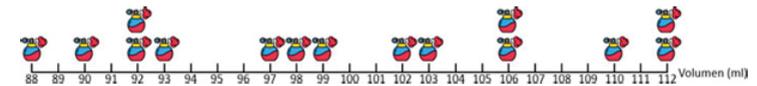
Pero; ¿qué pasaría si el guía te dice que la profundidad mínima es de 0,50 metros y la máxima es de 2,50 metros? ¿te animarías aún a cruzar el río? Probablemente no. En este caso, no es suficiente con conocer el promedio, si no también, el valor máximo y el mínimo.

Ejemplo 2:

Hace un tiempo, iba a comprar 10 000 frascos de perfumes para revenderlos a un mejor precio a mis amigos. El vendedor me dijo que, en promedio, los frascos traían 100 mililitros. Pero yo como yo sabía algo de estadística, decidí realizar un muestreo de 15 frascos,

y estos fueron los resultados que obtuve:

Ilustración 11 Muestreo



Fuente: Matemóvil

En el muestreo, efectivamente la media fue de 100 ml, pero hubo perfumes con 88 ml, con 106 ml y con 112 ml. La dispersión de los datos obtenidos era muy alta, iba a tener muchos problemas y quejas con mis amigos. Lo ideal hubiera sido que todos los perfumes traigan 100 ml, o al menos volúmenes muy cercanos a este valor, sin tanta dispersión o separación entre los valores, es decir, que los valores se encuentren estrechamente agrupados.

Felizmente que eran perfumes, porque si eran medicinas y salían al mercado, algunos pacientes hubieran recibido una dosis menor y otros una sobredosis.

Es por ello que son tan importantes las medidas de dispersión o variabilidad, pues estas nos indican que tan dispersos o separados se encuentran los datos.

### 11.1.1. Medidas de dispersión

Las medidas de dispersión son medidas estadísticas que muestran la variabilidad en la distribución de los datos.

Los valores de las medidas de dispersión nos permiten saber si los datos se encuentran estrechamente agrupados, si se encuentran ampliamente dispersos o si

son iguales.

En la siguiente gráfica se muestra gráficas de tres muestras con diferente dispersión en torno al centro. La primera tiene una medida más grande de dispersión, la tercera tiene la cantidad más pequeña, mientras que la segunda es intermedia con respecto a las otras dos.



*Fuente: Matemóvil*

En nuestro curso de estadística, estudiaremos las siguientes **medidas de dispersión**:

- Desviación estándar.
- La desviación media.
- Varianza.
- Desviación estándar.
- Coeficiente de variación.

## 12. Rango, amplitud o recorrido

El rango es la diferencia entre el valor máximo y mínimo de un conjunto de datos. Es la [medida de dispersión](#) o variabilidad más sencilla de calcular y tiene

las mismas unidades que los datos. También es llamado amplitud o recorrido.

### 12.1. La fórmula del rango es la siguiente:

$$\text{Rango} = \text{valor máximo} - \text{valor mínimo}$$

Ejemplo:

Dado el conjunto de datos, 1, 3, 5 y 7, encontrar el rango.

**Solución:**

De los datos, 1, 3, 5 y 7, el valor menor es el 1, y el valor mayor es el 7, por lo tanto, aplicamos la fórmula:

$$\text{➤ Rango} = \text{valor máximo} - \text{valor mínimo}$$

$$\text{➤ Rango} = 7 - 1$$

$$\text{➤ Rango} = 6$$

Ejemplo 2:

Las ganancias de la primera mitad del año pasado de una empresa que vende ositos de peluche en lata, son las siguientes:

Tabla 25 Ejercicio 2

	Enero	Fe- brero	Marzo	Abril	Mayo	Junio
Ga- nan- cias	\$ 16 800	\$ 34 500	\$ 17 300	\$ 12 500	\$ 14 000	\$ 18 600

Obtenido de: Matemovil.com

**Solución:**

- Rango = valor máximo – valor mínimo
- Rango = \$ 34 500 – \$ 12 500
- Rango = \$ 22 000

### 12.2. Características del rango

- Tiene la misma unidad de medida que las observaciones.
- Se utiliza para tener una idea rápida del grado de dispersión de un conjunto de datos.
- Es poco confiable.
- El rango muestral es muy inestable.
- El valor del rango no varía cuando se suma una constante K a cada observación de un conjunto de datos.
- El valor del rango si varía cuando se multiplica por constante K a cada observación de un conjunto de datos.

### 13. Varianza y desviación estándar

La varianza y la desviación estándar son medidas de dispersión o variabilidad, es decir, indican la dispersión o separación de un conjunto de datos. Hay que tener en cuenta que las fórmulas de la varianza y la desviación estándar son diferentes para una muestra que para una población.

A continuación, presentamos el resumen de fórmulas, las cuales analizaremos líneas abajo:

Ilustración 13 Fórmulas de varianza y desviación estándar

Fórmulas de Varianza y Desviación Estándar			
	Varianza	Desviación Estándar	Media
Población	$\sigma^2 = \frac{\sum_{i=1}^N (x_i - \mu)^2}{N}$	$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum_{i=1}^N (x_i - \mu)^2}{N}}$	$\mu = \frac{\sum_{i=1}^N x_i}{N}$
Muestra	$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$	$s = \sqrt{s^2} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$	$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$

Fuente: Matemóvil

#### 13.1. Varianza de la población ( $\sigma^2$ )

La varianza se define como la media aritmética de los cuadrados de las diferencias de los datos con su media aritmética.

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

### 13.2. Desviación estándar de la población ( $\sigma$ )

La desviación estándar es la raíz cuadrada positiva de la varianza.

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

Te recomendamos calcular primero la varianza de la población y luego sacar su raíz cuadrada para obtener la desviación estándar.

Ten en cuenta que, si tienes una serie de valores de una población y necesitas calcular su varianza y su desviación estándar, deberás calcular primero la media poblacional  $\mu$  con la siguiente fórmula:

$$\mu = \frac{\sum_{i=1}^N x_i}{N}$$

### 13.3. Varianza de la muestra ( $s^2$ )

La fórmula de la varianza de la muestra es diferente a la de varianza de la población.

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

### 13.4. Desviación estándar de la muestra ( $s$ )

Recuerda que la desviación estándar es la raíz cuadrada positiva de la varianza.

$$s = \sqrt{s^2} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

Te recomendamos calcular primero la varianza de la muestra y luego sacar su raíz cuadrada para obtener la desviación estándar.

Ten en cuenta que, si tienes una serie de valores de una muestra y necesitas calcular su varianza y su desviación estándar, deberás calcular primero la media poblacional  $\bar{x}$  con la siguiente fórmula:

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

En los ejercicios, se siguen los siguientes pasos:

- Se calcula la media.
- Se calcula la varianza.
- Se calcula la desviación estándar, que es la raíz cuadrada positiva de la varianza.

Ejemplo:

Calcular la varianza y la desviación estándar de los siguientes datos: 2, 4, 6 y 8 sabiendo que corresponden a una población.

**Solución:**

Nos indican que estos datos forman una población, por lo tanto, usaremos las fórmulas de varianza y desviación estándar para la población, teniendo en cuenta que tenemos 4 datos, es decir,  $N = 4$ .

Empezamos calculando la media poblacional:

$$\mu = \frac{\sum_{i=1}^N x_i}{N} = \frac{x_1 + x_2 + x_3 + x_4}{N} = \frac{2 + 4 + 6 + 8}{4} = \frac{20}{4} = 5$$

Ahora calculamos la varianza poblacional:

$$\sigma^2 = \frac{\sum_{i=1}^N (x_i - \mu)^2}{N} = \frac{(x_1 - \mu)^2 + (x_2 - \mu)^2 + (x_3 - \mu)^2 + (x_4 - \mu)^2}{N}$$

$$\sigma^2 = \frac{(2-5)^2 + (4-5)^2 + (6-5)^2 + (8-5)^2 + (-3)^2 + (-1)^2 + (1)^2 + (3)^2}{4} = \frac{9+1+1+9}{4} = \frac{20}{4} = 5$$

El valor de la varianza poblacional es de 5.

Ahora calculamos la desviación estándar, teniendo en cuenta que es la raíz cuadrada de la varianza.

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

$$\sigma = \sqrt{5} = 2,236$$

Ejemplo:

Calcular la varianza y la desviación estándar de los siguientes datos: 1, 3, 5, 7 y 9 sabiendo que corresponden a una muestra

Solución:

Nos indican que estos datos forman una muestra, por lo tanto, usaremos las fórmulas de varianza y desviación estándar para la muestra, teniendo en cuenta que tenemos 5 datos, es decir,  $n = 5$ .

Empezamos calculando la media de la muestra:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{x_1 + x_2 + x_3 + x_4 + x_5}{n} = \frac{1+3+5+7+9}{5} = \frac{25}{5} = 5$$

Ahora calculamos la varianza de la muestra:

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2 + (x_4 - \bar{x})^2 + (x_5 - \bar{x})^2}{n-1}$$

$$s^2 = \frac{(1-5)^2 + (3-5)^2 + (5-5)^2 + (7-5)^2 + (9-5)^2}{5-1}$$

$$s^2 = \frac{(-4)^2 + (-2)^2 + (0)^2 + (2)^2 + (4)^2}{4}$$

$$s^2 = \frac{16+4+0+4+16}{4} = \frac{40}{4} = 10$$

El valor de la varianza poblacional es de 10.

Ahora calculamos la desviación estándar, teniendo en cuenta que es la raíz cuadrada de la varianza.

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

$$s = \sqrt{10} = 3,162$$

Ejemplo:

Calcular la varianza y la desviación estándar de los siguientes datos: 10, 12, 13, 16, 9, 8, 12, 8, 6, 16 sabiendo que corresponden a una población.

Solución:

Empezaremos calculando la media y la varianza usando las fórmulas de la población.

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

En este caso, como tenemos muchos datos, recurriremos a una tabla para mantener el orden. Colocaremos los valores de los elementos de la población ( $x_i$ ) y sumaremos los valores.

Tabla 26 Datos (xi)

$x_i$
10
12
13
16
9
8
12
8
6
16
$\Sigma$
110

Fuente: Matemóvil

Teniendo en cuenta que tenemos 10 datos ( $N = 10$ ), calculamos la media:

$$\mu = \frac{\sum_{i=1}^N x_i}{N} = \frac{110}{10} = 11$$

Con el valor de la media, vamos en busca de la varianza poblacional:

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

Agregamos 2 columnas más a nuestra tabla para llegar a la forma de la varianza:

Tabla 27 Varianza

$x_i$	$x_i - \mu$	$(x_i - \mu)^2$
10	-1	1
12	1	1
13	2	4
16	5	25
9	-2	4
8	-3	9
12	1	1
8	-3	9
6	-5	25
16	5	25
$\Sigma$	110	104

Fuente: Matemóvil

Reemplazamos los valores en la fórmula:

$$\sigma^2 = \frac{\sum_{i=1}^N (x_i - \mu)^2}{N} = \frac{104}{10} = \mathbf{10,4}$$

La varianza tiene un valor de 10,4.

Finalmente calculamos la desviación estándar:

$$\sigma = \sqrt{\sigma^2} = \sqrt{10,4} = \mathbf{3,225}$$

La desviación estándar tiene un valor de 3,225.

#### 14. Varianza y desviación estándar de datos agrupados de variable discreta

Si trabajamos con una tabla de frecuencias con datos agrupados de una variable discreta, es decir, la variable toma valores puntuales, no intervalos de valores, podemos calcular la varianza y la desviación estándar

usando las fórmulas que veremos en esta clase.

Veamos los ejemplos y ejercicios que hemos preparado.

Fórmulas para la varianza y desviación estándar de datos agrupados

Ilustración 14 Fórmulas para varianza y desviación

Varianza y desviación estándar para datos agrupados				
	Varianza	Desviación estándar	Media	Número de elementos
<b>Población</b>	$\sigma^2 = \frac{\sum_{i=1}^k f_i(x_i - \mu)^2}{N}$	$\sigma = \sqrt{\sigma^2}$	$\mu = \frac{\sum_{i=1}^k x_i \cdot f_i}{N}$	$N = \sum_{i=1}^k f_i$
<b>Muestra</b>	$s^2 = \frac{\sum_{i=1}^k f_i(x_i - \bar{x})^2}{n - 1}$	$s = \sqrt{s^2}$	$\bar{x} = \frac{\sum_{i=1}^k x_i \cdot f_i}{n}$	$n = \sum_{i=1}^k f_i$

Fuente: Matemóvil

Donde:

- $f_i$ : frecuencia absoluta de cada valor, es decir, número de veces que aparece el valor en el estudio.
- $x_i$ : valor de los elementos de la población.
- $\sigma^2$ : varianza de la población.
- $\sigma$ : desviación estándar de la población.
- $\mu$ : media poblacional.
- $s^2$ : varianza de la muestra.

- $s$ : desviación estándar de la muestra.
- $\bar{x}$ : media de la muestra.
- $k$ : número de clases.

Tenemos siempre que fijarnos si estamos trabajando con datos que forman una población o con datos que forman una muestra, pues las fórmulas son diferentes.

Podemos ver también que la fórmula de la varianza se presenta de 2 formas diferentes, puedes tomar cualquiera de ellas, obtendrás el mismo resultado.

En los problemas, seguiremos los siguientes pasos:

0. Calculamos el número de elementos.

1. Calculamos la media.

2. Calculamos la varianza.

3. Calculamos la desviación estándar, que es la raíz cuadrada de la varianza.

Ejemplo 1:

Calcular la varianza y la desviación estándar de las edades de una población de niños que asisten a una fiesta infantil.

Tabla 28 Edades de una población de niños

Edad (años) $x_i$	Frecuencia $f_i$
3	6
4	12
5	4
6	12
7	6

Fuente: Matemóvil

Solución:

Para calcular la varianza y la desviación estándar, empezamos calculand  $N = \sum_{i=1}^k f_i$  número de elementos de la población:

En la tabla, sumamos las frecuencias  $f_i$ :

Tabla 29 Suma de frecuencias

Edad (años) $x_i$	Frecuencia $f_i$
3	6
4	12
5	4
6	12
7	6
$\Sigma$	40

Fuente: Matemóvil

Y así obtenemos el número de elementos de la

población (N).

$$N = \sum_{i=1}^k f_i = 40$$

A continuación, calculamos la media poblacional partiendo de su fórmula:

$$\mu = \frac{\sum_{i=1}^k x_i \cdot f_i}{N}$$

En la tabla de frecuencias agregamos la columna  $x_i \cdot f_i$

Edad (años) $x_i$	Frecuencia $f_i$	$x_i \cdot f_i$
3	6	18
4	12	48
5	4	20
6	12	72
7	6	42
$\Sigma$	40	200

Fuente: Matemóvil

Ahora sí, calculamos la media poblacional:

$$\mu = \frac{\sum_{i=1}^k x_i \cdot f_i}{N} = \frac{200}{40} = 5 \text{ años}$$

A continuación, recordamos la fórmula de la varianza de la población:

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

En la tabla, agregamos 3 columnas más, para buscar

la expresión de la fórmula:

Tabla 31 Expresión de la fórmula

Edad (años) $x_i$	Frecuencia $f_i$	$x_i \cdot f_i$	$x_i - \mu$	$(x_i - \mu)^2$	$f_i \cdot (x_i - \mu)^2$
3	6	18	-2	4	24
4	12	48	-1	1	12
5	4	20	0	0	0
6	12	72	1	1	12
7	6	42	2	4	24
$\Sigma$	40	200			72

Fuente: Matemóvil

Usamos la fórmula:

$$\sigma^2 = \frac{\sum_{i=1}^k f_i (x_i - \mu)^2}{N} = \frac{72}{40} = 1,8 \text{ (años)}^2$$

El valor de la varianza de esta población es de 1,8 (años)<sup>2</sup>. Ten en cuenta que la varianza queda expresada en las unidades originales elevadas al cuadrado, por ello, nos quedaría en (años)<sup>2</sup>.

Finalmente calculamos la desviación estándar, teniendo en cuenta que es la raíz cuadrada positiva de la varianza:

$$\sigma = \sqrt{\sigma^2} = \sqrt{1,8 \text{ (años)}^2} = 1,34 \text{ años}$$

El valor de la desviación estándar poblacional es de 1,34 años.

## 15. Coeficiente de variación

El coeficiente de variación es una medida de dispersión relativa (libre de unidades de medida), que se define como el cociente de la desviación estándar entre la media aritmética.

En esta clase, veremos cómo calcular su valor y también su utilidad.

### 15.1. Introducción

En ocasiones, necesitamos comparar la variabilidad o dispersión de dos conjuntos de datos, sin embargo, al hacerlo, puede que ambos conjuntos estén expresados en diferentes unidades de medida (por ejemplo, uno en metros, otro en litros), por lo tanto, no se podrán comparar sus varianzas o desviaciones estándar. También puede darse el caso de que estén expresados en la misma unidad de medida, pero nos interesa determinar la variación respecto a una base. Para estos casos, se utiliza el coeficiente de variación.

### 15.2. Definición y fórmula

El coeficiente de variación es una medida de dispersión relativa (libre de unidades de medida), que se define como el cociente de la desviación estándar entre la media aritmética. Su fórmula es la siguiente:

Ilustración 15 Fórmula coeficiente de variación

Fórmula del Coeficiente de Variación (CV)	
CV de la población	CV de la muestra
$CV = \frac{\text{desviación estándar}}{\text{media}} = \frac{\sigma}{\mu}$	$CV = \frac{\text{desviación estándar}}{\text{media}} = \frac{s}{\bar{x}}$

Fuente: Matemóvil

Donde:

$\sigma$ : desviación estándar de la población.

$\mu$ : media de la población.

$s$ : desviación estándar de la muestra.

$\bar{x}$ : media de la muestra.

Para la población y para la muestra, aunque tengan notación diferente, el coeficiente de variación se obtiene dividiendo la desviación estándar entre la muestra.

Algunos autores, expresan el coeficiente de variación de forma porcentual. En ese caso, solo multiplicamos la fórmula por el 100%.

Ilustración 16 Fórmula porcentual del coeficiente de variación

Fórmula Porcentual del Coeficiente de Variación (CV)	
CV de la población	CV de la muestra
$CV = \frac{\text{desviación estándar}}{\text{media}} \cdot 100\% = \frac{\sigma}{\mu} \cdot 100\%$	$CV = \frac{\text{desviación estándar}}{\text{media}} \cdot 100\% = \frac{s}{\bar{x}} \cdot 100\%$

Fuente: Matemóvil

El coeficiente de variación se suele usar para comparar

el grado de dispersión de dos o más conjuntos de datos; incluso si tienen medidas diferentes. Es de gran utilidad cuando se necesita comparar las dispersiones de dos conjuntos de datos cuyas medias son muy diferentes. De dos conjuntos de datos, el más homogéneo es el que tiene menor coeficiente de variación.

Esta medida de dispersión fue propuesta por Karl Pearson (1895) para comparar la variabilidad o dispersión entre varias distribuciones de frecuencias.

Ejemplo:

Una población de alumnos tiene una estatura media de 160 cm con una desviación estándar de 16 cm. Estos mismos alumnos, tienen un peso medio de 70 kg con una desviación estándar de 14 kg. ¿Cuál de las 2 variables presenta mayor variabilidad relativa?

**Solución:**

Vamos a comparar la dispersión de 2 variables, la estatura y el peso, usando el coeficiente de variación.

Ilustración 17 Comparación de dispersión entre dos variables

Estatura (E)	Peso (P)
$\mu_E = 160 \text{ cm} \quad \wedge \quad \sigma_E = 16 \text{ cm}$	$\bar{x}_P = 70 \text{ kg} \quad \wedge \quad s_P = 14 \text{ kg}$
$CV_E = \frac{\sigma_E}{\mu_E} = \frac{16 \text{ cm}}{160 \text{ cm}} = \frac{1}{10} = 0,1 = 10\%$	$CV_P = \frac{s_P}{\bar{x}_P} = \frac{14 \text{ kg}}{70 \text{ kg}} = \frac{1}{5} = 0,2 = 20\%$

$$D.M. = \frac{|x_1 - \bar{x}| + |x_2 - \bar{x}| + |x_3 - \bar{x}| + \dots + |x_n - \bar{x}|}{n} = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n}$$

Fuente: Matemóvil

Podemos que ver que  $CV_P > CV_E$ , por eso, el peso de esta población de alumnos tiene mayor variabilidad relativa que la estatura.

### 16. Desviación media

La desviación media de un conjunto de datos es la media aritmética de los valores absolutos de lo que se desvía cada valor respecto a la media aritmética.

La fórmula de la desviación media es la siguiente:

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n}$$

Donde:

$\bar{x}$ : media aritmética de los datos.

$x_1, x_2, x_3, \dots, x_n$ : datos.

$x_i$ : cada uno de los datos.

$n$ : número de datos.

Recuerda calcular la media aritmética  $\bar{x}$  antes de aplicar la fórmula de la desviación media. Su fórmula es esta:

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n} = \frac{2 + 4 + 6 + 8}{4} = \frac{20}{4} = 5$$

La desviación media también es llamada desviación promedio de la media o desviación absoluta promedio. Es una medida de dispersión poco usada debido a la dificultad de hacer cálculos con la función valor absoluto.

Ejemplo 1:

Calcular la desviación media de los siguientes datos: 2, 4, 6 y 8.

Solución:

Empezamos calculando la media aritmética de los datos, teniendo en cuenta que tenemos 4 datos ( $n = 4$ ).

El valor de la media aritmética es de 5.

Ahora aplicamos la fórmula de la desviación media:

$$D.M. = \frac{|x_1 - \bar{x}| + |x_2 - \bar{x}| + |x_3 - \bar{x}| + \dots + |x_n - \bar{x}|}{n} = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n}$$

$$D.M. = \frac{|2 - 5| + |4 - 5| + |6 - 5| + |8 - 5|}{4} = \frac{|-3| + |-1| + |1| + |3|}{4}$$

$$D.M. = \frac{3 + 1 + 1 + 3}{4} = \frac{8}{4} = 2$$

El valor de la desviación media es de 2.

### 17. Desviación media de datos agrupados de variable discreta

Si trabajamos con una tabla de frecuencias con datos agrupados de una variable discreta, es decir, la variable toma valores puntuales, no intervalos de valores, podemos **calcular la desviación media** usando las fórmulas que veremos en esta clase.

Las fórmulas que usaremos para calcular la desviación media con datos agrupados son las siguientes:

*Ilustración 18* Calculo de desviación media para datos agrupados

Desviación media para datos agrupados	
Desviación media	Media
$D.M. = \frac{\sum_{i=1}^k f_i  x_i - \bar{x} }{n}$	$\bar{x} = \frac{\sum_{i=1}^k x_i \cdot f_i}{n}$

Fuente: Matemóvil

Donde:

- $f_i$ : frecuencia absoluta de cada valor, es decir, número de veces que aparece el valor en el estudio.
- $x_i$ : cada uno de los datos.
- $k$ : número de clases.
- $D.M.$ : desviación media.

En los problemas, seguiremos los siguientes pasos:

1. Calculamos el número de elementos.
2. Calculamos la media.
3. Calculamos la desviación media.

Ejemplo:

Calcular la desviación media de las longitudes de la tabla:

*Tabla 31* Valores para la desviación media

Longitud (m)	Frecuencia $f_i$
3	7
4	11
5	7

Fuente: Matemóvil

Solución:

Empezamos calculando el número de valores o elementos “n”, simplemente sumando las frecuencias absolutas  $f_i$ .

Tabla 32 Suma de frecuencias absolutas

Longitud (m) $x_i$	Frecuencia $f_i$
3	7
4	11
5	7
$\Sigma$	25

$$n = \sum_{i=1}^k f_i = 25$$

Fuente: Matemóvil

El número de valores es 25.

A continuación, vamos a calcular la media  $\bar{x}$ , recordemos su fórmula:

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

En la tabla, iremos buscando las expresiones que aparecen en la fórmula. Por ello, agregamos una columna más a nuestra tabla, en la cual colocaremos los valores de  $x_i \cdot f_i$ .

Tabla 33 Valores de  $x_i \cdot f_i$

Longitud (m) $x_i$	Frecuencia	$x_i \cdot f_i$
3	7	21
4	11	44
5	7	35
$\Sigma$	25	100

Fuente: Matemóvil

Ahora sí, aplicamos la fórmula:

$$\bar{x} = \frac{\sum_{i=1}^k x_i \cdot f_i}{n} = \frac{100}{25} = 4 \text{ m}$$

El valor de la media es de 4 m.

La fórmula de la desviación media para datos agrupados es:

$$D.M. = \frac{\sum_{i=1}^k f_i |x_i - \bar{x}|}{n} = \frac{14}{25} = 0,56 \text{ m}$$

Agregamos 3 columnas más en la tabla de frecuencias:

Tabla 34 Frecuencias

Longitud (m) $x_i$	Frecuencia $f_i$	$x_i \cdot f_i$	$x_i - \bar{x}$	$ x_i - \bar{x} $	$f_i  x_i - \bar{x} $
3	7	21	-1	1	7
4	11	44	0	0	0
5	7	35	1	1	7
$\Sigma$	25	100			14

Fuente: Matemóvil

Aplicamos la fórmula:

$$D.M. = \frac{\sum_{i=1}^k f_i |x_i - \bar{x}|}{n}$$

El valor de la desviación media es de 0,56 m. Recuerda que la desviación media siempre queda expresada en las mismas unidades que los datos originales. En nuestro ejercicio, las longitudes estaban expresadas en metros, por ello, la desviación expresada en metros.

### 17. Desviación media para datos agrupados por intervalos

Si necesitamos calcular la desviación media de un

conjunto de datos agrupados por intervalos en una tabla de frecuencias, usaremos las fórmulas que revisaremos en esta clase.

Veamos los ejemplos que hemos preparado.

### 17.2. Fórmulas de la desviación media para datos agrupados

Ilustración 19 Fórmulas para la desviación media de datos agrupados

Desviación media para datos agrupados	
Desviación media	Media
$D.M. = \frac{\sum_{i=1}^k f_i  x_i - \bar{x} }{n}$	$\bar{x} = \frac{\sum_{i=1}^k x_i \cdot f_i}{n}$

Fuente: Matemóvil

Donde:

- $f_i$ : frecuencia absoluta de cada valor, es decir, el número de veces que aparece el valor en el estudio.
- $x_i$ : marca de clase. Es el punto medio del límite inferior y el límite superior de cada intervalo.
- $k$ : número de clases.
- $D.M.$ : desviación media.
- $\bar{x}$ : media aritmética de los datos.

En los problemas, seguiremos los siguientes pasos:

- Calculamos las marcas de clase  $x_i$ .
- Calculamos el número de elementos  $n$ .
- Calculamos la media  $\bar{x}$ .
- Calculamos la desviación media D.M..

Ejemplo:

Calcular la desviación media de las edades indicadas en la tabla de frecuencias:

Tabla 35 Edades

Edad (años)	Frecuencia $f_i$
[0 - 2)	7
[2 - 4)	8
[4 - 6)	8
[6 - 8]	7

Fuente: Matemóvil

Solución:

Lo primero que haremos será calcular las marcas de clase  $x_i$ . Recuerda que la marca de clase es el punto medio del límite inferior y del límite superior de cada intervalo, su fórmula es:

$$x_i = \frac{L_i + L_s}{2}$$

Agregamos una columna más a nuestra tabla para la marca de clase  $x_i$ :

Tabla 36 Marca de clase  $x_i$

Edad (años)	Marca de clase $x_i$	Frecuencia $f_i$
[0 - 2)	1	7
[2 - 4)	3	8
[4 - 6)	5	8
[6 - 8]	7	7

Fuente: Matemóvil

A continuación, calculamos el número de valores o número de datos “n”, solo tenemos que sumar las frecuencias:

Tabla 38 Número de datos

Edad (años)	Marca de clase $x_i$	Frecuencia $f_i$
[0 - 2)	1	7
[2 - 4)	3	8
[4 - 6)	5	8
[6 - 8]	7	7
	$\Sigma$	30

Fuente: Matemóvil

$$n = \sum_{i=1}^k f_i = 30$$

El número de valores es 30.

A continuación, vamos a calcular la media  $\bar{x}$ , recordemos su fórmula:

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

En la tabla, iremos buscando las expresiones que aparecen en la fórmula. Por ello, agregamos una columna más a nuestra tabla, en la cual colocaremos los valores de  $x_i \cdot f_i$ .

Tabla 39 Valores  $X_i \cdot F_i$

Edad (años)	Marca de clase $x_i$	Frecuencia $f_i$	$x_i \cdot f_i$
[0 - 2)	1	7	7
[2 - 4)	3	8	24
[4 - 6)	5	8	40
[6 - 8]	7	7	49
	$\Sigma$	30	120

Fuente: Matemóvil

Ahora sí, aplicamos la fórmula:

$$\bar{x} = \frac{\sum_{i=1}^k x_i \cdot f_i}{n} = \frac{120}{30} = 4 \text{ años}$$

El valor de la media es de 4 años.

La fórmula de la desviación media para datos agrupados es:

$$D.M. = \frac{\sum_{i=1}^k f_i |x_i - \bar{x}|}{n}$$

Agregamos 3 columnas más en la tabla de frecuencias:

Tabla 40 Frecuencias

Edad (años)	Marca de clase $x_i$	Frecuencia $f_i$	$x_i \cdot f_i$	$x_i - \bar{x}$	$ x_i - \bar{x} $	$f_i  x_i - \bar{x} $
[0 - 2]	1	7	7	-3	3	21
[2 - 4]	3	8	24	-1	1	8
[4 - 6]	5	8	40	1	1	8
[6 - 8]	7	7	49	3	3	21
	$\Sigma$	30	120			58

Fuente: Matemóvil

Aplicamos la fórmula:

$$D.M. = \frac{\sum_{i=1}^k f_i |x_i - \bar{x}|}{n} = \frac{58}{30} = 1,93 \text{ años}$$

El valor de la desviación media es de 1,93 años.



## CAPÍTULO IV

### TÉCNICAS DE CONTEO

## 18. Permutaciones y combinaciones, ejercicios resueltos

Una **permutación** de un conjunto de elementos es una disposición de dichos elementos **teniendo en cuenta el orden**. Una **combinación** de un conjunto de elementos es una selección de dichos elementos **sin tener en cuenta el orden**.

La diferencia entre permutaciones y combinaciones es que **en las permutaciones importa el orden de los elementos, mientras que en las combinaciones no importa el orden** en que se disponen los elementos (solo importa su presencia).

Veamos algunos conceptos adicionales, ejemplos y ejercicios resueltos.

### 18.1 Permutaciones

Una permutación de un conjunto de elementos es una disposición de dichos elementos **teniendo en cuenta el orden**. El número de permutaciones de “n” elementos tomados de “k” en “k” se calcula con la fórmula:

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

#### Ejemplo 1:

Eduardo, Carlos y Sergio se han presentado a un concurso de pintura. El concurso otorga \$200 al primer lugar y \$100 al segundo. ¿De cuántas formas se pueden repartir los premios de primer y segundo lugar?

**Solución:**

En este caso, si importa el orden, ya que no es lo mismo quedar en primer lugar que en segundo, además, los premios son diferentes. Por ejemplo, un arreglo o disposición, es que Carlos ocupe el primer lugar y Sergio el segundo. Otro arreglo, sería que Sergio ocupe el primer lugar y Eduardo el segundo. El número total de arreglos o formas lo calculamos con la fórmula:

$n = 3$  (número total de elementos);  $k = 2$  (tomados de dos en dos)

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

$$P_2^3 = \frac{3!}{(3-2)!} = \frac{3 \times 2 \times 1}{1!} = \frac{6}{1} = 6 \text{ formas}$$

**18.2. Combinaciones**

Una combinación de un conjunto de elementos es una selección de dichos elementos **sin tener en cuenta el orden**.

El número de combinaciones de “n” elementos tomados de “k” en “k” se calcula con la fórmula:

**Ejemplo 2:**

Un chef va a preparar una ensalada de verduras con tomate, zanahoria, papa y brócoli. ¿De cuántas formas se puede preparar la ensalada usando solo 2 ingredientes?

**Solución:**

En este caso, no importa el orden en que se tomen los ingredientes para la ensalada, pues da igual si es una ensalada de tomate con zanahoria, que una ensalada de zanahoria con tomate, ya que al final, el chef mezclará los dos ingredientes.

Un arreglo podría ser zanahoria y tomate, otro arreglo podría ser tomate y papa, otro arreglo podría ser papa y brócoli. El problema nos indica que solo se pueden usar 2 ingredientes en la ensalada. El número total de arreglos o formas lo calculamos con la fórmula:

$n = 4$  (número total de elementos);  $k = 2$  (tomados de dos en dos)

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

$$C_2^4 = \frac{4!}{(4-2)! 2!} = \frac{4 \times 3 \times 2 \times 1}{2! \times 2!} = \frac{24}{2 \times 2} = \frac{24}{4} = 6 \text{ formas}$$

**19. Permutación con repetición, ejercicios resueltos**

La permutación con repetición se usa cuando en un total de “n” elementos, el primero se repite “a” veces, el segundo “b” veces, el tercero “c” veces...

La fórmula para calcular el número de permutaciones u ordenamientos, es la siguiente:

$$P_{a;b;c;\dots}^n = \frac{n!}{a! b! c! \dots}$$

Donde:

- $n = a+b+c$

Hay tres condiciones en la permutación con repetición:

- Importa el orden.
- Hay elementos repetidos.
- Participan todos los elementos en los ordenamientos.

Ejemplo

¿Cuántas palabras diferentes se pueden formar con las letras de la palabra BANANA?

Solución:

Coloquemos algunas palabras como ejemplos:

- ANANAB.
- BANAAN.
- NAANAB.

En este caso, podemos ver que algunos elementos se repiten: la letra A aparece 3 veces en la palabra BANANA, mientras que la letra N aparece 2 veces. Además, importa el orden en el que se coloquen las letras, ya que ANANAB no es la misma palabra que BANAAN. Finalmente, participan todos los elementos en los ordenamientos, es decir, las 6 letras de la palabra BANANA.

Entonces, podemos aplicar la **fórmula de permutación con repetición**, teniendo en cuenta que:

- Número de veces que se repite la letra B = 1

- Número de veces que se repite la letra A = 3
- Número de veces que se repite la letra N = 2
- Número total de elementos:  $n = 1+3+2 \rightarrow n = 6$

$$P_{1;3;2}^6 = \frac{6!}{1!3!2!} = \frac{6 \times 5 \times 4 \times 3 \times 2 \times 1}{1 \times 3 \times 2 \times 1 \times 2 \times 1} = \frac{720}{12} = 60$$

En total, se pueden formar **60 palabras diferentes** con las letras de la palabra BANANA.

## 20. Permutación circular

La permutación circular, es un caso de permutación en el cual **los elementos se ordenan en círculo**. De modo que el primero elemento que se sitúa en el ordenamiento determina el principio y el final de la muestra.

La **fórmula** para calcular el número de permutaciones circulares es:

$$PC_n = (n - 1)!$$

Donde “n” es el número de elementos.

Hay 3 condiciones importantes que se cumplen en las permutaciones circulares:

- Importa el orden.
- Los elementos se ordenan en círculo.
- Participan todos los elementos en los ordenamientos.

Los problemas clásicos de ordenamientos circulares, nos preguntan de cuántas formas se pueden ordenar

«n» elementos alrededor de una piedra circular, de una mesa circular, de una fogata, entre otros.

Ejemplo:

¿De cuántas maneras se pueden sentar 5 amigos alrededor de una mesa circular?

**Solución:**

Número de elementos:  $n = 5$ .

Ahora calculamos el número de permutaciones circulares:

$$PC_n = (n - 1)!$$
$$PC_5 = (5 - 1)! = 4! = 4 \times 3 \times 2 \times 1 = 24$$

Los 5 amigos, se pueden sentar de 24 formas diferentes.

## 21. Variaciones, combinaciones y permutaciones.

Continuamos con nuestro curso de estadística, y para no tener complicaciones en la sesión de probabilidades, vamos a ver a detalle las **variaciones, combinaciones y permutaciones**.

### 21.1. Definiciones

➤ **Variación:** Es la disposición de una parte del total de elementos en un orden determinado. Aquí si importa el orden.

➤ **Combinación:** Disposición de una parte del total de elementos sin tener en cuenta el orden. Aquí no importa el orden de los elementos.

➤ **Permutación:** Es la disposición de todos los elementos en un orden determinado. Aquí si importa el orden. Por ejemplo, si quiero saber cuántos resultados posibles puede tener una carrera en la que participan 4 caballos, tengo que ordenar a todos los elementos, es decir, a los 4 caballos, como no es lo mismo salir primero que segundo en la carrera, aquí si importa el orden, y se necesita ordenar a todos los elementos, por ello, se trata de una permutación de 4 elementos.



## CAPÍTULO V

### PROBABILIDADES

## 22. Probabilidades

Probabilidad es un valor entre 0 y 1, que indica la posibilidad relativa de que ocurra un evento.

La fórmula de probabilidad es la siguiente:

$$P(A) = \frac{\text{número de casos favorables de } A}{\text{número total de casos posibles}}$$

Mientras más se acerca el valor de la probabilidad a 0, disminuye la posibilidad de que ocurra el evento. Mientras más se acerca el valor a 1, aumenta la posibilidad de que ocurra.



 Probabilidad es un valor entre 0 y 1.

La probabilidad de que ocurra un evento es 0, si es imposible que ocurra ese evento. Por otro lado, la probabilidad de que un ocurra un evento es 1, si es seguro que ocurrirá ese evento.

Ejemplo 1:

La moneda de México tiene 2 caras: águila y sello. ¿Cuál es la probabilidad de obtener águila al lanzar una moneda?

**Solución:**

Primero calculamos el número total de casos posibles que se dan al lanzar la moneda. En este problema, son 2 casos posibles, se obtiene águila o se obtiene sello.

Ahora, calculamos el número de casos favorables. Si lanzamos la moneda, tenemos 1 caso de águila. Por lo tanto, la probabilidad de obtener águila sería:

$$P(\text{águila}) = \frac{\text{número de casos favorables de águila}}{\text{número total de casos posibles}} = \frac{1}{2} = 0,5 = 50\%$$

Podemos colocar como respuesta: 0,5 o 50%.

**Ejemplo 2:**

¿Cuál es la probabilidad de obtener un 5 al lanzar un dado?

**Solución:**

Primero calculamos el número total de casos posibles que se dan al lanzar un dado. En este problema, son 6 casos posibles, ya que el dado puede arrojar 1, 2, 3, 4, 5 o 6.

Ahora, calculamos el número de casos favorables. Si lanzamos un dado, tenemos 1 caso en el que se obtiene 5. Por lo tanto, la probabilidad de obtener un 5 sería:

$$P(5) = \frac{\text{número de casos favorables de 5}}{\text{número total de casos posibles}} = \frac{1}{6} = 0,1667 = 16,67\%$$

La respuesta sería: 0,1667 o 16,67%.

**Ejemplo 3:**

Si se lanza una moneda de México al aire dos veces, ¿cuál es la probabilidad de obtener al menos 1 águila?

**Solución:**

Primero calculamos el número total de casos posibles. Los casos posibles del primer y segundo lanzamiento son:

- Águila – águila.
- Águila – sello.
- Sello – águila.
- Sello – sello.

En total, tenemos 4 casos posibles.

Ahora calculamos el número de casos en los cuáles se obtiene al menos 1 águila. Los casos son:

- Águila – águila.
- Águila – sello.
- Sello – águila.

Es decir, tenemos 3 casos favorables. Por lo tanto, la probabilidad de obtener al menos un águila es:

$$P(\text{al menos 1 águila}) = \frac{3}{4} = 75\%$$

La respuesta sería: 0,75 o 75%.

## 23. Experimento aleatorio, espacio muestral, evento y probabilidad

Veamos ahora la definición clásica de probabilidad, además de las definiciones de experimento aleatorio, espacio muestral y evento o suceso.

### 23.1. Experimento aleatorio

Es la reproducción controlada de un fenómeno; y cuyo resultado depende del azar.

Ejemplos

- Lanzamiento de un dado.
- Lanzamiento de una moneda.

Un experimento aleatorio puede ser repetido bajo las mismas condiciones, y se puede describir el número de resultados posibles.

### 23.2. Espacio muestral (S)

Es el conjunto de todos los resultados posibles de un experimento aleatorio.

Si se lanza un dado, el espacio muestral está compuesto por los siguientes elementos:  $S=\{1, 2, 3, 4, 5, 6\}$ .

Si se lanza una moneda que tiene dos caras: perro (P) y gato (G), el espacio muestral está compuesto por:  $S=\{P, G\}$ .

- Si se lanzan dos monedas, el espacio muestral está

compuesto por:  $S=\{(P, P), (P, G), (G, P), (G, G)\}$ .

• Si se lanza un dado y una moneda, el espacio muestral está compuesto por:  $S= \{(1,P), (1,G), (2,P), (2,G), (3,P), (3,G), (4,P), (4,G), (5,P), (5,G), (6,P), (6,G)\}$

### 23.3. Evento o suceso

Conjunto de uno o más resultados del experimento aleatorio.

- Si  $A = \{\text{obtener un número 5 al lanzar un dado}\}$ , entonces,  $A=\{5\}$ .
- Si  $B = \{\text{obtener un número mayor que 3 al lanzar un dado}\}$ , entonces,  $B=\{4, 5, 6\}$ .
- Si  $C = \{\text{obtener un número par al lanzar un dado}\}$ , entonces,  $C=\{2, 4, 6\}$ .

Si  $D = \{\text{obtener al menos 1 gato al lanzar 2 monedas}\}$ , entonces,  $D=\{(P, G), (G, P), (G, G)\}$

### 23.4. Probabilidad

Probabilidad es un valor entre 0 y 1, que indica la posibilidad relativa de que ocurra un evento. El valor de la probabilidad se calcula mediante la siguiente fórmula:

$$P(A) = \frac{\text{número de casos favorables de } A}{\text{número total de casos de } S} = \frac{n(A)}{n(S)}$$

Ilustración 21 Probabilidad



Fuente: Matemóvil

Recuerda que...

- El valor de la probabilidad siempre se encuentra entre 0 y 1 (incluidos ambos números), es decir,  $0 \leq P(A) \leq 1$ .
- La probabilidad de que ocurra un evento imposible es 0. Por ejemplo, la probabilidad de obtener un 8 al lanzar un dado numerado del 1 al 6 es 0, es decir,  $P(X)=0$ .
- La probabilidad de que ocurra un evento seguro es 1. Por ejemplo, la probabilidad de obtener un número menor que 7 al lanzar un dado numerado del 1 al 6 es 1, es decir,  $P(X)=1$ .

Ejemplo 1:

Calcular la probabilidad de obtener un 2 al lanzar un dado.

Solución:

Vamos a utilizar la fórmula de probabilidad:

$$P(A) = \frac{\text{número de casos favorables de } A}{\text{número total de casos de } S} = \frac{n(A)}{n(S)}$$

El experimento consiste en lanzar un dado. Luego, definimos los resultados o casos del espacio muestral.

$$S = \{1, 2, 3, 4, 5, 6\}$$

Definimos nuestro evento A, como obtener un 2 al lanzar un dado. Ahora, calculamos el número de casos favorables del evento A.

$$A = \{2\}$$

Ahora, empleamos la fórmula:

$$P(A) = \frac{1}{6} = 0,1667 = 16,67\%$$

## 24. Regla del complemento en probabilidades.

La regla de complementación o regla del complemento nos dice que, si tenemos un evento A, y su suceso contrario o complementario A, entonces, se cumple que:  $P(A) = 1 - P(A)$

También podemos expresarlo como: «la probabilidad de que no ocurra A, es 1 menos la probabilidad de que ocurra A».

Ejemplo 1

La probabilidad de que el FC Barcelona gane un partido jugando de local, es de 0,83. Calcular la probabilidad de que empate o pierda un partido jugando de local.

Suceso A: ganar el próximo partido de local. Del enunciado, conocemos el valor de P(A)

$$P(A) = 0,83$$

• Suceso A: no ganar el próximo partido de local, es decir, que empate o pierda. Si calculamos P(A)

$$P(\bar{A}) = 1 - P(A)$$

$$P(\bar{A}) = 1 - 0,83 = 0,17$$

La probabilidad de que el Barcelona empate o pierda en su próximo partido de local es de 0,17.

Ejemplo 2:

Calcular la probabilidad de no obtener un 1 al lanzar un dado al aire.

• Suceso A: obtener un 1 al lanzar un dado al aire.  $A = \{1\}$ . Si calculamos P(A):

$$P(A) = \frac{\text{número de casos favorables de } A}{\text{número total de casos posibles}} = \frac{1}{6}$$

• Suceso A: no obtener un 1 al lanzar un dado al aire, es decir, obtener un 2, 3, 4, 5 o 6.  $\bar{A} = \{2, 3, 4, 5, 6\}$ . Si calculamos P(A):

$$P(\bar{A}) = 1 - P(A)$$

$$P(\bar{A}) = 1 - \frac{1}{6} = \frac{6-1}{6} = \frac{5}{6}$$

## 25. Regla de la suma o adición de probabilidades

La regla de adición o regla de la suma establece que, si tenemos un evento A y un evento B, la probabilidad

de que ocurra el evento A o el evento B se calcula de la siguiente manera:

Fórmula

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

Donde:

- P(A): probabilidad de que ocurra el evento A.
- P(B): probabilidad de que ocurra el evento B.
- P(A ∪ B): probabilidad de que ocurra el evento A o el evento B.
- P(A ∩ B): probabilidad de que ocurra el evento A y el evento B a la vez.

➤ ¿Y si los eventos son mutuamente excluyentes?

Dos eventos A y B son mutuamente excluyentes si no pueden ocurrir al mismo tiempo, es decir, si no tienen elementos comunes. Por ejemplo, sacar una carta al azar de una baraja, y obtener un 5 y un 7, son eventos mutuamente excluyentes, ya que no hay ninguna carta que tenga un 5 y un 7 al mismo tiempo. Entonces  $P(A \cap B) = 0$ , por lo tanto, partiendo de la misma

fórmula, obtendríamos la siguiente expresión:

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

$$P(A \cup B) = P(A) + P(B) - 0$$

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

Ejemplo 1:

La probabilidad de que un día cualquiera, Carlos almuerce pollo frito es de 0,4. La probabilidad de que almuerce hamburguesa es de 0,3; mientras que la probabilidad de que almuerce pollo frito y hamburguesa el mismo día es de 0,1. Calcula la probabilidad de que un día cualquiera, Carlos almuerce pollo frito o hamburguesa.

**Solución:**

Definimos nuestras probabilidades:

- Probabilidad de que Carlos almuerce pollo frito:  $P(A) = 0,4$ .
- Probabilidad de que Carlos almuerce hamburguesa:  $P(B) = 0,3$ .
- Probabilidad de que Carlos almuerce pollo frito y hamburguesa el mismo día:  $P(A \cap B) = 0,1$ .
- Probabilidad de que Carlos almuerce pollo frito o hamburguesa:  $P(A \cup B) = ?$

Ahora, aplicamos nuestra fórmula

:

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

$$P(A \cup B) = 0,4 + 0,3 - 0,1$$

$$P(A \cup B) = 0,6$$

Ejemplo 2:

La probabilidad de que, al tirar un dado, salga 1, es de  $1/6$ . La probabilidad de que salga 3, es de  $1/6$ . Calcular la probabilidad de que, al tirar un dado, salga 1 o 3.

**Solución:**

Definimos nuestros eventos:

- Probabilidad de que salga 1:  $P(A) = 1/6$ .
- Probabilidad de que salga 3:  $P(B) = 1/6$ .
- Probabilidad de que salga 1 y 3 al mismo tiempo  $P(A \cap B) = 0$ . Es cero, ya que son eventos mutuamente excluyentes. Si sale 1, ya no puede salir 3.
- Probabilidad de que salga 1 o 3:  $P(A \cup B) = ?$

26. Probabilidad condicional.

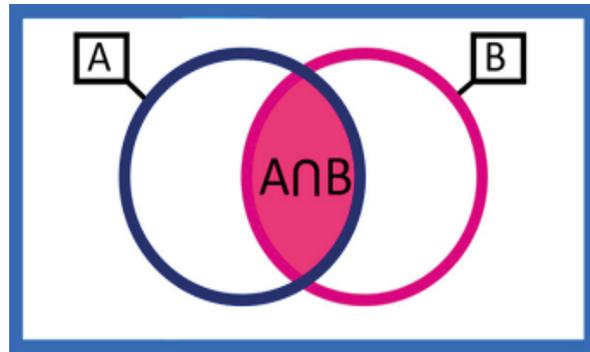
Si tenemos dos eventos, A y B, la probabilidad condicional de que ocurra el evento A, dado que ha ocurrido el evento B, se representa como  $P(A|B)$ , y se calcula de la siguiente manera:

**Fórmula:**

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

En un diagrama de Venn, veríamos los eventos A y B de la siguiente manera:

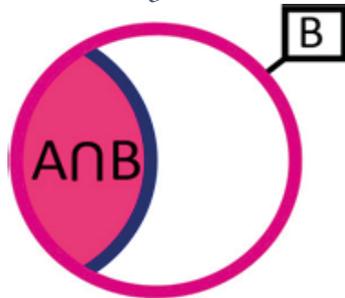
Ilustración 22 Diagrama de Venn, evento A y B



Fuente: Matemóvil

La condición, es que se ha realizado en el evento B, por lo tanto, nuestro diagrama de Venn quedaría reducido a:

Ilustración 23 Diagrama de Venn, evento B



Fuente: Matemóvil

Por ello, podemos ver que el universo está representado por la probabilidad de B, y dentro de ese universo, la probabilidad de que ocurra A, está representada por la probabilidad de  $A \cap B$ .

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

En algunos problemas, puede que sea necesario

calcular la probabilidad de que ocurra el evento B, dado que ha ocurrido A. En ese caso, simplemente invertimos el orden de las variables:

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

Ejemplo 1:

Si  $P(A) = 0,6$  ;  $P(B) = 0,4$  y  $P(A \cap B) = 0,18$ .

Calcular:

- $P(A|B)$
- $P(B|A)$

En este problema, simplemente vamos a reemplazar los datos en la fórmula.

a) Usamos la fórmula de probabilidad condicional:

$$P(A|B) = \frac{0,18}{0,4} = 0,45 = 45\%$$

b) Usamos la fórmula de probabilidad condicional, teniendo en cuenta que vamos a calcular la probabilidad de que ocurra B, dado que ha ocurrido A.

$$P(B|A) = \frac{0,18}{0,6} = 0,3 = 30\%$$

Ejemplo 2:

Al 25% de tus amigos le gusta la fresa y el chocolate, mientras que al 60% le gusta el chocolate. ¿Cuál es la probabilidad de que a un amigo que le gusta el chocolate, le guste la fresa?

**Solución:**

Vamos a trabajar con 2 eventos: que a un amigo le guste la fresa, y que a un amigo le guste el chocolate.

- Evento A: que a un amigo le gusten los fresa.  $P(A) = ?$
- Evento B: que a un amigo le guste el chocolate.  $P(B) = 60 \%$ .
- Evento A y B: que a un amigo le guste la fresa y el chocolate.  $P(A \cap B) = 25 \%$ .

Ahora calculamos la probabilidad de que a un amigo le guste la fresa, dado que le gusta el chocolate.

$$P(A|B) = \frac{25\%}{60\%} = \frac{25}{60} = \frac{5}{12} = 0,4167 = 41,67\%$$

La probabilidad de que a un amigo le guste la fresa dado que le gusta el chocolate es del 41,67 %.

**Ejemplo 3:**

El 76 % de los estudiantes de Ingeniería Civil han aprobado resistencia de materiales y el 45 % aprobaron estática. Además, el 30 % aprobaron resistencia de materiales y estática. Si Camilo aprobó resistencia de materiales, ¿qué probabilidad tiene de haber aprobado también estática?

**Solución:**

Vamos a trabajar con 2 eventos: aprobar resistencia de materiales, y aprobar estática.

- Evento A: aprobar resistencia de materiales.  $P(A) = 76 \%$ .
- Evento B: aprobar estática.  $P(B) = 45 \%$ .
- Evento A y B: aprobar resistencia de materiales y estática.  $P(A \cap B) = 30 \%$ , y es lo mismo que:  $P(B \cap A) = 30 \%$

Ahora calculamos la probabilidad de aprobar estática, dado que se aprobó resistencia de materiales.

$$P(B|A) = \frac{30\%}{76\%} = \frac{30}{76} = \frac{15}{38} = 0,3947 = 39,47\%$$

Para Camilo, la probabilidad de aprobar estática, dado que aprobó resistencia de materiales es de 39,47 %.

**27. Regla de la multiplicación o producto de probabilidades**

La regla de la multiplicación o regla del producto permite encontrar la probabilidad de que ocurra el evento A y el evento B al mismo tiempo (probabilidad conjunta). Esta regla depende de si los eventos son dependientes o independientes.

**Eventos dependientes**

Dos eventos A y B son dependientes, si la ocurrencia de uno de ellos afecta la ocurrencia del otro. Para eventos dependientes, la regla de la multiplicación establece que:

$$P(A \cap B) = P(A) \times P(B|A)$$

$$P(A \cap B) = P(B) \times P(A|B)$$

Ejemplo 1:

Una caja contiene 2 canicas azules y 3 rojas. Si se extraen dos canicas al azar sin reposición, ¿cuál es la probabilidad de que las dos sean azules?

Solución:

Dado que las canicas serán extraídas de la misma caja, y que las canicas que se extraigan, no serán devueltas a la caja (no hay reposición), entonces, se trata de eventos dependientes.

- Evento A: obtener una canica azul en la primera extracción.
- Evento B: obtener una canica azul en la segunda extracción.

Por la regla de la multiplicación, sabemos que:

$$P(A \cap B) = \frac{2}{5} \cdot \frac{1}{4} = \frac{2}{20} = 0,1 = 10 \%$$

Eventos independientes

Dos eventos A y B son independientes, si la ocurrencia de uno de ellos no afecta la ocurrencia del otro, es decir, cuando los eventos A y B no están relacionados. Para eventos independientes, la regla de la multiplicación establece que:

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

Esto se debe, a que, en los eventos independientes, la

ocurrencia de un evento no afecta a la ocurrencia del otro:

$$P(A|B) = P(A) \quad \wedge \quad P(B|A) = P(B)$$

Ejemplo 2:

En un colegio, la probabilidad de que un alumno seleccionado al azar hable inglés es de 0,20; mientras que la probabilidad de que un alumno juegue fútbol es de 0,80.

El hecho de que un alumno hable inglés no afecta en nada que juegue fútbol; por lo tanto, se trata de eventos independientes.

- Evento A: que el alumno hable inglés.  $P(A) = 0,20$ .
- Evento B: que el alumno juegue fútbol.  $P(B) = 0,80$ .

Usamos la regla de la multiplicación para eventos independientes:

$$P(A \cap B) = 0,20 \times 0,80 = 0,16 = 16 \%$$

Ejemplo 3:

Sabiendo que  $P(A) = 0,70$ ;  $P(B) = 0,50$ ; y además,  $P(A \cap B) = 0,40$ ; determinar si son eventos dependientes o independientes.

Solución:

En los eventos independientes, se cumple que:

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

En este caso:

- $P(A \cap B) = 0,40$ .
- $P(A) \times P(B) = 0,70 \times 0,50 = 0,35$ .

Podemos ver que 0,40 es diferente de 0,35; entonces:

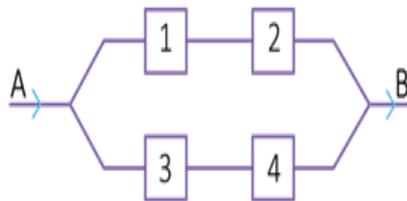
$$P(A \cap B) \neq P(A) \times P(B)$$

Podemos concluir que no son eventos independientes, es decir, **son eventos dependientes**.

#### Ejemplo 4:

Un sistema consta de 4 componentes como se ilustra en la figura.

Ilustración 24 Sistema de cuatro componentes



Fuente: Matemóvil

Todo el sistema funcionará el subsistema 1-2 funciona o si el subsistema 3-4 funciona (porque los dos subsistemas están conectados en paralelo). Como los dos componentes de cada subsistema están conectados en serie, un subsistema funcionará sólo si ambos componentes de cada subsistema funcionan. Si los

componentes funcionan o fallan de modo independiente uno de otro y si cada uno funciona con probabilidad de 0,85, ¿cuál es la probabilidad de que todo el sistema funcione (coeficiente de confiabilidad del sistema)?

Solución:

Imagina que por este circuito circula agua, esta puede ir de A hasta B por el camino de arriba 1-2 o por el camino de abajo 3-4. No necesita que los 2 caminos funcionen a la vez, con que uno funcione, será suficiente para que llegue el agua hasta el final.

Para que el sistema funcione, tiene que funcionar el subsistema 1-2 o el subsistema 3-4. Entonces, la probabilidad de que funcione el sistema será:

$$P(\text{funcione el sistema}) = P(1 \cap 2) \cup P(3 \cap 4)$$

A continuación, realizaré varios cálculos empleando la regla de la multiplicación para eventos independientes:

Calculamos primero la probabilidad de que funcionen 1 y 2:

$$P(1 \cap 2) = P(1) \times P(2) = 0,85 \times 0,85$$

$$P(1 \cap 2) = 0,7225$$

Luego la proba  $P(3 \cap 4) = 0,7225$  en 3 y 4:

$$P(3 \cap 4) = P(3) \times P(4) = 0,85 \times 0,85$$

Además, nos servirá la probabilidad de que funcione el subsistema de arriba (1 y 2) y el subsistema de abajo (3 y 4) al mismo tiempo.

$$P(1 \cap 2 \cap 3 \cap 4) = P(1) \times P(2) \times P(3) \times P(4)$$

$$P(1 \cap 2 \cap 3 \cap 4) = 0,85 \times 0,85 \times 0,85 \times 0,85$$

$$P(1 \cap 2 \cap 3 \cap 4) = 0,5220$$

Nos será de utilidad la regla de la suma:

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

Para que el sistema funcione, tiene que funcionar el subsistema 1-2 o el subsistema 3-4.

$$P(\text{funcione el sistema}) = P(1 \cap 2) \text{ o } P(3 \cap 4)$$

$$P(\text{funcione el sistema}) = P(1 \cap 2) + P(3 \cap 4) - P(1 \cap 2 \cap 3 \cap 4)$$

$$P(\text{funcione el sistema}) = 0,7225 + 0,7225 - 0,5220$$

$$P(\text{funcione el sistema}) = 0,923$$

La probabilidad de que funcione el sistema es de 0,923.

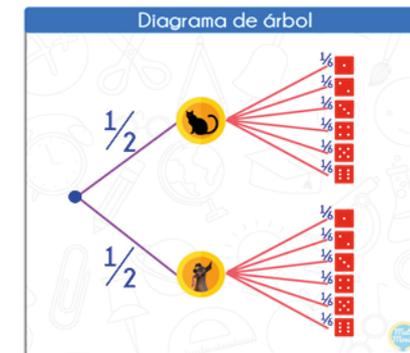
### 28. Diagrama de árbol (probabilidades)

Un diagrama de árbol es una representación gráfica de los posibles resultados de un experimento que tiene varios pasos. Nos permite calcular la probabilidad de que ocurra un evento de una manera muy sencilla.

Aquí tenemos un clásico diagrama de árbol, en el cual graficamos los posibles resultados de un experimento

que consiste en lanzar una moneda y un dado.

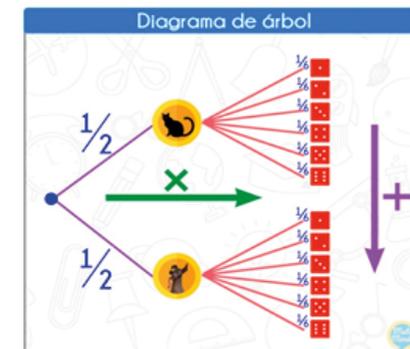
Ilustración 25 Diagrama de árbol



Fuente: Matemóvil

Para el cálculo de las probabilidades, usaremos un truco, si para calcular cierta probabilidad avanzamos hacia la derecha, entonces multiplicamos. Por otro lado, si para calcular cierta probabilidad avanzamos hacia abajo, entonces sumamos.

Ilustración 26 Métodos de cálculo en el diagrama de árbol



Fuente: Matemóvil

**Ejercicio 1:**

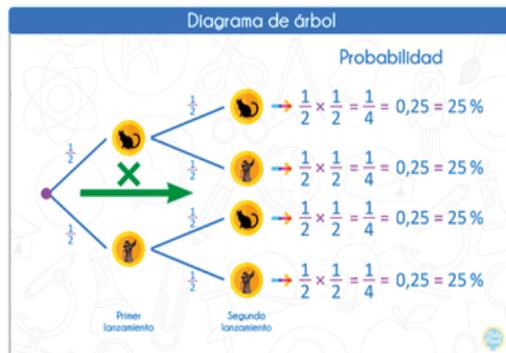
Una moneda tiene en sus caras un gato y un perro. Se se lanza 2 veces la moneda, calcular:

- La probabilidad de obtener 2 gatos.
- La probabilidad de obtener solo 1 gato.

**Solución:**

Vamos a elaborar el diagrama de árbol para este experimento. Calculamos la probabilidad para cada uno de los posibles casos, cuando avanzamos a la derecha, multiplicamos.

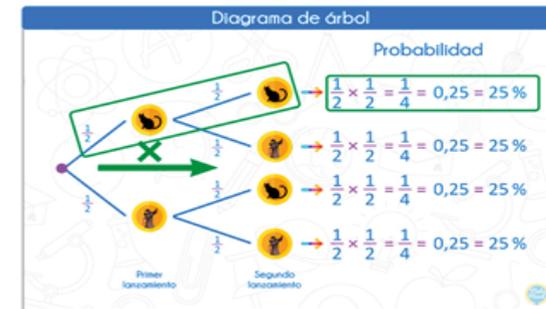
Ilustración 27 Cálculo de probabilidad mediante diagrama de árbol



Obtenido de: Matemovil.com

a) La probabilidad de obtener 2 gatos, la podemos observar en el gráfico.

Ilustración 28 Cálculo de probabilidad para obtener dos gatos



Fuente: Matemóvil

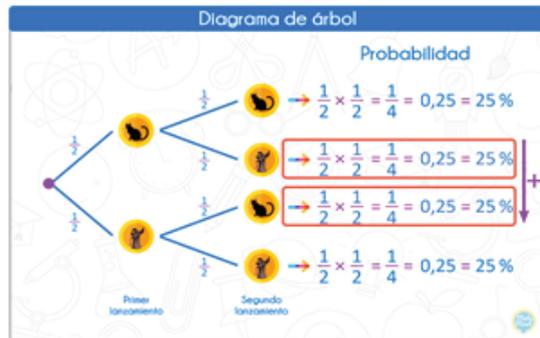
$$P(2 \text{ gatos}) = \frac{1}{4} = 0,25 = 25 \%$$

b) La probabilidad de obtener solo 1 gato, se calcula sumando 2 probabilidades, ya que hay 2 maneras de obtener solo 1 gato:

- Obtener gato y perro.
- Obtener perro y gato.

Recuerda que cuando avanzamos hacia abajo, entonces sumamos:

Ilustración 29 Método de suma en el diagrama de árbol



Fuente: Matemóvil

Por lo tanto, la probabilidad de obtener 1 solo gato será:

$$P(1 \text{ gato}) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2} = 0,5 = 50 \%$$

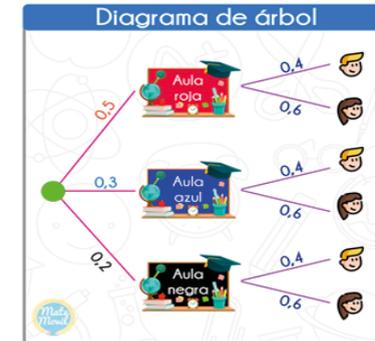
**Ejercicio 2:**

En una academia hay 3 aulas: el aula roja, el aula azul y el aula negra. El aula roja tiene al 50 % de los estudiantes de la academia, el aula azul al 30 % y el aula negra al 20 %. Además, en cada aula hay un 40 % de hombres. Si se selecciona un estudiante al azar, ¿cuál es la probabilidad de que sea un estudiante hombre del aula azul?

**Solución**

Empezamos con nuestro diagrama de árbol a partir de la información del problema.

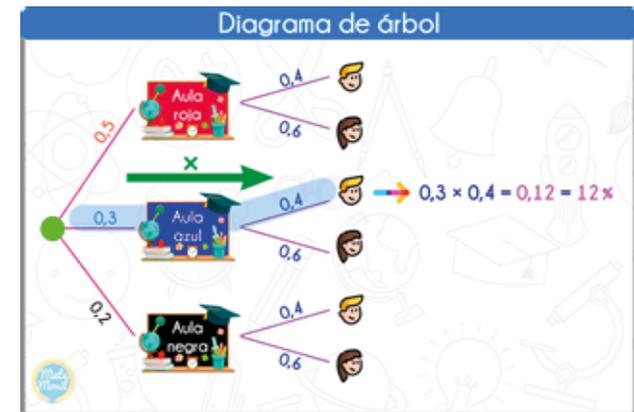
Ilustración 30 Diagrama de árbol, ejercicio 2



Fuente: Matemóvil

Ahora calculamos la probabilidad de que, si se selecciona un estudiante al azar, este sea un hombre del aula azul.

Ilustración 31 Probabilidad de estudiante al azar mediante el diagrama de árbol



Fuente: Matemóvil

La probabilidad sería de 0,12 o 12 %.

## 29. Teorema de la probabilidad total

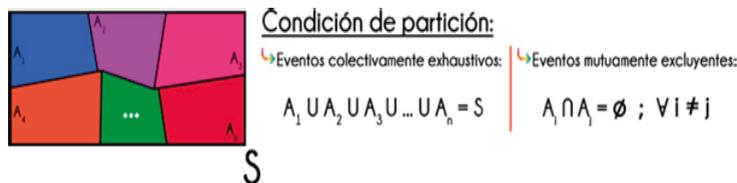
El teorema de la probabilidad total permite calcular la probabilidad de que ocurra un evento, que se puede realizar a través de varios caminos. Antes de revisar el teorema de probabilidad total, es necesario definir el concepto de «partición».

Veamos los ejemplos y ejercicios del teorema de la probabilidad total.

### 29.1. Partición

Sean  $A_1, A_2, A_3, \dots, A_n$ , eventos de un mismo espacio muestral  $S$ . Dichos eventos forman una partición de  $S$  si son mutuamente excluyentes y colectivamente exhaustivos. Es decir, si cumplen con las condiciones siguientes:

Ilustración 32 Condición de partición



Fuente: Matemóvil

### 29.2. Teorema de la probabilidad total

El teorema de la probabilidad total establece que:

Sean  $A_1, A_2, A_3, \dots, A_n$ , eventos que forman una partición del espacio muestral  $S$ , y sea  $B$  otro evento cualquiera del

espacio muestral  $S$ , entonces la probabilidad del evento  $B$  se puede obtener de la siguiente manera:

Ilustración 33 Teorema de probabilidad total

$$P(B) = P(A_1) \cdot P(B|A_1) + P(A_2) \cdot P(B|A_2) + \dots + P(A_n) \cdot P(B|A_n) = \sum_{i=1}^n P(A_i) \cdot P(B|A_i)$$

Fuente: Matemóvil

Los problemas de probabilidades que requieren al teorema de probabilidad total, también se pueden resolver de manera sencilla usando el diagrama de árbol. Vamos a revisar 1 ejercicio, y lo vamos a resolver con el diagrama de árbol y con el teorema.

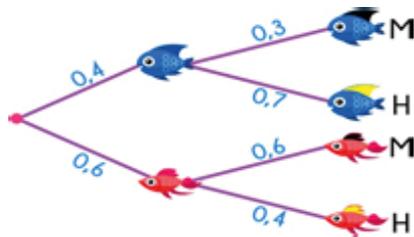
Ejercicio 1:

En un acuario se tienen solo 2 especies de peces, el 40% son de la especie azul y el 60% son de la especie roja. De la especie azul, el 30% son machos; mientras que, de la especie roja, el 40% son hembras. ¿Cuál es la probabilidad de que un pez elegido aleatoriamente en el acuario sea macho?

Solución con el diagrama de árbol:

A partir de los datos del enunciado, vamos a elaborar el diagrama de árbol.

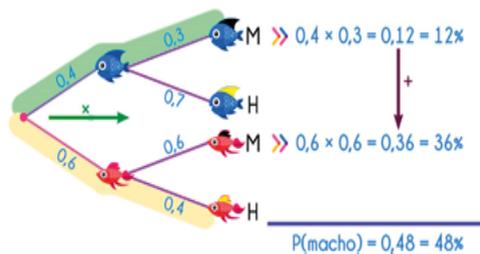
Ilustración 34 Diagrama de árbol, ejercicio 1



Fuente: Matemóvil

Recuerda el truco que usamos para calcular probabilidades usando el diagrama de árbol: cuando avanzamos de izquierda a derecha, multiplicamos las probabilidades; cuando avanzamos de arriba hacia abajo, sumamos las probabilidades.

Ilustración 35 Método de multiplicación en diagrama de árbol



Fuente: Matemóvil

La probabilidad de encontrar un macho, seleccionando un pez de forma aleatoria es de 0,48 o 48%.

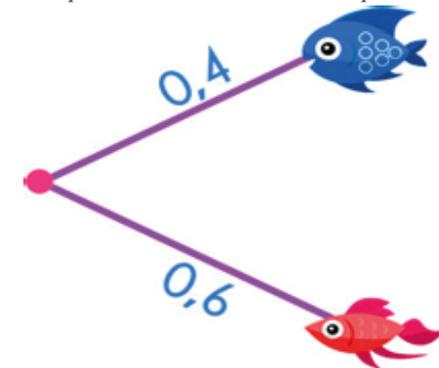
### 29.3. Solución con el teorema de probabilidad total:

Ahora resolvemos el mismo problema usando el teorema. Tenemos 2 eventos  $A_1$  y  $A_2$ , que forman una

partición del espacio muestral S (peces del acuario):

- $A_1$ : que un pez elegido aleatoriamente sea de la especie azul.
- $A_2$ : que un pez elegido aleatoriamente sea de la especie roja.

Ilustración 36 Aplicación de teorema de probabilidad total



Fuente: Matemóvil

A partir del gráfico, sabemos que:

$$P(A_1) = 0,4 \quad \wedge \quad P(A_2) = 0,6$$

También tenemos al evento B:

- B: que un pez elegido aleatoriamente sea macho.

Nos dicen que, de la especie azul, el 30% son machos. Por ello, sabemos que la probabilidad de que un pez sea macho, dado que es de la especie azul, es de:

$$P(B|A_1) = 0,3$$

Nos dicen que el 40% de los peces de la especie roja

son hembras, por ello, el 60% serán machos. Entonces, sabemos que la probabilidad de que un pez sea macho, dado que es de la especie roja:

$$P(B|A_2) = 0,6$$

En nuestro caso, tenemos una partición del espacio muestral S, formada solo por 2 eventos: A<sub>1</sub> y A<sub>2</sub>.

$$P(B) = P(A_1).P(B|A_1) + P(A_2).P(B|A_2) = \sum_{i=1}^n P(i).P(B|A_i)$$

Reemplazando nuestros valores:

$$P(B) = P(A_1).P(B|A_1) + P(A_2).P(B|A_2)$$

$$P(B) = (0,4)(0,3) + (0,6)(0,6)$$

$$P(B) = 0,12 + 0,36$$

$$P(B) = 0,48$$

$$P(B) = 0,48 = 48 \%$$

La probabilidad de que un pez elegido aleatoriamente sea macho es de 0,48 o 48 %.

### 30. Teorema de Bayes

El teorema de Bayes expresa la probabilidad de que ocurra el evento A, dado que ha ocurrido B, en función de la probabilidad de que ocurra B dado que ha ocurrido A, de la probabilidad de A y de la probabilidad de B. La fórmula del teorema de Bayes es la siguiente:

Ilustración 37 Fórmula teorema de Bayes

Fuente: Matemóvil

Como se puede apreciar, el teorema de Bayes permite calcular la probabilidad de que ocurra un evento, a partir de valores conocidos de otras probabilidades relacionadas al evento.

El teorema o regla de Bayes fue planteado por el matemático y religioso inglés Thomas Bayes. Este teorema fue publicado en el año 1763, dos años después de la muerte de Bayes.

Este teorema lo encontramos de dos formas diferentes, en su forma simple y en su forma extendida, las cuales revisaremos a continuación, además de muchos ejemplos y ejercicios.

#### 30.1. Forma simple del teorema de Bayes:

La forma simple de este teorema es la que mencionamos al inicio del artículo y es la siguiente:

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

Donde:

- A y B son eventos, y además: P(B) ≠ 0.

- $P(A|B)$ : es la probabilidad de que ocurra A, dado que ha ocurrido B.
- $P(B|A)$ : es la probabilidad de que ocurra B, dado que ha ocurrido A.
- $P(A)$ : es la probabilidad de que ocurra A.
- $P(B)$ : es la probabilidad de que ocurra B.

El teorema de Bayes expresa la probabilidad de que ocurra el evento A, dado que ha ocurrido B, en función de la probabilidad de que ocurra B dado que ha ocurrido A, de la probabilidad de A y de la probabilidad de B.

En la práctica tiene muchísimas aplicaciones, por ejemplo, conociendo la probabilidad de que una persona tenga fiebre dado que tiene gripe, nos permite calcular la probabilidad de que una persona que tiene gripe, dado que tiene fiebre. Tiene, además, aplicaciones importantísimas en la detección del cáncer y otras enfermedades.

### Ejemplo 1:

En la academia de Matemóvil, la probabilidad de que a un alumno seleccionado al azar le guste el helado es del 60 %, mientras que la probabilidad de que a un alumno le guste la torta es del 36 %. Además, se sabe que la probabilidad de que a un alumno le guste la torta dado que le gusta el helado es del 40 %. Calcular la probabilidad de que a un alumno le guste el helado, dado que le gusta la torta.

### Solución:

Primero definimos los 2 eventos con los que vamos a trabajar:

- h: que a un alumno le guste el helado.
- t: que a un alumno le guste la torta.

Tenemos los siguientes datos:

- $P(h) = 0,6$ .
- $P(t) = 0,36$ .
- $P(t|h) = 0,4$ .

Nos piden calcular  $P(h|t)$ .

Aplicamos el teorema de Bayes:

$$P(h|t) = \frac{P(h) \cdot P(t|h)}{P(t)}$$

$$P(h|t) = \frac{0,6 \cdot 0,4}{0,36} = \frac{0,24}{0,36} = \frac{24}{36} = \frac{2}{3} = 0,6667 = 66,67 \%$$

Entonces, la probabilidad de que un alumno le guste el helado dado que le gusta la torta es de 0,6667 o 66,67 %.

### 30.2. Forma extendida del teorema de Bayes:

Ilustración 38 Fórmula extendida del teorema de Bayes

**Teorema de Bayes**

Sean  $A_1, A_2, A_3, \dots, A_n$  eventos mutuamente excluyentes y exhaustivos, con  $P(A_i) \neq 0$  para cada  $A_i$ . Sea  $B$  cualquier evento con  $P(B) \neq 0$ , entonces:

$$P(A_i|B) = \frac{P(A_i) \cdot P(B|A_i)}{\sum_{j=1}^n P(A_j) \cdot P(B|A_j)}$$


*Fuente: Matemóvil*

Esta forma extendida es la que encontrarás en la mayoría de los libros de estadística. Emplea las particiones del espacio muestral.



## CAPÍTULO VI

### VARIABLES ALEATORIAS Y DISTRIBUCIONES DE PROBABILIDAD

### 31. Variables aleatorias discretas y continuas

Una **variable aleatoria** es una función que asigna un valor numérico, al resultado de un experimento aleatorio. Una variable aleatoria puede ser discreta o continua. **Las variables aleatorias discretas** son aquellas que presentan un número contable de valores; por ejemplo, el número de personas que viven en una casa (pueden ser 3, 5 o 9). **Las variables aleatorias continuas** son aquellas que presentan un número incontable de valores; por ejemplo, el peso de las vacas en una granja (una vaca puede pesar 632,12 kg, otra puede pesar 583,12312 kg, otro 253,12012 kg, otro 198,0876 kg y nunca terminaríamos de enumerar todos los posibles valores).

Como estas definiciones son muy difíciles de entender a simple vista, vamos a explicarlas a detalle.

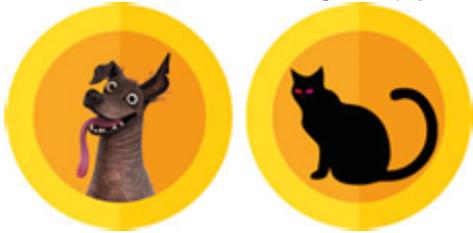
#### 31.1. Variable aleatoria

Una variable aleatoria es una función que asigna un valor numérico, al resultado de un experimento aleatorio. Recordemos que el resultado de un experimento aleatorio depende del azar.

Ejemplo.

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

Ilustración 39 Moneda de perro y gato



Fuente: Matemóvil

Vamos a realizar un experimento aleatorio que consiste en lanzar 2 monedas. Colocaremos los resultados en el siguiente gráfico:

Ilustración 40 Resultados de experimentos aleatorios de dos monedas

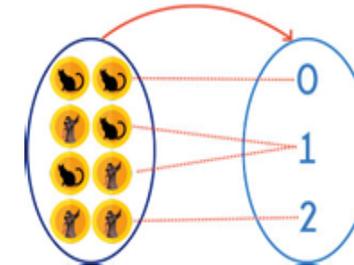


Fuente: Matemóvil

Definimos nuestra variable aleatoria X:

- X = número de perros.

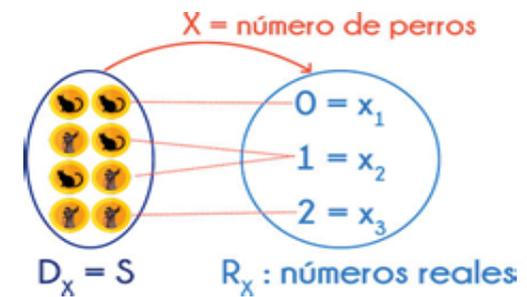
Ilustración 41 Variable aleatoria X  
X = número de perros



Fuente: Matemóvil

Ten en cuenta que la variable aleatoria siempre va con letras mayúsculas (en este caso X), mientras que los valores de su rango siempre con letras minúsculas (en este caso  $x_1, x_2, x_3$ ).

Ilustración 42 Variables aleatorias y valores de rango



Fuente: Matemóvil

Los valores del rango de esta variable aleatoria son:

$$x_1 = 0$$

$$x_2 = 1$$

$$x_3 = 2$$

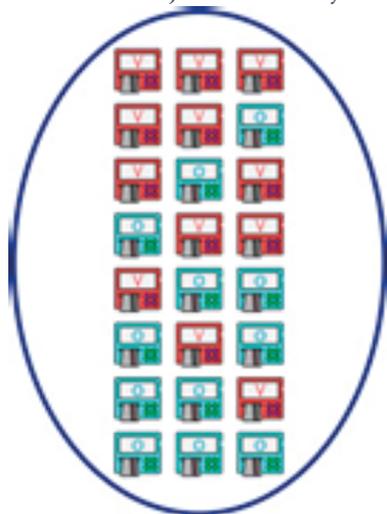
En el dominio de la función tenemos el **espacio muestral**, es decir, todos los resultados posibles de nuestro experimento aleatorio. Mientras que el rango tenemos un conjunto de **números reales**.

Ejemplo 2 de variable aleatoria

En un banco hay 3 cajeros automáticos. Vamos a realizar un experimento aleatorio que consiste en ir al banco a una hora al azar del día y ver qué cajeros están ocupados y qué cajeros están vacíos.

Colocamos en el siguiente gráfico los resultados, los cajeros vacíos (V) irán de color rojo y los ocupados (O) de color verde.

Ilustración 43 Cajeros vacíos y ocupados

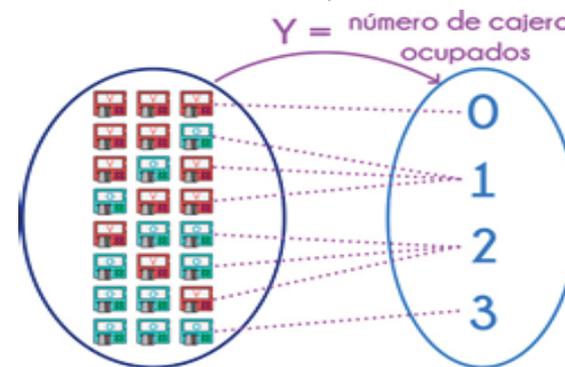


Obtenido de: Matemovil.com

Definimos nuestra variable aleatoria Y:

- Y = número de cajeros automáticos ocupados.

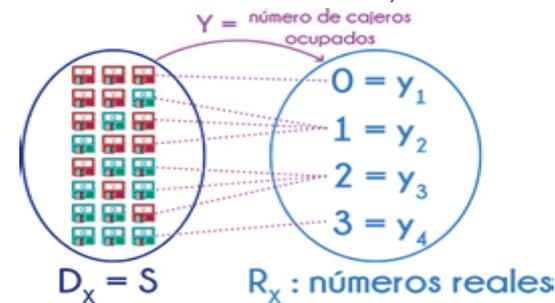
Ilustración 44 Números de cajeros automáticos ocupados



Fuente: Matemóvil

Ten en cuenta que **la variable aleatoria siempre va con letras mayúsculas** (en este caso Y), mientras que **los valores de su rango siempre con letras minúsculas** (en este caso  $y_1, y_2, y_3, y_4$ ).

Ilustración 45 Variables aleatorias y valores de rango



Fuente: Matemóvil

Los valores del rango de esta variable aleatoria son:

- $y_1 = 0$
- $y_2 = 1$
- $y_3 = 2$
- $y_4 = 3$

Las variables aleatorias se clasifican en **discretas** o **continuas** en función de los valores numéricos que asumen. Veamos esto a detalle.

### 31.2. Variable aleatoria discreta

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

Por ejemplo, si realizamos el experimento de salir a calle y seleccionar 10 personas al azar para un examen sorpresa de matemáticas, podemos definir la variable aleatoria A:

A = número de personas que aprobaron el examen.

Los valores que asume A (en su rango), van del 0 al 10 (0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10). El rango lo expresaríamos de la siguiente manera:

$$\text{➤ } R_A = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$$

La variable aleatoria A asume un número contable de valores, por ello, es una variable aleatoria discreta.

Otro ejemplo, vamos a realizar el experimento de registrar los automóviles a una caseta de peaje. Podemos

definir la variable aleatoria V:

V = número de vehículos que llegan durante el periodo de un día.

Los valores que asume V (en su rango), son 0, 1, 2, 3, 4, 5, ...; así sean muchos vehículos los que llegan, siempre podremos contar la cantidad de valores que asume V. Por ello, la variable V es una variable aleatoria discreta.

### 31.3. Variable aleatoria continua

Una variable aleatoria continua, es aquella que puede asumir un número incontable de valores.

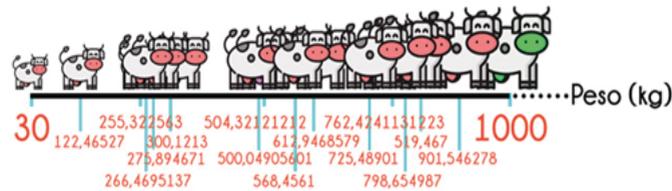
Por ejemplo, si realizamos el experimento de ir a mi granja y estudiamos las características de las vaquitas, podemos definir la variable aleatoria C:

B = peso de una vaca en la granja de Jorge (en kilogramos).

Alguna vaca puede pesar 425,1872 kg; otra puede pesar 612,5874541 kg; otra puede pesar 545,897512121 kg. Si tomamos más vacas, podríamos tener más valores y nunca terminaríamos.

Se conoce que el becerro más pequeño tiene un peso de 30 kg, y la vaca más grande tiene un peso de 1000 kg.

Ilustración 46 Peso de diferentes vacas



» B = peso de una vaca en la granja de Jorge (en kg).

↳  $R_b : 30 \leq b \leq 1000$

Fuente: Matemóvil

Y así, tendríamos un número incontable de valores para el rango de esta variable. El rango de esta variable puede ser cualquier valor dentro del intervalo que va desde 30 kg hasta 1000 kg.

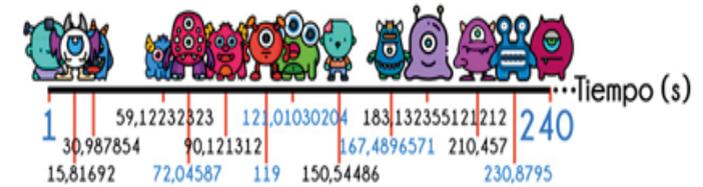
Por ello, se trata de una *variable aleatoria continua*.

Otro ejemplo, si vamos a una agencia del banco y registramos los datos de atención a los clientes, podemos definir la variable aleatoria D:

- D = tiempo de atención a los clientes del banco (en segundos).

Un cliente puede ser atendido en 24,123 s; otro cliente en 72,32142 s; otro en 51,123123 s. Si seguimos tomando más clientes, tendríamos más valores. Se conoce además que el tiempo mínimo de atención en ventanilla es de 1 s y el tiempo máximo es de 240 s.

Ilustración 47 Datos de atención al cliente



» D = tiempo de atención en ventanilla (en s).

↳  $R_d : 1 \leq d \leq 240$

Fuente: Matemóvil

Y así, tendríamos un número incontable de valores para el rango de esta variable. El rango de esta variable puede ser cualquier valor dentro del intervalo que va desde 1 s hasta 240 s. Por ello, se trata de una variable aleatoria continua.

*En general, las variables aleatorias discretas representan datos que provienen del conteo del número de elementos, mientras que, las variables aleatorias continuas representan datos que provienen de mediciones, por ejemplo, tiempo, peso, longitud, etc.*

### 32. Función de probabilidad de una variable aleatoria discreta

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

➤ Introducción

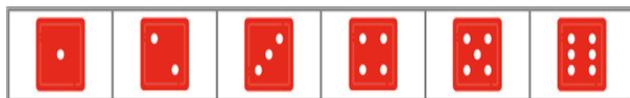
La fórmula que usaremos para la función de probabilidad de una variable aleatoria discreta es la siguiente:

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

Como suena complicado, mejor lo aprendemos con algunos ejemplos.

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

Ilustración 48 Posibles resultados al lanzar un dado



Fuente: Matemóvil

Definimos la variable aleatoria discreta X:

X = resultado de lanzar un dado.

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

Ilustración 49 Valores de rango

x	1	2	3	4	5	6

Fuente: Matemóvil

Vamos a asociar cada valor con una probabilidad. En un dado, las probabilidades son bastante sencillas de

calcular usando la fórmula de probabilidad:

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

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

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

⋮

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

Ahora colocamos las probabilidades en la tablita:

Ilustración 50 Probabilidades

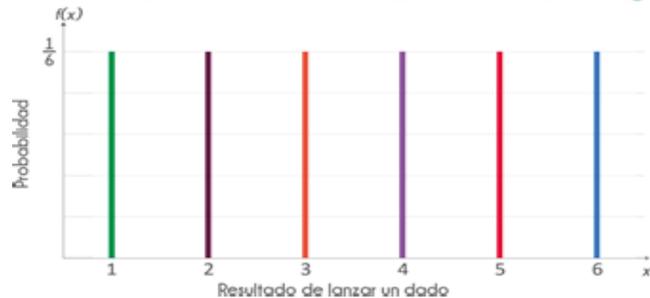
x	1	2	3	4	5	6
f(x)	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Fuente: Matemóvil

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

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

Ilustración 51 Función de probabilidad  
Gráfica de función de probabilidad



Obtenido de: Matemovil.com

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

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

Ilustración 52 Función de la probabilidad

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

Fuente: Matemóvil

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

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

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

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

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

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

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

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

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

$$\sum_{\text{toda } x} f(x) = 1$$

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

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

Ejercicios

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

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

Solución:

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

Empezamos colocando una tabla con los valores de  $x$ :

Tabla 41 Valores  $X$

$x$	1	2	3	4
-----	---	---	---	---

Fuente: Matemóvil

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

Tabla 42 Valores de  $f(x)$

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

Fuente: Matemóvil

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

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

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

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

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

Veamos si se cumple realizando la sumatoria:

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

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

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

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

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

Tabla 43 Simplificación de fracciones

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

Fuente: Matemóvil

Colocamos la misma tabla, pero ya ordenada y pasaremos las fracciones a decimales:

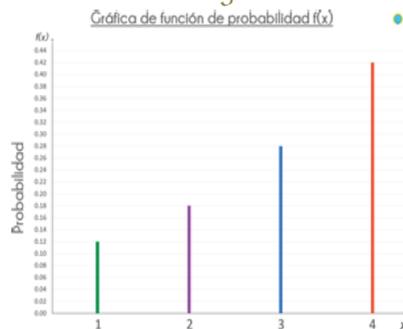
Tabla 44 Transformación de fracciones a decimales

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

Fuente: Matemóvil

Y terminamos con el diagrama de barras de la función:

Ilustración 53 Diagrama de barras f(x)



Fuente: Matemóvil

### 33. Función de distribución acumulativa de una variable aleatoria discreta

La función de distribución acumulativa especifica la probabilidad de que una variable aleatoria sea menor o igual a un valor dado. La función de distribución acumulativa de la variable aleatoria  $X$  es la función  $F(x) = P(X \leq x)$ . Para que se entienda mejor, vamos a revisar algunos ejemplos y ejercicios.

#### Introducción

La función de probabilidad de una variable aleatoria discreta que estudiamos en un capítulo anterior, nos indica la probabilidad de que la variable aleatoria sea igual a un valor determinado. Pero la función de distribución acumulativa especifica la probabilidad de que una variable aleatoria sea menor o igual que un valor dado. A la función de distribución acumulativa la denominamos  $F(x)$  y su fórmula es la siguiente:

$$F(x) = P(X \leq x)$$

Vamos a mencionar 2 propiedades importantes:

- Los valores de  $F(x)$  se encuentran siempre en este intervalo:  $0 \leq F(x) \leq 1$ .
- $F(x)$  no es una función decreciente.

Además, para la función de distribución acumulativa de la variable aleatoria discreta, se cumple que:

$$P(a < X \leq b) = F(b) - F(a)$$

Ejemplo 1:

Una tienda vende discos duros de 1 TB, 2 TB y 3 TB de capacidad. Encontrar la función de distribución de acumulativa de  $X$ , sabiendo que  $X$  = la capacidad de memoria en un disco duro comprado:

Tabla 45 Datos de la tienda

$x$	$f(x)$
1	0,5
2	0,3
3	0,2

Fuente: Matemóvil

Solución:

Nos piden encontrar la función de distribución acumulativa de  $X$ , entonces recordemos la fórmula:

$$F(x) = P(X \leq x)$$

Primero vamos a encontrar  $F(x)$  para cada uno de los posibles valores de  $X$ :

- $F(1) = P(X \leq 1) = f(1) = 0,5$
- $F(2) = P(X \leq 2) = P(X = 1 \text{ o } 2) = f(1) + f(2) = 0,5 + 0,3 = 0,8$
- $F(3) = P(X \leq 3) = P(X = 1, 2 \text{ o } 3) = f(1) + f(2) + f(3) = 0,5 + 0,3 + 0,2 = 1$

Completamos la tabla:

Tabla 46  $F(x)$  y valores de  $x$

$x$	$f(x)$	$F(x)$
1	0,5	0,5
2	0,3	0,8
3	0,2	1
$\Sigma$	1	

Fuente: Matemóvil

Vamos a borrar la columna  $f(x)$  para no generar confusión:

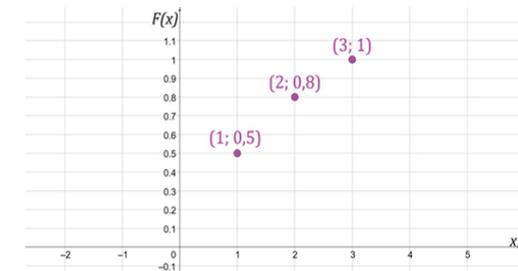
Tabla 47  $f(x)$

$x$	$F(x)$
1	0,5
2	0,8
3	1

Fuente: Matemóvil

A continuación, elaboramos la gráfica de  $x$  vs  $F(x)$ :

Ilustración 54 Representación gráfica de  $X$  vs  $F(x)$

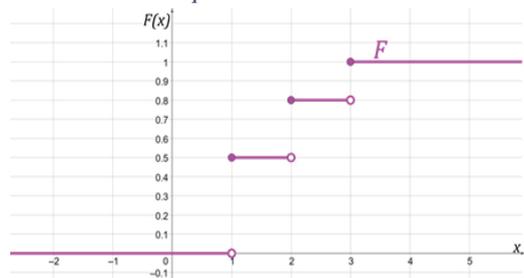


Fuente: Matemóvil

En la gráfica, colocamos líneas continuas entre los

valores, partiendo desde el lado izquierdo:

Ilustración 55 Representación en líneas continuas



Fuente: Matemóvil

Y listo, solo nos queda definir la función  $F(x)$  a partir de la gráfica:

$$F(x) = \begin{cases} 0; & x < 1 \\ 0,5; & 1 \leq x < 2 \\ 0,8; & 2 \leq x < 3 \\ 1; & x \geq 3 \end{cases}$$

Además, pudimos comprobar que los valores de  $F(x)$  se encuentran entre 0 y 1, y también, que  $F(x)$  no es una función decreciente. Estas son 2 propiedades importantes de la función de probabilidad.

Ejemplo 2:

Una moneda tiene en sus caras, un gato y un perro. Encontrar la función de distribución acumulativa del número de perros obtenidos al lanzar 2 monedas y graficar la función.

Ilustración 56 Moneda de dos caras (Perro, gato)



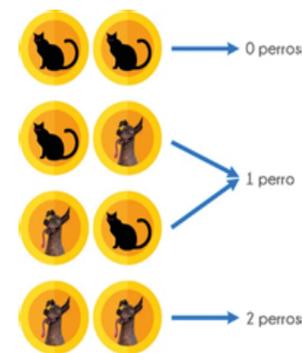
Fuente: Matemóvil

Solución:

Arrancamos definiendo la variable aleatoria discreta con la que vamos a trabajar:

- $X$  = número de perros obtenidos al lanzar 2 monedas.
- Si lanzamos dos monedas, podemos obtener los siguientes resultados: gato con gato, gato con perro, perro con gato o, por último, perro con perro.

Ilustración 57 Resultado del lanzamiento de dos monedas



Obtenido de: Matemovil.com

A continuación, colocamos en una tablita vertical, los valores de la variable aleatoria  $X$  (número de perros

obtenidos al lanzar 2 monedas):

Tabla 48 Valores de la variable X

x
0
1
2

Fuente: Matemóvil

A cada valor de la variable aleatoria, le asignamos una probabilidad:

$$\text{Probabilidad (0 perros)} = \frac{\text{número de casos favorables}}{\text{número total de casos posibles}} = \frac{1}{4} = 0,25$$

$$\text{Probabilidad (1 perro)} = \frac{\text{número de casos favorables}}{\text{número total de casos posibles}} = \frac{2}{4} = 0,50$$

$$\text{Probabilidad (2 perros)} = \frac{\text{número de casos favorables}}{\text{número total de casos posibles}} = \frac{1}{4} = 0,25$$

Y completamos la tablita que venimos construyendo, agregando la columna de la función de probabilidad:

Tabla 49 Función de probabilidad

x	f(x)
0	0,25
1	0,50
2	0,25

Fuente: Matemóvil

Ahora sí, vamos con la función de distribución acumulativa, recordando su fórmula:

$$F(x) = P(X \leq x)$$

i) Calculamos los valores de  $F(x)$ , empezando con  $F(0)$ .

$$F(0) = P(X \leq 0) = 0,25$$

La probabilidad de obtener 0 perros o menos al lanzar 2 monedas, es 0,25.

ii) Continuamos con  $F(1)$ :

$$F(1) = P(X \leq 1) = 0,25 + 0,50 = 0,75$$

La probabilidad de obtener 1 perro o menos al lanzar 2 monedas, es 0,75.

iii) Seguimos con  $F(2)$ :

$$F(2) = P(X \leq 2) = 0,25 + 0,50 + 0,25 = 1$$

La probabilidad de obtener 2 perros o menos al lanzar 2 monedas, es 1.

Completamos nuestra tablita:

Tabla 50 Resultados

x	f(x)	F(x)
0	0,25	0,25
1	0,50	0,75
2	0,25	1

Fuente: Matemóvil

Puedes ver en la tabla como los valores de  $f(x)$  se van acumulando para obtener los valores de  $F(x)$ .

Agradecemos a  $f(x)$  por su participación, pero ya no la necesitamos. Nos quedamos con  $x$  y  $F(x)$ :

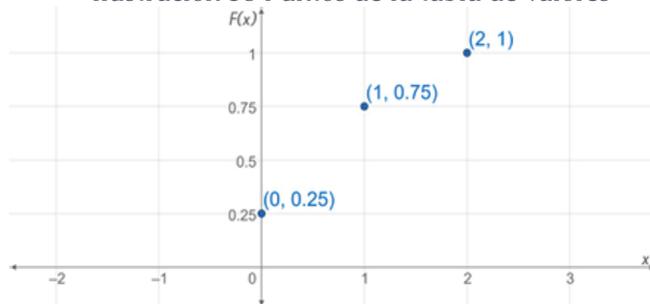
Tabla 51  $X$  y  $F(x)$

$x$	$F(x)$
0	0,25
1	0,75
2	1

Fuente: Matemóvil

Y empezamos nuestra gráfica de  $F(x)$ , marcando los puntos de la tabla:

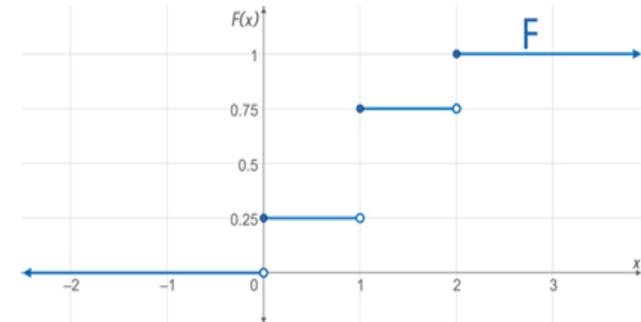
Ilustración 58 Puntos de la tabla de valores



Fuente: Matemóvil

En la gráfica, colocamos líneas continuas entre los valores, partiendo desde el lado izquierdo:

Ilustración 59 Representación en líneas continuas



Y ahora sí, solo nos queda pasar la información de la gráfica a la definición de  $F(x)$ :

$$F(x) = \begin{cases} 0; & x < 0 \\ 0,25; & 0 \leq x < 1 \\ 0,75; & 1 \leq x < 2 \\ 1; & x \geq 2 \end{cases}$$

Ejemplo 3:

Sea  $X$  el número de defectos importantes en un auto seleccionado al azar de cierta marca. La función de distribución acumulativa de  $X$  es la siguiente:

$$F(x) = \begin{cases} 0; & x < 1 \\ 0,06; & 1 \leq x < 2 \\ 0,46; & 2 \leq x < 3 \\ 0,96; & 3 \leq x < 4 \\ 1; & x \geq 4 \end{cases}$$

Calcule las siguientes probabilidades:

- a)  $f(2)$
- b)  $P(X \leq 3)$
- c)  $P(1 < X \leq 3)$
- d)  $P(1 < X \leq 4)$   $P(X \geq 2,35)$

**Solución:**

A partir de la función de distribución acumulativa, vamos a elaborar la tabla de la función de probabilidad. Empezamos colocando una tabla de 3 columnas. Los valores de X los obtenemos de los extremos de los intervalos que aparecen en la función de distribución acumulativa. No tomes en cuenta la primera línea de la función de distribución acumulativa.

Ilustración 60 Función de probabilidad

$$F(x) = \begin{cases} 0; & x < 1 \\ 0,06; & 1 \leq x < 2 \\ 0,46; & 2 \leq x < 3 \\ 0,96; & 3 \leq x < 4 \\ 1; & x \geq 4 \end{cases}$$

x	f(x)	F(x)
1		0,06
2		0,46
3		0,96
4		1

Fuente: Matemóvil

A continuación, encontraremos los valores de  $f(x)$  a partir de los valores de la función de distribución acumulativa  $F(x)$ . Esto se parece mucho a lo que hicimos en el capítulo de tablas de frecuencia. Arrancamos colocando el primer valor de  $F(x)$  en la columna de  $f(x)$ .

Ilustración 61 Valores de  $f(x)$

x	f(x)	F(x)
1	0,06	0,06
2		0,46
3		0,96
4		1

Fuente: Matemóvil

Después calculamos el resto de los valores, ten en cuenta que los valores de  $f(x)$  se van acumulando para

dar paso a los valores de  $F(x)$  y los agregamos a la tabla:

Ilustración 62 Valores de  $F(x)$

$$f(2) = 0,46 - 0,06 = 0,40$$

$$f(3) = 0,96 - 0,46 = 0,50$$

$$f(4) = 1 - 0,96 = 0,04$$

x	f(x)	F(x)
1	0,06	0,06
2	0,40	0,46
3	0,50	0,96
4	0,04	1

Fuente: Matemóvil

Calculamos lo que pide el problema:

a)  $f(2)$

Lo obtenemos a partir de la tabla:

$$f(2) = 0,40 = 0,4$$

b)  $P(X \leq 3)$

Partimos de la fórmula de la función de distribución acumulativa, reemplazando x por 3.

$$F(x) = P(X \leq x)$$

$$F(3) = P(X \leq 3)$$

Le damos la vuelta a la ecuación:

$$P(X \leq 3) = F(3)$$

A partir de la tabla:

$$P(X \leq 3) = F(3) = 0,96$$

c)  $P(1 < X \leq 3)$

En este apartado, usaremos la siguiente propiedad:

$$P(a < x \leq b) = F(b) - F(a)$$

$$P(1 < x \leq 3) = F(3) - F(1)$$

$$P(1 < x \leq 3) = 0,96 - 0,06$$

$$P(1 < x \leq 3) = 0,90 = \mathbf{0,9}$$

d)  $P(1 < X \leq 4)$

Usamos la misma propiedad:

$$P(a < x \leq b) = F(b) - F(a)$$

$$P(1 < x \leq 4) = F(4) - F(1)$$

$$P(1 < x \leq 4) = 1 - 0,06$$

$$P(1 < x \leq 4) = \mathbf{0,94}$$

e)  $P(X \geq 2,35)$

A partir de la tabla, sabemos que la variable aleatoria  $X$  puede tomar los valores: 1, 2, 3 o 4. Si  $X$  tiene que ser mayor a 2,35; puede ser 3 o 4.

$$P(X \geq 2,35) = P(X = 3 \text{ o } 4) = P(X = 3) + P(X = 4)$$

$$P(X \geq 2,35) = f(3) + f(4)$$

$$P(X \geq 2,35) = 0,50 + 0,04$$

$$P(X \geq 2,35) = \mathbf{0,54}$$

### 34. Media (valor esperado), varianza y desviación estándar de una variable aleatoria discreta

El día de hoy veremos cómo calcular la media (esperanza o valor esperado), la varianza y la desviación

estándar de una variable aleatoria discreta a partir de su función o distribución de probabilidad.

#### 34.1. Valor esperado

La media « $\mu$ » o valor esperado « $E(X)$ » es un promedio ponderado de los valores que asume la variable aleatoria cuando los pesos son las probabilidades. Es una medida de tendencia central. Cuando trabajamos con una variable aleatoria discreta, la media o valor esperado se calcula mediante la siguiente fórmula:

$$\mu = E(X) = x_1 \cdot f(x_1) + x_2 \cdot f(x_2) + \dots + x_n \cdot f(x_n) = \sum x \cdot f(x)$$

Como verás, la media  $\mu$  de una variable aleatoria discreta  $X$  se encuentra al multiplicar cada posible valor de  $X$  por su propia probabilidad y luego sumar todos los productos.

El valor esperado o media no tiene que ser un valor que la variable aleatoria pueda asumir.

#### 34.2. Varianza

La varianza,  $\sigma^2$  o  $V(X)$ , es un promedio ponderado de las desviaciones al cuadrado de una variable aleatoria de su media. Los pesos son las probabilidades. La varianza  $\sigma^2$  o  $V(X)$  se calcula con la siguiente fórmula:

$$\sigma^2 = V(X) = \sum (x - \mu)^2 \cdot f(x)$$

Pero es más fácil y rápido usar esta fórmula, equivalente a la anterior:

$$\sigma^2 = V(X) = \sum (x - \mu)^2 \cdot f(x)$$

Debemos recordar que la varianza es una medida de variabilidad o dispersión.

### 34.3. Desviación estándar

La desviación estándar  $\sigma$  es la raíz cuadrada positiva de la varianza:

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

Recuerda que la varianza se mide en las mismas unidades que la variable aleatoria original.

#### Ejemplo 1

Sea X el número de clientes que visitan una tienda por día. Calcular el valor esperado de X a partir de su función de probabilidad:

Tabla 52 Cálculo de valor esperado

$x$	0	1	2	3
$f(x)$	0,60	0,25	0,10	0,05

Fuente: Matemóvil

Solución:

Recordemos la fórmula de la media o valor esperado:

$$\mu = E(X) = x_1 \cdot f(x_1) + x_2 \cdot f(x_2) + \dots + x_n \cdot f(x_n) = \sum x \cdot f(x)$$

Entonces reemplazamos los valores:

$$\mu = E(X) = x_1 \cdot f(x_1) + x_2 \cdot f(x_2) + x_3 \cdot f(x_3) + x_4 \cdot f(x_4)$$

$$\mu = E(X) = 0 \cdot 0,60 + 1 \cdot 0,25 + 2 \cdot 0,10 + 3 \cdot 0,05$$

$$\mu = E(X) = 0 + 0,25 + 0,20 + 0,15$$

$$\mu = E(X) = \mathbf{0,60}$$

Y listo, la media o valor esperado es de 0,60. Recuerda que el valor esperado o media no tiene que ser un valor que la variable aleatoria pueda asumir.

#### Ejemplo 2

Encuentra la media, varianza y desviación estándar de la variable aleatoria X:

Tabla 53 Media, varianza y desviación estándar

$x$	0	1	2	3
$f(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Fuente: Matemóvil

Solución

Cuando no solo nos piden calcular la media, sino también la varianza y la desviación estándar, es mejor emplear una tabla en la cual iremos colocando todo lo que indican las fórmulas. Empezamos colocando la tabla de la función de manera vertical.

Tabla 54 Función  $f(x)$

$x$	$f(x)$
0	$\frac{1}{8}$
1	$\frac{3}{8}$
2	$\frac{3}{8}$
3	$\frac{1}{8}$

Fuente: Matemóvil

Para calcular la media  $\mu$  o valor esperado  $E(X)$ , usamos su fórmula:

$$\mu = E(X) = \sum x \cdot f(x)$$

Agregamos una columna más a la tabla, para encontrar los productos  $x \cdot f(x)$ . Luego, sumamos estos productos.

Tabla 55 Producto de  $X \cdot f(x)$

$x$	$f(x)$	$x \cdot f(x)$
0	$\frac{1}{8}$	0
1	$\frac{3}{8}$	$\frac{3}{8}$
2	$\frac{3}{8}$	$\frac{6}{8}$
3	$\frac{1}{8}$	$\frac{3}{8}$
$\Sigma$		$\frac{12}{8} = 1,5$

Fuente: Matemóvil

Listo el valor de la media, obtuvimos que:

$$\mu = E(X) = \sum x \cdot f(x) = \frac{12}{8} = \underline{1,5}$$

Llega el turno de la varianza, usaremos la fórmula rápida:

$$\sigma^2 = V(X) = \left[ \sum x^2 \cdot f(x) \right] - \mu^2$$

Entonces, agregamos algunas columnas más a nuestra tabla para encontrar lo que nos pide la fórmula:

Tabla 56  $X^2 \cdot f(x)$

$x$	$f(x)$	$x^2$	$x^2 \cdot f(x)$
0	$\frac{1}{8}$	0	0
1	$\frac{3}{8}$	1	$\frac{3}{8}$
2	$\frac{3}{8}$	4	$\frac{12}{8}$
3	$\frac{1}{8}$	9	$\frac{9}{8}$
$\Sigma$			$\frac{24}{8} = 3$

Ya casi terminamos con la varianza:

$$\sigma^2 = V(X) = \left[ \sum x^2 \cdot f(x) \right] - \mu^2$$

$$\sigma^2 = V(X) = 3 - 1,5^2 = 3 - 2,25 = \underline{0,75}$$

Y, por último, viene el valor de la desviación estándar, que es la raíz cuadrada positiva de la varianza.

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

$$\sigma = \sqrt{0,75}$$

$$\sigma = 0,866$$

### 35. Función de densidad de probabilidad con variable continua

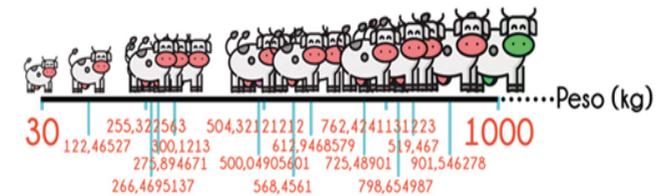
La función de densidad de probabilidad de una variable aleatoria continua es una fórmula que ayuda a calcular la probabilidad de que una variable aleatoria continua tenga un valor que esté dentro de un intervalo específico. Como todavía no se entiende mucho, veamos los detalles y los ejercicios resueltos.

#### 35.1. Variable aleatoria continua

Una variable aleatoria continua, es aquella que puede asumir un número incontable de valores. Por ejemplo, si realizamos el experimento de ir a una granja y estudiamos las características de las vaquitas, podemos definir la variable aleatoria  $B =$  peso de una vaca en la granja de Jorge (en kilogramos).

Alguna vaquita puede pesar 425,1872 kg; otra puede pesar 612,5874541 kg; otra puede pesar 545,897512121 kg. Si tomamos más vacas, podríamos tener más valores y nunca terminaríamos. Se conoce que el becerro más pequeño tiene un peso de 30 kg, y la vaca más grande tiene un peso de 1000 kg.

Ilustración 63 Distintos pesos en Kg de vacas



»  $B =$  peso de una vaca en la granja de Jorge (en kg).

$$R_B : 30 \leq b \leq 1000$$

Fuente: Matemóvil

Y así, tendríamos un número incontable de valores para el rango de esta variable. El rango de esta variable puede ser cualquier valor dentro del intervalo que va desde 30 kg hasta 1000 kg, por ello, es una variable aleatoria continua.

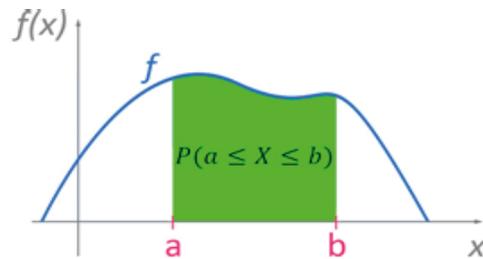
#### 35.2. Función de densidad de probabilidad

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

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

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

Ilustración 64 Curva de densidad



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

Fuente: Matemóvil

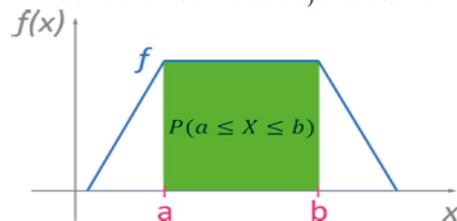
La gráfica de  $f(x)$  se suele llamar curva de densidad.

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

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

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

Ilustración 65 Área bajo la curva



Fuente: Matemóvil

### Ejemplo 1

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

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

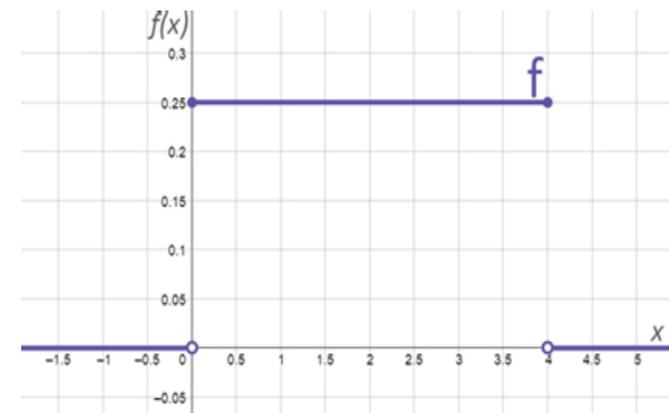
Solución:

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

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

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

Ilustración 66 Función de  $f(x)$

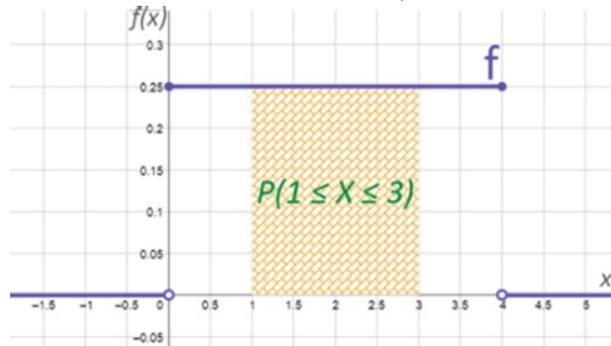


Fuente: Matemóvil

Como queremos calcular la probabilidad de que

nuestra variable aleatoria discreta  $X$  tome un valor entre 1 y 3, entonces sombreamos el área bajo la función  $f(x)$  en ese intervalo:

Ilustración 67 Área bajo la función

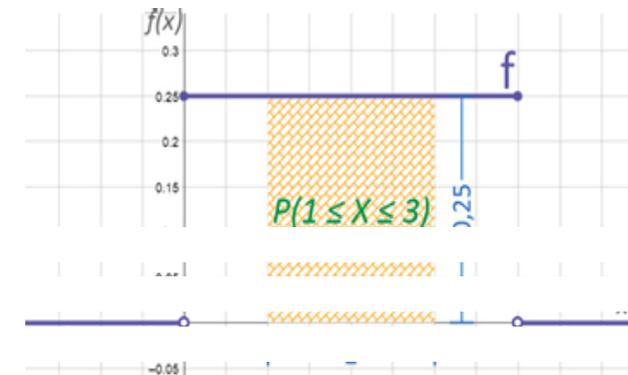


Fuente: Matemóvil

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

### 35.2. Con áreas:

Ilustración 68 Cálculo de área



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

Fuente: Matemóvil

### 35.3. Con integrales:

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

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

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

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

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

#### Propiedad importante

En la función de probabilidad de una variable aleatoria continua sucede algo bien interesante si queremos calcular la probabilidad de que la variable

aleatoria sea igual a un valor puntual  $c$ .

$$P(X = c)$$

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

$$P(X = c) = \int_c^c f(x)dx = 0$$

El hecho de que  $P(X = c) = 0$  cuando  $X$  es continua nos permite afirmar que la probabilidad de que  $X$  quede en algún intervalo entre  $a$  y  $b$  no depende de sus límites.

$$P(a \leq X \leq b) = P(a < X < b) = P(a < X \leq b) = P(a \leq x < b)$$

### 36. Función de distribución acumulativa de una variable aleatoria continua

La función de distribución acumulativa es la función que para un valor  $x$ , nos da la probabilidad de que la variable aleatoria sea menor o igual que dicho valor  $x$ . La función de distribución acumulativa de la variable aleatoria  $X$  es la función  $F(x) = P(X \leq x)$ . Como aún quedan muchas dudas, mejor veamos algunos ejemplos y ejercicios.

#### Introducción

En esta ocasión, vamos a trabajar con los siguientes elementos:

- $X$ : variable aleatoria continua.
- $f(x)$ : función de densidad de probabilidad.

- $F(x)$ : función de distribución acumulativa.

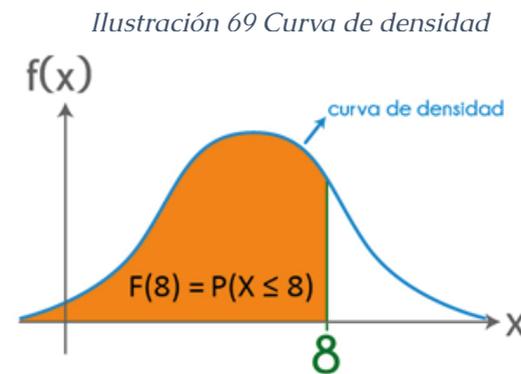
La función de distribución acumulativa es la función que para un valor  $x$ , nos da la probabilidad de que la variable aleatoria sea menor o igual que dicho valor  $x$ . Recuerda que a la función de distribución acumulativa la denominamos  $F(x)$ .

$$F(x) = P(X \leq x)$$

Por ejemplo, si tenemos que calcular  $F(8)$ , haríamos lo siguiente:

$$F(8) = P(X \leq 8)$$

En la siguiente gráfica, tenemos de color celeste a la curva de densidad (gráfica de  $f(x)$ ), y de color naranja, a  $F(8)$ .



Fuente: Matemóvil

Es decir, tenemos de color naranja a:

$$F(8) = P(X \leq 8)$$

Para calcular el valor del área bajo la curva de densidad a la izquierda de 8, tenemos que usar la integral definida:

$(8) = P(X \leq 8) = \text{área bajo la curva de densidad a la izquierda de 8}$

$$F(8) = P(X \leq 8) = \int_{-\infty}^8 f(x)dx$$

Esto fue solo un ejemplo, ahora vamos a ver la forma general y la definición formal.

### 36.1. Función de distribución acumulativa

La función de distribución acumulativa es la función que para un valor  $x$ , nos da la probabilidad de que la variable aleatoria sea menor o igual que dicho valor  $x$ . A la función de distribución acumulativa la denominamos  $F(x)$ .

A continuación, viene la definición formal:

Sea  $X$  una variable aleatoria continua con función de densidad de probabilidad  $f(x)$ . La función de distribución acumulativa de  $X$  es la función:

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t)dt$$

De forma gráfica,  $F(x)$  es el área bajo la curva de densidad a la izquierda de  $x$ . Recordemos que cuando trabajamos con la función de densidad, área es probabilidad.

Además, tenemos algunas fórmulas interesantes que nos servirán para resolver los problemas.

$$P(X > a) = 1 - F(a)$$

$$P(a \leq X \leq b) = F(b) - F(a) \quad ; \quad a < b$$

La siguiente fórmula nos permite pasar de la función de distribución acumulativa a la función de densidad de probabilidad:

$$f(x) = \frac{dF(x)}{dx}$$

Recuerda también que, al ser una función acumulativa, esta no puede ser decreciente.

Ejemplo

La variable aleatoria continua  $X$  tiene la siguiente función de densidad:

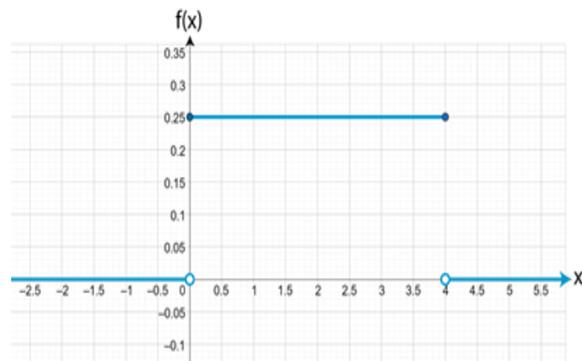
$$f(x) = \begin{cases} 0,25; & \text{si } 0 \leq x \leq 4 \\ 0 & ; \text{ en caso contrario} \end{cases}$$

Definir y graficar la función de distribución acumulativa de  $X$ .

Solución:

Iniciamos graficando la función de densidad para no meternos en problemas.

Ilustración 70 Función de densidad



Fuente: Matemóvil

Como se ve en la gráfica, la curva de densidad tiene 3 tramos bien marcados. Vamos a calcular el valor de  $F(x)$  en cada uno de los tramos empleando la siguiente fórmula:

$$F(x) = \int_{-\infty}^x f(t) dt$$

Vuelvo a colocar la misma función de densidad  $f(x)$ , pero esta vez como  $f(t)$ , es lo mismo, no te preocupes.

$$f(t) = \begin{cases} 0,25; & \text{si } 0 \leq t \leq 4 \\ 0 & ; \text{ en caso contrario} \end{cases}$$

Partimos con el primer tramo:

Parte 1: si  $x > 4$

$$F(x) = \int_{-\infty}^x f(t) dt = \int_{-\infty}^x 0 dt = 0$$

Parte 2: si  $0 \leq x \leq 4$

$$F(x) = \int_{-\infty}^x f(t) dt = \int_{-\infty}^0 f(t) dt + \int_0^x f(t) dt$$

$$F(x) = \int_{-\infty}^0 0 dt + \int_0^x 0,25 dt$$

$$F(x) = 0 + 0,25t \Big|_0^x = 0,25(x - 0) = 0,25x$$

Parte 3: si  $x > 4$

$$F(x) = \int_{-\infty}^x f(t) dt$$

$$F(x) = \int_{-\infty}^0 f(t) dt + \int_0^4 f(t) dt + \int_4^x f(t) dt$$

$$F(x) = \int_{-\infty}^0 0 dt + \int_0^4 0,25 dt + \int_4^x 0 dt$$

$$F(x) = 0 + 0,25t \Big|_0^4 + 0$$

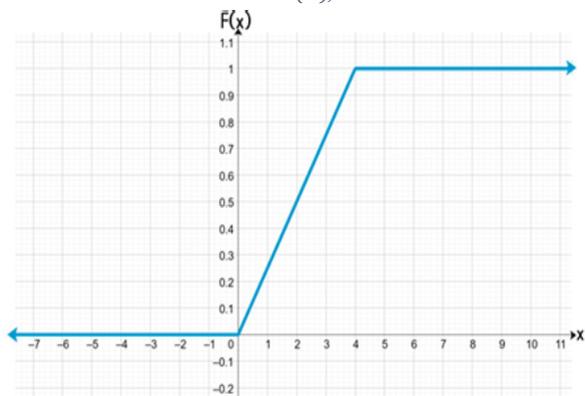
$$F(x) = 0,25(4 - 0) = 1$$

Finalmente, definimos  $F(x)$ .

$$F(x) = \begin{cases} 0 & ; \text{ si } x < 0 \\ 0,25x & ; \text{ si } 0 \leq x \leq 4 \\ 1 & ; \text{ si } x > 4 \end{cases}$$

Terminamos con la gráfica de  $F(x)$ .

Ilustración 71 Gráfica  $f(x)$ , curva de densidad



Fuente: Matemóvil

### 37. Media o valor esperado, varianza y desviación estándar de una variable aleatoria continua

El día de hoy veremos cómo calcular la media (esperanza o valor esperado), la varianza y la desviación estándar de una variable aleatoria continua a partir de su función de densidad de probabilidad.

#### 37.1. Media o valor esperado

La media, llamada también valor esperado o esperanza, se denota con  $\mu$  o  $E(X)$  y su fórmula es:

$$\mu = E(X) = \int_{-\infty}^{\infty} x \cdot f(x) dx$$

- Varianza

Es una medida de dispersión, se representa con  $\sigma^2$  o  $V(X)$  y su fórmula es:

$$\sigma^2 = V(X) = \int_{-\infty}^{\infty} (x - \mu)^2 \cdot f(x) dx$$

Una fórmula alternativa y mucho más rápida es la siguiente:

$$\sigma^2 = V(X) = \int_{-\infty}^{\infty} x^2 \cdot f(x) dx - \mu^2$$

- Desviación estándar

Es una medida de dispersión. Se representa con  $\sigma$  y se calcula teniendo en cuenta que es la raíz cuadrada (positiva) de la varianza:

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

La varianza y la desviación estándar dan medidas cuantitativas de cuánta dispersión hay en la distribución o población de valores  $x$ .

#### Ejemplo 1

La variable aleatoria continua  $X$  tiene la siguiente función de densidad de probabilidad:

$$f(x) = \begin{cases} 1 - \frac{x}{2} & \text{si } 0 \leq x \leq 2 \\ 0 & \text{en caso contrario} \end{cases}$$

Calcular el valor esperado, la varianza y la desviación estándar.

#### Solución:

Empezaremos graficando esta función de densidad de probabilidad. Yo lo haré con ayuda de GeoGebra, pero puedes hacerlo manualmente.

Ilustración 72 Función de densidad de probabilidad



Fuente: Matemóvil

Ten en cuenta que la función de densidad obtenida tiene 3 tramos.

Ahora sí, calculamos la media o valor esperado:

$$\mu = E(X) = \int_{-\infty}^{\infty} x \cdot f(x) dx$$

Como la función tiene 3 tramos, expresaremos la integral original como la suma de 3 integrales:

$$\begin{aligned} \sigma^2 &= V(X) = \int_{-\infty}^{\infty} x^2 \cdot f(x) dx - \mu^2 \\ \sigma^2 &= V(X) = \int_{-\infty}^0 x^2 \cdot f(x) dx + \int_0^2 x^2 \cdot f(x) dx + \int_2^{\infty} x^2 \cdot f(x) dx - \mu^2 \\ \sigma^2 &= V(X) = \int_{-\infty}^0 x^2 \cdot 0 dx + \int_0^2 x^2 \cdot dx + \int_2^{\infty} x^2 \cdot 0 dx - \mu^2 \\ \sigma^2 &= V(X) = \int_{-\infty}^0 0 dx + \int_0^2 x^2 \cdot \left(1 - \frac{x}{2}\right) dx + \int_2^{\infty} 0 dx - \left(\frac{2}{3}\right)^2 \\ \sigma^2 &= V(X) = 0 + \int_0^2 \left(x^2 - \frac{x^3}{2}\right) dx + 0 - \frac{4}{9} \\ \sigma^2 &= V(X) = \left(\frac{x^3}{3} - \frac{x^4}{8}\right) \Big|_0^2 - \frac{4}{9} \\ \sigma^2 &= V(X) = \frac{2^3}{3} - \frac{2^4}{8} - \left(\frac{0^3}{3} - \frac{0^4}{8}\right) - \frac{4}{9} \\ \sigma^2 &= V(X) = \frac{8}{3} - \frac{16}{8} - 0 - \frac{4}{9} \\ \sigma^2 &= V(X) = \frac{8}{3} - 2 - \frac{4}{9} \\ \sigma^2 &= V(X) = \frac{24 - 18 - 4}{9} \\ \sigma^2 &= V(X) = \frac{2}{9} \end{aligned}$$

Luego, calcularemos la varianza. Es mejor usar la fórmula alternativa:

$$\begin{aligned} \mu &= E(X) = \int_{-\infty}^0 x \cdot f(x) dx + \int_0^2 x \cdot f(x) dx + \int_2^{\infty} x \cdot f(x) dx \\ \mu &= E(X) = \int_{-\infty}^0 x \cdot 0 dx + \int_0^2 x \cdot \left(1 - \frac{x}{2}\right) dx + \int_2^{\infty} x \cdot 0 dx \\ \mu &= E(X) = \int_{-\infty}^0 0 dx + \int_0^2 \left(x - \frac{x^2}{2}\right) dx + \int_2^{\infty} 0 dx \\ \mu &= E(X) = 0 + \left(\frac{x^2}{2} - \frac{x^3}{6}\right) \Big|_0^2 + 0 \\ \mu &= E(X) = \frac{2^2}{2} - \frac{2^3}{6} - \left(\frac{0^2}{2} - \frac{0^3}{6}\right) \\ \mu &= E(X) = \frac{4}{2} - \frac{8}{6} = 2 - \frac{4}{3} \\ \mu &= E(X) = \frac{2}{3} \end{aligned}$$

Terminamos con la desviación estándar, que es la raíz cuadrada positiva de la varianza:

$$\sigma = \sqrt{\frac{2}{9}}$$

$$\sigma = \frac{\sqrt{2}}{3}$$



## CAPÍTULO VII

### DISTRIBUCIÓN DE BERNOULLI

La distribución de Bernoulli es una distribución de probabilidad discreta que abre el camino a la distribución binomial, la distribución geométrica y la distribución binomial negativa. Lleva ese nombre por Jakob Bernoulli, un gran matemático y científico suizo.

Veamos los detalles de esta distribución.

#### ➤ Introducción

Vamos a partir de un ensayo de Bernoulli, es decir, de un experimento que tiene solo 2 resultados posibles, a uno de ellos lo llamaremos éxito y al otro fracaso.

Por ejemplo, voy a realizar un experimento bien facilito que consiste en preguntar a un suscriptor de este canal seleccionado al azar si le gusta la pizza. Si me dice que si, lo considero un éxito, si me dice que no, lo considero un fracaso. Otro ensayo de Bernoulli sería lanzar mi moneda que de un lado tiene gato y del otro lado tiene perro. Si defino perro como éxito, obtener un gato será un fracaso.

*Ilustración 73 Moneda de perro y gato*



*Fuente: Matemóvil*

*Se denomina ensayo de Bernoulli a todo experimento*

aleatorio que tiene solo dos resultados posibles mutuamente excluyentes, generalmente llamados éxito y fracaso.

En un ensayo de Bernoulli vamos a definir la variable aleatoria  $X$  de la siguiente manera:

- si obtenemos un éxito, la variable aleatoria  $X$  vale 1.
- caso contrario, si el ensayo termina en fracaso, la variable aleatoria  $X$  vale 0.

De manera sencilla podemos decir que  $X$  cuenta el número de éxitos. Si sale fracaso, el número de éxitos sería 0, no hay ningún éxito. Si el experimento termina en éxito, hay 1 éxito.

A continuación, vamos a elaborar una tabla muy sencilla con los valores de  $X$  en un ensayo de Bernoulli

Tabla 56 Valores de  $X$

	Fracaso	Éxito
$x$	0	1

Fuente: Matemóvil

La probabilidad de éxito se denota por  $p$ , por ello, la probabilidad de fracaso será  $1-p$ . Agregamos esta información a la tabla:

Tabla 57 Probabilidad de fracaso

	Fracaso	Éxito
$x$	0	1
$f(x)$	$1-p$	$p$

Fuente: Matemóvil

Y listo, ya tenemos la distribución de probabilidad de Bernoulli representada mediante una tabla. Si queremos la función de probabilidad (una fórmula), sería la siguiente:

$$f(x) = P(X = x) = \begin{cases} 1-p; & x = 0 \\ p; & x = 1 \end{cases}$$

### 38.1. Distribución de Bernoulli

Una variable aleatoria discreta  $X$  tiene una distribución de Bernoulli si su función de probabilidad está dada por:

$$f(x) = P(X = x) = \begin{cases} 1-p; & x = 0 \\ p; & x = 1 \end{cases}$$

De forma resumida, se puede colocar de la siguiente manera:

$$f(x) = P(X = x) = p^x(1-p)^{1-x}; \text{ con } x = 0; 1$$

Para indicar que la variable aleatoria  $X$  sigue una distribución de Bernoulli de parámetro  $p$  usamos la siguiente notación:

$$X \sim Be(p)$$

### 38.2. Media y varianza de una variable aleatoria de Bernoulli

Si  $X$  tiene una distribución de Bernoulli de parámetro  $p$ , entonces:

$$\text{Media: } \mu = E(x) = p$$

$$\text{Varianza: } \sigma^2 = p(1-p)$$

Listo, ya fue demasiada teoría, vamos con los

problemas:

Ejemplo

Un juego consiste en lanzar un dado una sola vez, si se obtiene un 3 se gana el juego, si se obtiene otro valor, se pierde. Sea  $X =$  el número de veces que sale un 3. Determinar la distribución de probabilidad de  $X$ .

Solución

Vamos a considerar un éxito si se obtiene un 3, con cualquier otro resultado, consideraremos un fracaso. Como  $X$  cuenta el número de veces que sale un 3, puede tomar dos valores:

- $X = 0$ , cuando no se obtenga 3.
- $X = 1$ , cuando se obtenga un 3.

Armamos nuestra tabla:

Tabla 58 Fracaso; Éxito

	Fracaso (■;■;■;■;■)	Éxito (■)
$x$	0	1

Fuente: Matemóvil

Ahora asignamos probabilidades, calculamos primero la probabilidad de éxito  $p$ , es decir, la probabilidad de obtener un 3 al lanzar un dado:

$$p = \frac{\text{n}^\circ \text{ de casos favorables}}{\text{n}^\circ \text{ total de casos posibles}}$$

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

Recuerda que la probabilidad de fracaso será  $1 - p$ . Completamos la tabla asignando probabilidades  $f(x)$ .

Tabla 59 probabilidades de  $f(x)$

	Fracaso (■;■;■;■;■)	Éxito (■)
$x$	0	1
$f(x)$	$1 - \frac{1}{6} = \frac{5}{6}$	$\frac{1}{6}$

Fuente: Matemóvil

Ya tenemos la distribución de probabilidad expresada con una tabla.

Vamos en búsqueda de la función de probabilidad, sabemos de la tabla que:

$$f(x) = P(X = x) = \begin{cases} \frac{5}{6}; & x = 0 \\ \frac{1}{6}; & x = 1 \end{cases}$$

También le podemos dar la siguiente forma compacta equivalente:

$$f(x) = P(X = x) = \frac{1^x}{6} \left(\frac{5}{6}\right)^{1-x}; \quad x = 0; 1$$

Como vemos, la función de probabilidad de  $X$  tiene la forma de una función de probabilidad de Bernoulli con probabilidad de éxito  $p$  de  $1/6$ . Por lo tanto:

Esa sería la respuesta, la variable aleatoria  $X$  sigue una distribución de Bernoulli con una probabilidad de éxito de  $1/6$ .

$$X \sim Be\left(\frac{1}{6}\right)$$

El problema no lo pide, pero, así como hemos

representado la función de probabilidad de  $X$  mediante una tabla y mediante una fórmula, también lo podemos hacer mediante un diagrama de barras. Colocamos un eje horizontal para los valores de  $X$  y otro vertical para las probabilidades  $f(x)$ .



Fuente: Matemóvil



## CAPÍTULO VIII

### DISTRIBUCIÓN BINOMIAL

### 39. Distribución binomial (Introducción)

La distribución de probabilidad binomial es una distribución discreta que tiene muchísimas aplicaciones. Se asocia con un experimento de múltiples pasos que se llama experimento binomial. La palabra binomial viene de otra palabra que significa “dos nombres” y esto nos hará recordar que en cada ensayo que veremos en este tema siempre habrá dos resultados: éxito y fracaso. Si es que respondes una pregunta de alternativas al azar, la respuesta es correcta o incorrecta. Si es que realizas un control de calidad a un producto, este será defectuoso o no defectuoso. Si es que lanzas una moneda, sale cara o sale cruz.

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

#### 39.1. Experimento binomial

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

- El experimento consta de una secuencia de  $n$  ensayos idénticos. En cada ensayo hay dos resultados posibles. A uno de ellos se le llama éxito y al otro, fracaso.
- La probabilidad de éxito es constante de un ensayo a otro, nunca cambia y se denota por  $p$ .
- Por ello, la probabilidad de fracaso será  $1 - p$ . Esto se

debe a que la probabilidad de éxito más la probabilidad de fracaso suman 1.

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

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

### Las hamburguesas de Jorge

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

Ilustración 75 Matehamburguesas



Fuente: Matemóvil

La probabilidad de que a un cliente nuevo le guste la matehamburguesa es de 0,8. Si vienen 3 nuevos clientes a mi cafetería ¿cuál será la probabilidad de que solo a

dos de ellos les guste la hamburguesa?

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

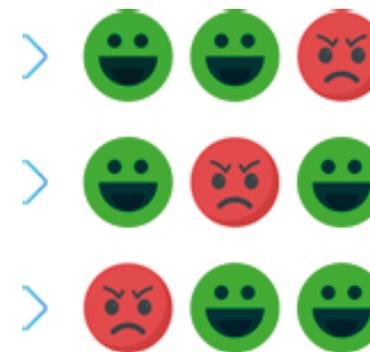
Ilustración 76 Satisfacción del cliente



Fuente: Matemóvil

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

Ilustración 77 Satisfacción del cliente



Obtenido de: Matemovil.com

Si la probabilidad de que a un cliente le guste la mate

hamburguesas es de 0,8, entonces, la probabilidad de que no le gusten será:

Ilustración 78 Probabilidad de satisfacción

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

$$P(\text{😊}) + P(\text{😡}) = 1$$

$$P(\text{😡}) = 1 - P(\text{😊})$$

$$P(\text{😡}) = 1 - 0,8$$

$$P(\text{😡}) = 0,2$$

Fuente: Matemóvil

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

Ilustración 79 Cálculo de probabilidad

$$> \text{😊😊😡} > 0,8 \times 0,8 \times 0,2$$

Fuente: Matemóvil

Coloco ahora todas las opciones.

Ilustración 80 Cálculo de probabilidad

$$> \text{😊😊😡} > 0,8 \times 0,8 \times 0,2$$

$$> \text{😊😡😊} > 0,8 \times 0,2 \times 0,8$$

$$> \text{😡😊😊} > 0,2 \times 0,8 \times 0,8$$

Fuente: Matemóvil

Esas mismas probabilidades, las colocaré en forma de potencias:

Ilustración 81 Cálculo de probabilidad

$$\text{😊😊😡} > 0,8 \times 0,8 \times 0,2 = (0,8)^2 \times (0,2)^1$$

$$\text{😊😡😊} > 0,8 \times 0,2 \times 0,8 = (0,8)^2 \times (0,2)^1$$

$$\text{😡😊😊} > 0,2 \times 0,8 \times 0,8 = (0,8)^2 \times (0,2)^1$$

Fuente: Matemóvil

Me quedaré solo con las potencias, eliminando el resto.

Ilustración 82 Potencias

$$\text{😊😊😡} > (0,8)^2 \times (0,2)^1$$

$$\text{😊😡😊} > (0,8)^2 \times (0,2)^1$$

$$\text{😡😊😊} > (0,8)^2 \times (0,2)^1$$

Fuente: Matemóvil

A continuación, calculamos la probabilidad total de

que solo a 2 de los 3 clientes les guste mi hamburguesa empleando la regla de la suma de probabilidades para eventos mutuamente excluyentes.

Ilustración 83 Cálculo de probabilidad

$$\begin{array}{l}
 \text{😊😊😡} > (0,8)^2 \times (0,2)^1 + \\
 \text{😊😡😊} > (0,8)^2 \times (0,2)^1 + \\
 \text{😡😊😊} > (0,8)^2 \times (0,2)^1 + \\
 \hline
 & 3 \times (0,8)^2 \times (0,2)^1
 \end{array}$$

Fuente: Matemóvil

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

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

### 39.2. Función de probabilidad binomial (fórmula)

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

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

Siendo el coeficiente binomial o número combinatorio:

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

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

$$f(x) = P(X = x) = \frac{n!}{x! (n - x)!} (p)^x (1 - p)^{n-x}, \text{ para } x = 0, 1, 2, \dots, n$$

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

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

### 39.3. Media, varianza y desviación estándar de la distribución binomial

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

Media:

$$\mu = np$$

Varianza:

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

Desviación estándar:

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

Ejemplos

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

Solución:

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

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

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

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

Aplicaremos la fórmula binomial:

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

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

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

Reemplazamos estos valores en la fórmula:

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

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

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

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

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

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

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

La respuesta sería **0,2048**.

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

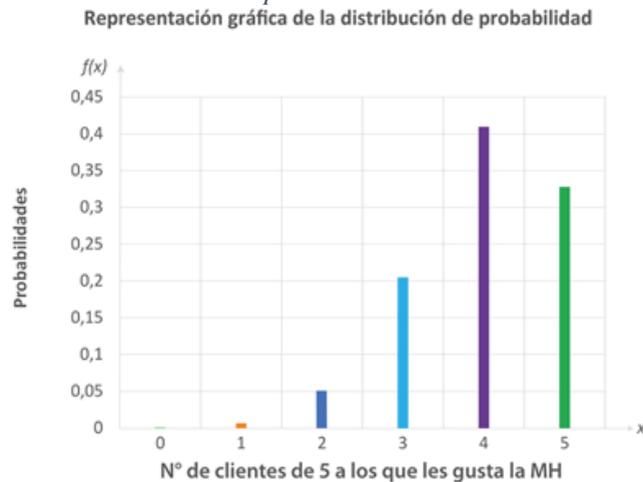
Tabla 60 distribución de probabilidad

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

Fuente: Matemóvil

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

Ilustración 84 Diagrama de barras para una distribución de probabilidad



Fuente: Matemóvil

Algunos usan los histogramas, también es válido.

Ejemplo 2:

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

Solución:

Antes de aplicar la fórmula, verificamos que se trate de un experimento binomial. Para ello, tiene que

cumplir con las 4 condiciones que mencionamos arriba. Efectivamente, se trata de un experimento binomial.

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

$X$  = número de flores de 10 que sobreviven

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

Aplicaremos la fórmula binomial:

$$P(X \geq 9)$$

Nos piden calcular la probabilidad de 9 o más sobrevivan.

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

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

– ¿Y no pueden sobrevivir 11 flores?

– No se puede, porque solo se plantaron 10.

– ¿Y si las flores tienen hijitos bonitos?

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

Ahora colocamos los valores de  $n$ ,  $k$  y  $p$ . Recuerda

que  $n$  es el número de ensayos,  $k$  el número de éxitos y  $p$  la probabilidad de éxito. En este caso:

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

Regresamos con la fórmula de arriba:

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

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

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

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

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

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

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

Continuamos con  $P(X = 10)$ .

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

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

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

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

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

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

Regresamos con esta fórmula:

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

Y reemplazamos lo calculado:

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

$$P(X \geq 9) = 0,7361$$

### Ejemplo 3

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

- Calcular la probabilidad de no obtener ningún éxito.
- Calcular la probabilidad de obtener al menos 1 éxito.

**Solución:**

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

$X =$  número de éxitos en  $n$  ensayos.

$x = 0; 1; 2$ .

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

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

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

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

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

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

$$f(x) = P(X = 0) = (1)(1)(0,6)^2$$

$$f(x) = (X = 0) = \mathbf{0,36}$$

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

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

Aquí nos piden calcular:

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

Pero:

$$f(1) = P(X = 1) = \mathbf{0,48}$$

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

Ade

$$f(2) = P(X = 2) = \mathbf{0,16}$$

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

Reem

$$P(X \geq 1) = 0,48 + 0,16$$

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

Esa sería la respuesta: 0,64.

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

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

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

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

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

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

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

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

#### 40. Media, varianza y desviación estándar de una variable aleatoria binomial

Para encontrar la media, la varianza y la desviación estándar de una variable aleatoria binomial, solo tienes que aplicar algunas fórmulas muy sencillas que veremos aquí.

Veamos los ejercicios que hemos preparado.

##### 40.1. Media o valor esperado de una variable aleatoria binomial

La media la representamos mediante  $\mu$  y es igual al valor esperado  $E(X)$ . La fórmula es la siguiente:

$$\mu = E(X) = np$$

No hay que olvidar que la media es una medida de tendencia central. Su interpretación es interesante y algo enredada, te sugiero ver el video que viene líneas abajo.

#### 40.2. Varianza de una variable aleatoria binomial

La varianza la representamos mediante  $\sigma^2$  o  $V(X)$  y la calculamos con la siguiente fórmula.

$$\sigma^2 = V(X) = np(1 - p)$$

#### 40.3. Desviación estándar de una variable aleatoria binomial

La desviación estándar la representamos mediante  $\sigma$  y es la raíz cuadrada de la varianza.

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

Recuerda que la varianza y la desviación estándar son medidas de dispersión.

##### Ejemplo

En una tienda, el 75% de las compras se hacen con tarjeta de crédito. Sea  $X =$  el número entre diez compras seleccionadas al azar realizadas con tarjeta de crédito. Entonces  $X \sim B(10; 0,75)$ .

##### Calcular:

- El valor esperado de  $X$ .
- La varianza de  $X$ .
- La desviación estándar de  $X$ .

##### Solución:

Para calcular los parámetros que nos piden, usaremos los valores de  $n$  y  $p$ . Siento:

- $n$ : número de ensayos. En este caso,  $n = 10$ .
- $p$ : probabilidad de éxito. En este caso,  $p = 0,75$ .

Estos valores son datos que nos da el enunciado.

#### 40.4. Media o valor esperado

Usamos la siguiente fórmula:

$$\mu = E(X) = np$$

$$\mu = E(X) = 10 \cdot 0,75$$

$$\mu = E(X) = 7,5$$

Aunque  $X$ , el número entre diez compras seleccionadas al azar realizadas con tarjeta de crédito, solo puede tomar valores enteros, su media o valor esperado no tiene que ser un valor entero.

Si se realiza un gran número de experimentos binomiales independientes, siempre con 10 ensayos y una probabilidad de éxito de 0,75, entonces el promedio de los éxitos por experimento se acerca, tiende o se aproxima a 7,5.

#### 40.5. La varianza de $X$ .

La varianza la calculamos usando esta fórmula:

$$\sigma^2 = V(X) = np(1 - p)$$

$$\sigma^2 = V(X) = 10 \cdot 0,75 \cdot (1 - 0,75)$$

$$\sigma^2 = V(X) = 10 \cdot 0,75 \cdot 0,25$$

$$\sigma^2 = V(X) = 1,875$$

#### 40.6. La desviación estándar de X.

Si ya tenemos el valor de la varianza, solo recordaremos que la desviación estándar es la raíz cuadrada positiva de la varianza.

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

$$\sigma = \sqrt{1,875}$$

$$\sigma = 1,3693$$

#### 41. Función de distribución acumulativa de una variable aleatoria binomial

Veamos la función de distribución acumulativa de una variable aleatoria binomial, la cual nos permite resolver de manera rápida muchos problemas de distribución binomial.

Veamos lo que hemos preparado.

Para una variable aleatoria  $X \sim B(n, p)$ , la función de distribución acumulativa está definida por:

$$F(x) = P(X \leq x) = P(X = 0) + P(X = 1) + \dots + P(X = x)$$

Ten en cuenta que la función de distribución acumulativa de una variable aleatoria binomial se representa mediante  $F(x)$ , mientras que la función de probabilidad binomial se presenta mediante  $f(x)$ .

Para calcular las probabilidades acumuladas tienes 2 opciones:

Calcular cada probabilidad de forma individual.

Usar la tabla de probabilidades acumuladas (método recomendado).

#### 41.1. Tabla de probabilidades acumuladas de la distribución binomial

Desde el siguiente link, podrás descargar la tabla de probabilidades acumuladas de la distribución binomial:

#### 41.2. ¿Cómo usar la tabla?

La tabla te permite calcular la probabilidad de obtener  $x$  éxitos o menos en  $n$  ensayos independientes, cada uno con una probabilidad de éxito  $p$ .

Por ejemplo, vamos a calcular la probabilidad de obtener dos éxitos o menos en cuatro ensayos con probabilidad de éxito de  $0.25$ :

$$n = 4 ; x = 2 ; p = 0.25$$

Y ahora buscamos en la tabla:

Ilustración 85 Tabla de probabilidades

**Tabla de probabilidades ac**

La tabla muestra la probabilidad de obtener  $x$  éxitos o menos en  $n$  ensayos independientes, cada uno con probabilidad de éxito

$n$	$x$	0.01	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40
2	0	0.9801	0.9025	0.8100	0.7225	0.6400	0.5625	0.4900	0.4225	0.3600
	1	0.9999	0.9975	0.9900	0.9775	0.9600	0.9375	0.9100	0.8775	0.8400
	2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3	0	0.9703	0.8574	0.7290	0.6141	0.5120	0.4219	0.3430	0.2746	0.2160
	1	0.9997	0.9928	0.9720	0.9393	0.8960	0.8438	0.7840	0.7183	0.6480
	2	1.0000	0.9999	0.9990	0.9966	0.9920	0.9844	0.9730	0.9571	0.9360
4	0	0.9606	0.8145	0.6561	0.5220	0.4096	0.3164	0.2401	0.1785	0.1296
	1	0.9994	0.9860	0.9477	0.8905	0.8192	0.7383	0.6517	0.5630	0.4752
	2	1.0000	0.9995	0.9963	0.9880	0.9728	0.9492	0.9163	0.8735	0.8208
5	0	0.9510	0.7738	0.5905	0.4437	0.3277	0.2373	0.1681	0.1160	0.0778
	1	0.9990	0.9774	0.9185	0.8352	0.7373	0.6328	0.5282	0.4284	0.3370
	2	1.0000	0.9988	0.9914	0.9734	0.9421	0.8965	0.8369	0.7648	0.6826
	3	1.0000	1.0000	0.9995	0.9978	0.9933	0.9844	0.9692	0.9460	0.9130

Fuente: Matemóvil

La probabilidad que buscamos es **0.9492**.

Ejemplo:

Una moneda cuenta con 2 caras: gato y perro. Si se lanza 4 veces la moneda, calcular la probabilidad de obtener 2 o menos perros.

Solución:

En este ejemplo, un ensayo consiste en lanzar una moneda. Vamos a considerar un éxito si obtenemos un perro, caso contrario, consideraremos un fracaso si obtenemos un gato. El experimento binomial consta de una secuencia de 4 ensayos de estos, pues la moneda se lanza 4 veces. También sabemos que la probabilidad de

obtener un éxito (perro) al lanzar una moneda es de 50 % o 0,5.

$$n = 4 \wedge p = 0,5$$

• A partir de aquí definimos nuestra variable aleatoria de interés como:

•  $X =$  número de perros obtenidos al lanzar la moneda 4 veces.

Entonces:

$$X \sim B(4; 0,5)$$

$$x = 0; 1; 2; 3; 4$$

Recordemos la fórmula:

$$F(x) = P(X \leq x) = P(X = 0) + P(X = 1) + \dots + P(X = x)$$

En este caso, nos piden la probabilidad de obtener 2 o menos perros, por ello calcularemos  $P(X \leq 2)$ , es decir, cambiaremos  $x$  por 2.

$$F(2) = P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2)$$

A partir de aquí hay dos caminos, podemos calcular las probabilidades de forma individual o podemos usar la tabla. Por ahora, usaremos la **tabla de probabilidades acumuladas** con los siguientes datos:

$$n = 4; x = 2; p = 0,5$$

Ilustración 86 Tabla de probabilidad para obtener X y N  
(Éxitos ensayos)

La tabla muestra la probabilidad de obtener x éxitos o menos en n ensayos independientes, cada uno con probabilidad de éxito p. Por ejemplo, la probabilidad de

n	x	p											
		0.01	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55
2	0	0.9801	0.9025	0.8100	0.7225	0.6400	0.5625	0.4900	0.4225	0.3600	0.3025	0.2500	0.2025
	1	0.9999	0.9975	0.9900	0.9775	0.9600	0.9375	0.9100	0.8775	0.8400	0.7975	0.7500	0.6975
	2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3		0.01	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55
	0	0.9703	0.8574	0.7290	0.6141	0.5120	0.4219	0.3430	0.2746	0.2160	0.1664	0.1250	0.0911
	1	0.9997	0.9928	0.9720	0.9393	0.8960	0.8438	0.7840	0.7183	0.6480	0.5748	0.5000	0.4253
	2	1.0000	0.9999	0.9990	0.9966	0.9920	0.9844	0.9730	0.9571	0.9360	0.9089	0.8750	0.8336
4	3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
		0.01	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55
	0	0.9606	0.8145	0.6561	0.5220	0.4096	0.3164	0.2401	0.1785	0.1296	0.0915	0.0625	0.0410
	1	0.9994	0.9860	0.9477	0.8905	0.8192	0.7383	0.6517	0.5630	0.4752	0.3910	0.3125	0.2415
	2	1.0000	0.9995	0.9963	0.9880	0.9728	0.9492	0.9163	0.8735	0.8208	0.7585	0.6875	0.6090
3	1.0000	1.0000	0.9999	0.9995	0.9984	0.9961	0.9919	0.9850	0.9744	0.9590	0.9375	0.9085	
4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	

Fuente: Matemóvil

Y listo:

$$F(2) = P(X \leq 2) = 0,6875$$

Otra forma

Si no te gusta usar la tabla, también puedes realizar el cálculo a mano, pero te tomará mucho más tiempo:

$$F(2) = P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2)$$

$$P(X \leq 2) = \binom{4}{0} 0,5^0 (1 - 0,5)^{4-0} + \binom{4}{1} 0,5^1 (1 - 0,5)^{4-1} + \binom{4}{2} 0,5^2 (1 - 0,5)^{4-2}$$

$$P(X \leq 2) = (1)(1)(0,5)^4 + (4)(0,5)(0,5)^3 + (6)(0,5)^2(0,5)^2$$

$$P(X \leq 2) = 0,0625 + 0,25 + 0,375$$

$$P(X \leq 2) = 0,6875$$

## 42. Distribución binomial en Excel

Hoy revisamos como calcular probabilidades binomiales usando Excel. Solo es necesario usar la función DISTR.BINOM.N.

Veamos los detalles.

La función DISTR.BINOM.N

Esta función usa los siguientes argumentos:

- Núm\_éxito: número de éxitos. En nuestras clases, es el valor de la variable aleatoria binomial X.

- Ensayos: número de ensayos. En clase, le asignamos la “n”.

- Prob\_éxito: probabilidad de un éxito en un ensayo. Le asignamos la “p”.

- Acumulada: si quieres el valor de una probabilidad puntual, coloca “Falso”. Si deseas calcular una probabilidad acumulada, coloca “Verdadero”.

Veamos algunos ejemplos.

## 43. Función de probabilidad binomial

Por ejemplo, si deseas calcular la probabilidad de obtener 2 éxitos en un experimento binomial de 4 ensayos, con una probabilidad de éxito de 0.5 en cada ensayo, coloca esta fórmula:

$$=DISTR.BINOM.N(2; 4; 0.5; FALSO)$$

El valor obtenido es de 0.375.

Para calcular esta probabilidad, Excel hace uso de la clásica fórmula de la función de probabilidad binomial.

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

#### 44. Función de distribución acumulativa

Si deseas calcular una probabilidad binomial acumulada, usa la misma función, pero coloca "VERDADERO" al último argumento.

Por ejemplo, si deseas calcular la probabilidad de obtener 2 éxitos o menos en un experimento binomial de 4 ensayos, con una probabilidad de éxito de 0.5 en cada ensayo, coloca esta fórmula:

$$=DISTR.BINOM.N(2; 4; 0.5; VERDADERO)$$

El valor obtenido es de 0.6875.

Para calcular esta probabilidad, Excel hace uso de la clásica fórmula de la función de distribución acumulativa de una variable aleatoria binomial.

$$F(x) = P(X \leq x) = P(X = 0) + P(X = 1) + \dots + P(X = x)$$

#### 45. Distribución de Poisson

La distribución de Poisson fue propuesta por primera vez por Simeón Poisson en libro publicado en 1837. A medida que pasaron los años, el número de aplicaciones comenzó a aumentar, sobre todo el siglo XX y con la aparición de las computadoras en el siglo XXI permitió

incrementarlas aún más. La distribución de Poisson es una distribución de probabilidad discreta, que describe el número de veces que ocurre un evento durante un intervalo específico; el cual puede ser de tiempo, distancia, área, volumen, entre otros. Por ejemplo, son variables de Poisson: el número de llamadas que recibe una central telefónica en el período de 1 minuto, el número de bacterias en un volumen de 1 litro de agua o el número de fallas en la superficie de una pieza de cerámica rectangular.

#### 45.1 Distribución de Poisson Introducción

Como les había contado antes, de profesor de matemáticas me va muy mal, por eso paso la mayor parte de mi día trabajando en una veterinaria. Todos los días llegan muchos pacientes, sobre todo perros y gatos, pero han llegado también algunas tortugas, mapaches, serpientes, tiburones, vacas, dinosaurios y patos. Esperamos cada día la llegada de 4 pacientes, pero algunos días llegan más, otros menos.

¿Será posible calcular la probabilidad de que lleguen 3 pacientes en un día? o ¿5 pacientes en un día?

La respuesta es que sí, gracias a la fórmula de la distribución de Poisson. Eso sí, antes de aplicar la fórmula de la distribución de Poisson, tenemos que verificar que nos encontramos ante un experimento de Poisson. Si no estamos en un experimento de Poisson, no podemos aplicar la fórmula.

## 45.2. Experimento de Poisson

En una distribución de Poisson, se cumplen siempre los siguientes supuestos:

i) la variable aleatoria es el número de veces que ocurre un evento durante un intervalo definido. El intervalo puede ser tiempo, distancia, área, volumen o alguna unidad similar.

$X$  = número de veces que ocurre un evento durante un intervalo definido

ii) la probabilidad de ocurrencia es la misma para cualesquiera 2 intervalos de igual longitud.

iii) la ocurrencia o no ocurrencia en cualquier intervalo es independiente de la ocurrencia o no ocurrencia en cualquier otro intervalo.

iv) dos eventos no pueden ocurrir exactamente al mismo tiempo.

Si se cumplen esas condiciones, la variable aleatoria discreta  $X$  sigue una distribución de Poisson y podemos aplicar la fórmula de la distribución de Poisson.

Aquí algunos ejemplos típicos de variables aleatorias que siguen una distribución de Poisson:

- El número de clientes que ingresan a un supermercado en un día.
- El número de accidentes registrados en una fábrica

durante una semana.

- El número de llamadas que recibe una central telefónica en el período de un minuto.
- El número de bacterias en un volumen de un litro de agua.
- El número de vehículos que llegan a una gasolinera en una hora.
- El número de fallas en la superficie de una pieza de cerámica rectangular.
- El número de toxinas en partes por millón encontradas en un litro de agua de un río.

## 45.3. Distribución de Poisson

Se dice que una variable aleatoria discreta  $X$  tiene una distribución de Poisson con parámetro  $\mu$  ( $\mu > 0$ ) si la función de probabilidad de  $X$  es:

$$f(x) = P(X = x) = \frac{e^{-\mu} \cdot \mu^x}{x!}; \quad x = 0; 1; 2; 3; 4; \dots$$

Donde:

- $f(x) = P(X=x)$ : probabilidad de  $x$  ocurrencias en un intervalo.
- $\mu$ : valor esperado o media de  $X$ .
- $e$ : base de los logaritmos naturales. Su valor es

2,71828...

Además, en una distribución de Poisson:

- La media de  $X$  es igual a  $\mu$ .
- La varianza de  $X$  es igual a  $\mu$ .

Ejemplos

La veterinaria de Jorge recibe un promedio de  $\mu = 4$  pacientes por día. Sabiendo que el número de pacientes que llegan en un día sigue una distribución de Poisson, calcular:

- La probabilidad de que lleguen 3 pacientes en un día.

Primero definimos nuestra variable aleatoria:

$X =$  número de pacientes que llegan en un día.

Además, nos indican que esta variable aleatoria  $X$  sigue una distribución de Poisson, entonces podemos aplicar la fórmula:

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

Calculamos la probabilidad de que lleguen 3 pacientes en un día, es decir,  $f(3)$ . Además, el enunciado nos indica que llegan en promedio 4 pacientes por día, es decir,  $\mu = 4$ .

$$f(3) = P(X = 3) = \frac{e^{-4} \cdot 4^3}{3!}$$

$$f(3) = P(X = 3) = \frac{0,01832 \cdot 64}{3 \cdot 2 \cdot 1}$$

$$f(3) = P(X = 3) = 0,1954$$

La probabilidad de que lleguen 3 pacientes en un día es de 0,1954 o 19,54 %.

- La probabilidad de que lleguen 5 pacientes en un día.

Calculamos la probabilidad de que lleguen 5 pacientes en un día, es decir,  $f(5)$ . Además, el enunciado nos indica que el enunciado que llegan en promedio 4 pacientes por día, es decir,  $\mu = 4$ .

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

$$f(5) = P(X = 5) = \frac{e^{-4} \cdot 4^5}{5!}$$

$$f(5) = P(X = 5) = \frac{0,01832 \cdot 1024}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}$$

$$f(5) = P(X = 5) = 0,1563$$

La probabilidad de que lleguen 5 pacientes en un día es de 0,1563 o 15,63 %.



## CAPÍTULO IX

### DISTRIBUCIÓN NORMAL

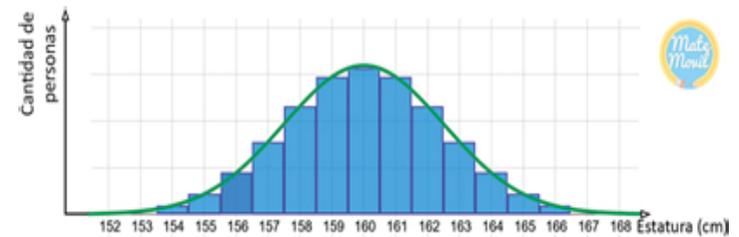
## 46. Distribución normal

La distribución normal, distribución de Gauss o distribución gaussiana, es la **distribución de probabilidad** individual más importante. La distribución normal nos permite crear modelos de muchísimas variables y fenómenos, como, por ejemplo, la estatura de los habitantes de un país, la temperatura ambiental de una ciudad, los errores de medición y muchos otros fenómenos naturales, sociales y hasta psicológicos. Por ello, hoy vamos a revisar sus características y muchísimos problemas resueltos en 3 niveles de dificultad.

### Repaso

¿Qué pasaría si se realiza una encuesta en una ciudad a personas adultas consultando su estatura? A partir de los resultados obtenidos, se puede elaborar un histograma que tendría la siguiente forma:

*Ilustración 87 Estatura de personas adultas*

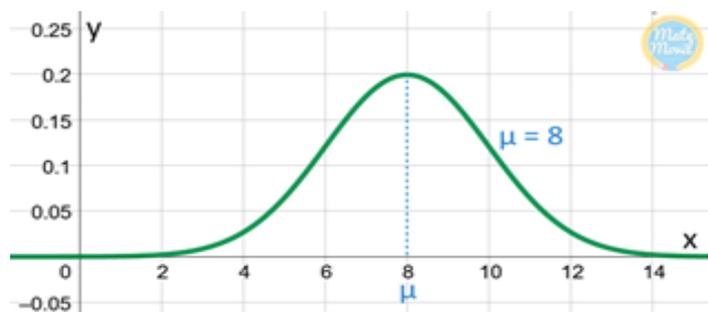


*Fuente: Matemóvil*

Como vemos, el histograma tiene **forma de campana**, una característica importante de la distribución normal.

Un parámetro muy importante es la **media** ( $\mu$ ) y siempre estará al centro de la curva con forma de campana. Por ejemplo, aquí tenemos la gráfica de una distribución normal con **media igual a 8**.

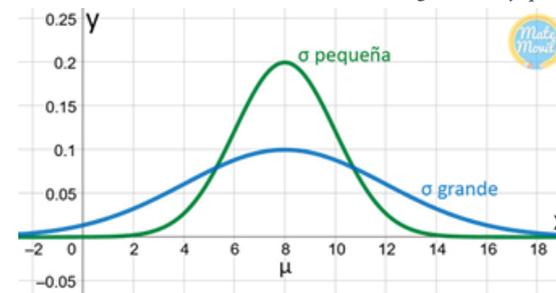
Ilustración 88 Distribución normal



Fuente: Matemóvil

Además de la media, existe otro parámetro muy importante, se trata de la **desviación estándar**, representada con la letra griega  $\sigma$ . La desviación estándar es la medida de variabilidad más utilizada y nos indica que tan dispersos se encuentran los datos. Por ejemplo, aquí veremos dos curvas normales, una con desviación estándar pequeña, y otra con desviación estándar grande. Cuando la desviación estándar es pequeña, los datos tienen una dispersión baja y se agrupan alrededor de la media. En cambio, cuando la desviación estándar es alta, los datos tienen una dispersión alta y se alejan de la media.

Ilustración 89 Desviación estándar grande y pequeña



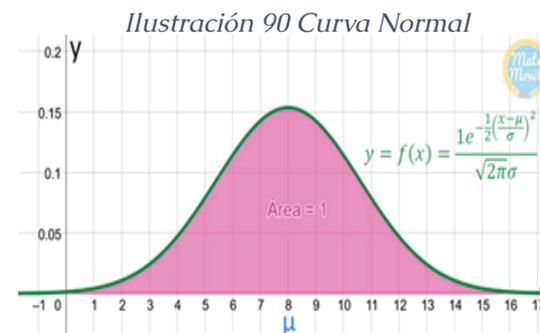
Fuente: Matemóvil

### 46.1. Características de la distribución normal

Existe una función matemática con forma de campana que tiene la siguiente definición:

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

Si se grafica esta función, se obtiene como resultado la **curva normal**:



Fuente: Matemóvil

Además, tiene las siguientes características:

- Toma en cuenta la **media**( $\mu$ ) y la **desviación**

estándar( $\sigma$ ).

- El área bajo la curva es igual a 1.
- Es simétrica respecto al centro, o a la media.
- 50% de los valores son mayores que la media, y 50% de los valores son menores que la media.
- La media es igual a la mediana y a la moda.
- Tiene una asíntota en  $y = 0$  (eje x).

Algunos otros ejemplos de variables con distribución normal:

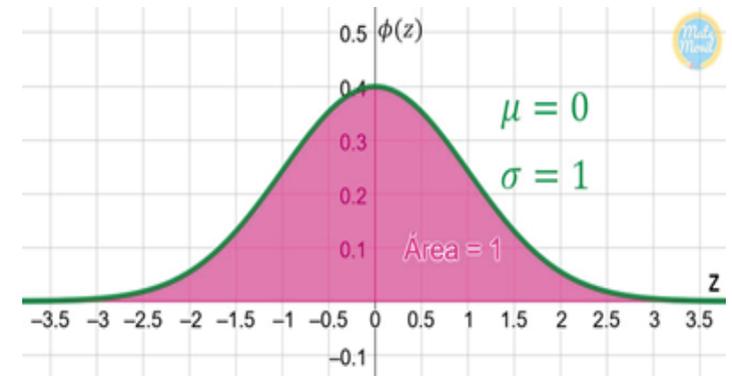
- Notas en un examen.
- Errores de medida.
- Presión sanguínea.
- Tamaño de las piezas producidas por una máquina.

Para encontrar las probabilidades o cantidad de datos entre determinados valores de la variable, se calcula el área bajo la curva normal, que se encuentra en la tabla z o tabla de áreas bajo la curva normal estandarizada.

#### 46.2. La distribución normal estándar

La distribución normal estándar, es aquella distribución normal que tiene una media igual a cero, y una desviación estándar igual a uno.

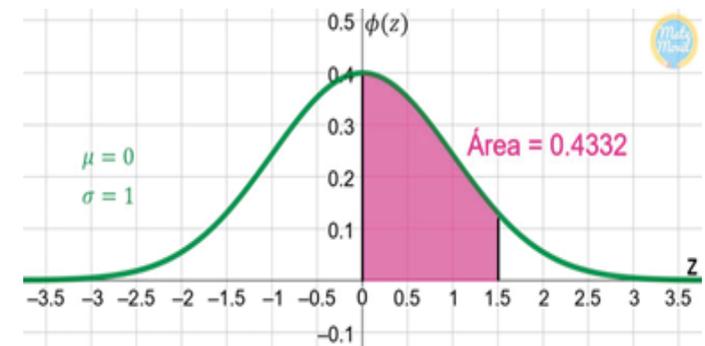
Ilustración 91 Función de densidad normal estandarizada



Fuente: Matemóvil

Por ejemplo, si se desea encontrar la probabilidad de que la variable estandarizada  $z$ , tome un valor entre 0 y 1,50; hay que encontrar el área bajo la curva entre  $z = 0$  y  $z = 1,50$ .

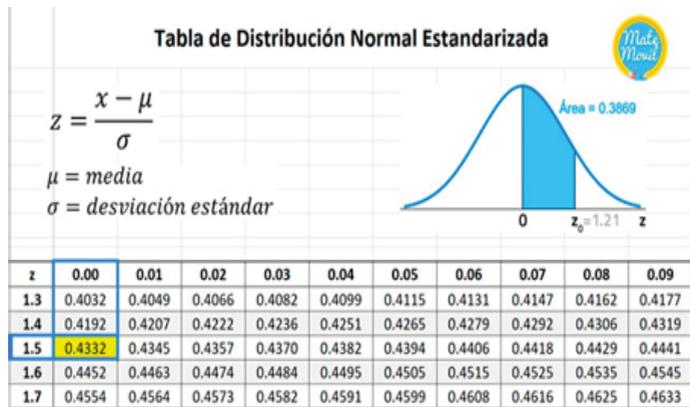
Ilustración 92 Área bajo la curva



Fuente: Matemóvil

Para calcular el valor de esta área, se utiliza la tabla z y se busca el valor de 1,50:

Ilustración 93 Tabla de distribución normal estandarizada



Fuente: Matemóvil

Como vemos, el valor del área bajo la curva es de 0,4332, y esa sería la probabilidad de que la variable estandarizada z, tome un valor comprendido entre 0 y 1,50.

¿Y si mi distribución normal no es estandarizada?

En la mayoría de los problemas, cuando se analizan diferentes variables x, la distribución normal no tiene la forma estandarizada, es decir, la media no es cero y la desviación estándar no es uno. En esos casos, se convierten los valores de la variable (x) a z, es decir, se estandarizan los valores de la variable (x).

La fórmula de la variable estandarizada «z», la cual indica cuántas desviaciones estándar se aleja el valor x de la media, es la siguiente:

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

Y luego, con los valores de z, se utiliza la tabla y se calculan las áreas bajo la curva, porcentajes o probabilidades. En los videos que vienen líneas abajo, encontrarás muchos

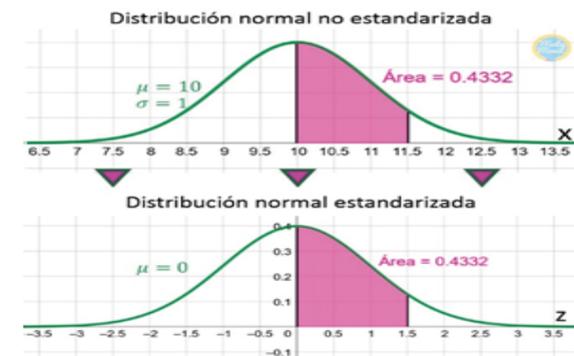
Por ejemplo, si tenemos una variable aleatoria continua X con una distribución normal no estandarizada, con media igual a 10 y desviación estándar igual a 1, y el problema pide calcular la probabilidad de que la variable X tome un valor entre 10 y 11,50, hay que estandarizar los valores de la variable X aplicando la fórmula de z:

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

$$z = \frac{11,50 - 10}{1} = \frac{1,50}{1} = 1,50$$

Ahora, veamos la gráfica de esta distribución normal:

Ilustración 94 Distribución normal estandarizada y no estandarizada

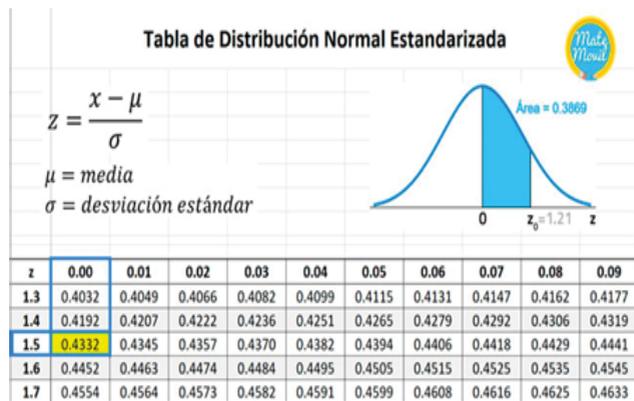


Fuente: Matemóvil

Y usando la tabla z, se calcula el área bajo la curva. Cuando z es igual a 1,50, el área bajo la curva es

de 0,4332.

Ilustración 95 Cálculo área bajo la curva



Fuente: Matemóvil

Podemos concluir que la probabilidad de que la variable estandarizada  $z$ , tome un valor comprendido entre 10 y 11,50 es de 0,4332.

### 46.3. Calculadora de distribución normal en GeoGebra

Usando GeoGebra. Hemos preparado una aplicación que te permitirá resolver cualquier problema al instante y calcular áreas bajo la curva normal, porcentajes y probabilidades.

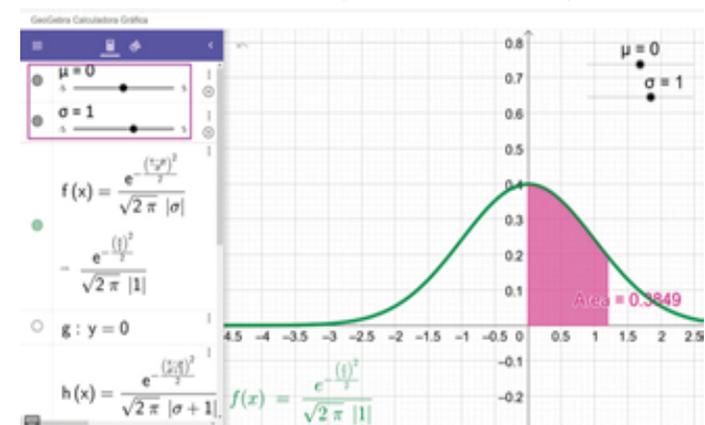
Recuerda que, en la distribución normal, el área bajo la curva normal, nos indica la probabilidad o porcentaje.

### 46.3. Pasos para calcular el área bajo la curva

- En la barra de entradas, modifica el valor de la media, desviación estándar, límite izquierdo y límite

derecho.

Ilustración 96 Geogebra calculadora gráfica



Fuente: Matemóvil

Recuerda que, al momento de cambiar el valor de la media, la curva normal se desplazará hacia el nuevo valor de la media. Te recomendamos ver el video que se encuentra líneas arriba.

- Podrás ver el valor del área, probabilidad o porcentaje en el gráfico.

### 47. Distribución normal en calculadoras Casio ClassWiz 570 y 991

Hoy vamos a revisar como resolver problemas de distribución normal al instante usando la calculadora Casio ClassWiz 570 o 991. Esta calculadora tiene muchísimas funciones estadísticas, y vamos a sacarle el máximo provecho. Veamos cómo trabajan estas funciones.

### 47.1. Función de distribución normal en Casio

Las calculadoras Casio ClassWiz 570 o 991, pueden calcular el área bajo la curva, probabilidad o porcentaje en una distribución normal de una manera muy rápida. En el siguiente video, te mostraremos como usar esta función de distribución normal en las calculadoras, es realmente muy sencillo. Sigue estos pasos:

1. Presiona el botón Menú ubicado arriba a la derecha de la calculadora.
2. Presiona el número 7 para ingresar al modo Distribución.

Ilustración 97 Pasos para calcular el área bajo la curva



Fuente: Matemóvil

1. Presiona el número 2 para ingresar a la función de distribución normal.
2. Coloca los valores del límite inferior, límite superior, media y desviación estándar.

Ilustración 98 Encontrar el valor de probabilidad



Fuente: Matemóvil

¡Y listo! Ahora podrás obtener el valor de la probabilidad al instante con tu calculadora Casio ClassWiz 570 o 991.

### 47.2. Función de distribución normal inversa

Por otro lado, si deseas calcular un valor de la variable  $X$  a partir del área de la cola izquierda bajo la curva normal estándar, puedes usar la función de distribución normal inversa. Sigue estos pasos:

1. Presiona el botón Menú ubicado arriba a la derecha de la calculadora.
2. Presiona el número 7 para ingresar al modo Distribución.

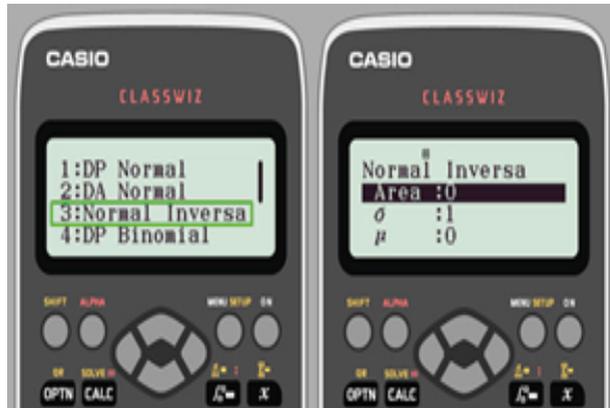
Ilustración 99 Pasos para encontrar la distribución normal inversa



Fuente: Matemóvil

3. Presiona el número 3 para ingresar a la función de distribución normal inversa.
4. Coloca los valores del área de la cola izquierda bajo la curva normal estándar, media y desviación estándar.

Ilustración 100 Encontrar el valor de la variable  $X$



Fuente: Matemóvil

#### 48. Tabla de distribución normal estandarizada

#### (tabla z)

Hoy traemos la tabla de áreas bajo la curva normal estandarizada, conocida también como tabla z. Esta tabla te permitirá encontrar áreas y por supuesto, **probabilidades en la distribución normal**. Existen muchas tablas de este tipo, la diferencia entre ellas se encuentra en la zona sombreada. Pero no te preocupes, al descargar la tabla podrás ver el área bajo la curva de manera sombreada.

#### 48.1. Tabla t de Student

La tabla t de Student es muy importante para resolver problemas de intervalos de confianza y pruebas de hipótesis. Recordemos que la distribución t de Student es una distribución de probabilidad que permite estimar la media de una población con distribución normal, pero con **un tamaño de muestra pequeño**.

La tabla te permite encontrar el valor de la variable t, conociendo el área de la cola derecha y los grados de libertad.

## Referencias bibliográficas

Estadística. (2015, 30 de enero). MateMóvil; Matemóvil.

Autor, A. A., & Autor, B. B. (Año). Título del libro en cursiva. Lugar de publicación: Editorial.

Agresti, A., & Franklin, C. (2019). Estadística: inferencia y métodos. Cengage Learning.

Autor, A. A., Autor, B. B., & Autor, C. C. (Año). Título del libro en cursiva. Lugar de publicación: Editorial.

DeGroot, M. H., & Schervish, M. J. (2012). Probability and Statistics. Pearson.

Autor, A. A., Autor, B. B., & Autor, C. C. (Año). Título del libro en cursiva. Lugar de publicación: Editorial.

Hogg, R. V., McKean, J., & Craig, A. T. (2018). Introduction to Mathematical Statistics. Pearson.

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2019). Introduction to Linear Regression Analysis. Wiley.

Devore, J. L., & Berk, K. N. (2018). Modern Mathematical Statistics with Applications. Springer.

Rice, J. A. (2006). Mathematical Statistics and Data Analysis. Cengage Learning.

ISBN: 978-9942-621-56-6



9 789942 621566