Unidad V

Estadística

Introducción

La estadística es una ciencia matemática aplicada que se ocupa de emplear una serie de métodos y procedimientos para recoger, clasificar, organizar, resumir y analizar datos, así como para sacar conclusiones válidas y tomar decisiones razonables basadas en tal análisis. Es empleada en los diferentes ámbitos de la investigación, especialmente la cuantitativa, por eso la presente unidad tiene el propósito de profundizar en el estudio de la estadística para desarrollar en los estudiantes la habilidad de comunicar información basada en datos cuantitativos y conocer el comportamiento y la posibilidad de ocurrencia de ciertos eventos, conocimiento importante en diferentes campos de estudio.

Los temas que abordaremos en esta unidad son: medidas de dispersión, cálculo de varianza y desviación estándar, el conteo y la probabilidad.

Esperamos que los conocimientos que adquiera en esta unidad los aplique en la vida cotidiana y le sirvan de base para su estudios posteriores en educación superior.

¡ÉXITOS!

¿Qué vamos a aprender?

Competencias	Objetivos	Contenido
1. Conocer y analizar las medidas de dispersión.	1. Aplicar las medidas de dispersión en la clasificación y colección de datos.	1. Medidas de dispersión • Media • Mediana • Moda 2. Cálculo de varianza y desviación estándar 3. Aplicación de la estadística en la vida diaria
2. Determinar la probabilidad de eventos simples y compuestos.	2. Calcular la probabilidad de ocurrencia de eventos en situaciones de la vida cotidiana. 3. Conocer la tecnica de conteo para poder determinar el tamaño de una muestra.	4. Conteo y probabilidad 5. Aplicación de los principios de conteo 6. Determinación de las fórmulas de probabilidad de un evento 7. Propiedades de la probabilidad 8. Cálculo de la probabilidad en eventos simples y compuestos
3. Usar el principio fundamental de conteo para determinar el tamaño del espacio de muestra para eventos simples y compuestos.	4. Aplicar el principio fundamental del conteo al determinar el tamaño de los espacios de la muestra en eventos simples y compuestos.	9. Principios fundamentales del conteo 10. Resolución de problemas de aplicación

Mis conocimientos previos

Instrucciones: lea el siguiente texto y conteste las preguntas que aparecen al final. Luego realice una plenaria con sus compañeros y tutor para compartir sus respuestas y llegar a conclusiones.

Según el informe Estado de la Población Mundial 2012 del Fondo de Población de las Naciones Unidas (UNFPA), Honduras posee una tasa de mortalidad de 100 por cada 100,000 nacidos vivos. En América Latina Honduras ocupa el segundo lugar en embarazos adolescentes. El 30 % de la población embarazada son menores de 18 años, es decir, 108 por cada mil mujeres de 15 a 19 años. Según las Naciones Unidas, en 2012, de los 80 millones de embarazos no deseados 40 millones terminaron en aborto. Se estima que 5 de cada 10 niñas entre los 10 y 14 años salen embarazadas y 2 mensualmente se infectan de VIH, según estadísticas del hospital de El Progreso. Actualmente hay 21 niñas embarazadas menores de 14 años. La mayoría de ellas han sido abusadas sexualmente por sus padres.

1. ¿Cuáles son las ideas principales de la lectura anterior?

   ____

   ____

2. ¿Qué significa para nuestro país ocupar el segundo lugar en América Latina en embarazos adolescentes?

   ____

   ____

3. ¿Para qué sirven estas estadísticas?

   ____

   ____

4. ¿Mencione otras áreas que usted conoce donde se aplican las estadísticas?

   ____

   ____

5. ¿Por qué es importante el estudio de la estadística?

   ____

   ____

6. ¿Cite los temas que recuerda haber aprendido sobre estadística en el espacio curricular anterior de Matemáticas II?

   ____

   ____

Introducción a las medidas de dispersión

Los datos, igual que los estudiantes, se congregan alrededor de puntos de encuentro. Parece que los estudiantes acuden en grandes cantidades a sitios tales como partidos de fútbol, bares populares, cines y en algunas ocasiones a bibliotecas. De igual forma, los números parecen disfrutar de la misma compañía de otros números y están propensos a reunirse alrededor de un punto central denominado medida de la tendencia central o, más comúnmente, media. Una medida de tendencia central ubica e identifica el punto alrededor del cual se centran los datos.

Pero ¿por qué es importante una medida de la tendencia?

Un conjunto grande de datos puede ser rápidamente descrito de manera precisa con un solo número. Si el profesor de Matemáticas III dice que el promedio de la clase en el último examen fue de 94 %, esto indica algo: que los educandos estudiaron, pusieron atención en clases, realizaron todas sus tareas, etc.

Pero si el promedio fuese de 48 % ¿que expresa? Antes de ver este punto, tendremos que recordar que el objetivo es alcanzar una nota que sea mayor a 59 y en otros casos mayor a 69, para poder aprobar la clase, partiendo de esto si el promedio es de 48 % diremos que los educandos no estudiaron, no resolvieron las guías de tarea, no entraron a clases regularmente, etc.

Las medidas de dispersión indican el punto hasta el cual las observaciones individuales se esparcen alrededor de un punto central. También miden la dispersión o la variabilidad de los datos y reflejan la tendencia de las observaciones individuales a desviarse de dicho punto central.

Analicemos un pequeño ejemplo relacionado con lo que antes hablamos:

Cuadro de notas parcial y final de estudiantes

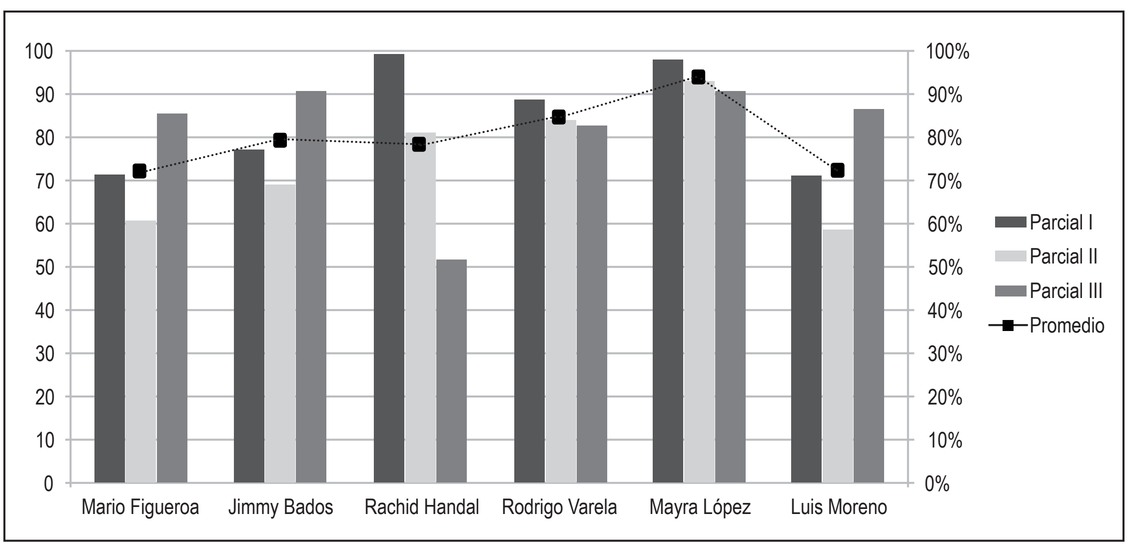
No.	Nombre alumno	Parcial I	Parcial II	Parcial III	Promedio
1	Mario Figueroa	69	59	83	70%
2	Jimmy Bados	75	67	88	77%
3	Rachid Handal	99	79	50	76%
4	Rodrigo Varela	86	81	80	82%
5	Mayra López	95	90	88	91%
6	Luis Moreno	69	57	84	70%
Promedio por parcial		82%	72%	79%	77.72%

Promedio final general

Acabamos de leer una pequeña tabla que contiene las notas parciales de seis estudiantes de la clase de Matemáticas III. En esta tabla encontramos los promedios por estudiantes, los promedios por parcial y, por último, el promedio general de la clase.

El siguiente gráfico refleja el promedio que cada uno de los estudiantes tuvo en los tres parciales y cuál es el promedio final de la clase de Matemáticas III.

Notas por parcial en Matemáticas III



La tabla y el gráfico son formas de agrupar los datos, lo que se interprete con ellos dependerá de las necesidades de cada persona, de momento solo se pretende mostrar a los alumnos lo que se puede obtener con algunos pocos datos.

Metodos de identificación de datos

Existen tres métodos comunes para identificar el centro de un conjunto de datos:

- La media
- La mediana
- La moda

Su cálculo e interpretación dependerá de la definición de "centro"; en cada caso, se ubican en el punto alrededor del cual se aglomeran los datos.

La media

La media o media arítmetica resulta al efectuar una serie de operaciones con un conjunto de números y que en determinadas condiciones puede representar por si solo todo el conjunto. Por ejemplo, para calcular la media de los últimos 5 exámenes de Matemáticas III, simplemente se suman las notas obtenidas y el resultado se divide entre 5.

Media aritmética: es la medida de la tendencia central que normalmente era considerada como el promedio.

La media de una población se representa por μ (que se pronuncia miu). Si hay N observaciones en el conjunto de datos de la población, la media se calcula así:

$$\mu = \frac{x_1 + x_2 + x_3 + \cdots + xn}{n} = \frac{\sum_{i=1}^{N} X_i}{n}$$

Media poblacional [5.1]

La letra griega mayúscula ∑ es el signo de sumatoria que indica que se suman todas las observaciones de 1 a N. Las Xi denotan las observaciones individuales. En aras de la simplicidad, el subíndice y el superíndice se reducirán y aparecerá el signo de la sumatoria como simplemente ∑.

La media de una muestra se representa de la siguiente forma: $\overline{X}$ (que se lee X – barra). Con *n* observaciones en el conjunto de datos de la muestra, la media se determina así:

$$X = \frac{x_1 + x_2 + x_3 + \cdots + xn}{n} = \frac{\sum_{i=1}^{N} X_i}{n}$$

Media muestral [5.2]

Veremos a continuación un pequeño ejemplo de cómo resolver la fórmula de la media muestral:

El tesorero de la empresa Jugos y Más reporta los ingresos mensuales por ventas de jugos y agua purificada.

$$\overline{X} = \frac{56 + 67 + 52 + 45 + 67}{5} = 57.4$$

La mediana

La mediana algunas veces es llamada media posicional, porque queda exactamente en la mitad del conjunto de datos después de que las observaciones se han colocado en serie ordenada. La mitad de las observaciones estará por encima de la mediana, la otra mitad estará por debajo de ella.

Mediana: es la observación de la mitad después de que se han colocado los datos en una serie ordenada.

Si el conjunto de datos tiene un número impar de observaciones, la posición de la mediana es:

$$\textit{Posición de la mediana} = \frac{n+1}{2} \qquad [5.3]$$

Dado el ejemplo anterior sobre los ingresos por ventas, se deben colocar primero los datos en serie ordenada:

45, 52, 56, 67, 67 $\quad$ *Posición de la mediana* $= \dfrac{5+1}{2} = 3$

La posición del valor de la mediana se vuelve a la tercera posición de la serie ordenada anterior, se puede observar que valor es 56.

Con un conjunto de datos que contiene un número par de observaciones es necesario promediar los dos valores medios. Por ejemplo, si los ingresos para un sexto mes son de 35 y se adiciona esta cifra al conjunto de datos, entonces la serie ordenada se vuelve 35, 45, 52, 56, 67, 67. La posición de la mediana es de:

$$\textit{Posición de la mediana} = \frac{6+1}{2} = 3.5$$

Es decir la posición tres y medio; los dos valores de las posiciones tercera y cuarta se promedian para producir una mediana de 52 + 56 = 54. Esto significa que en la mitad de los meses las ventas estuvieron por debajo de L 54,000 y en la mitad de los meses los ingresos excedieron dicha suma.

La moda

La observación modal es la observación que ocurre con mayor frecuencia. Utilizando las seis observaciones anteriores de 35, 45, 52, 56, 67, 67, la moda seria 67. Si una séptima observación de 56 se agrega, el conjunto de datos sería bimodal, con modas de 56 y 67.

Ejemplo 1:

la emisión de la revista el *Forjador* del 26 de noviembre de 2014 reportó que en 2013 las utilidades en millones de lempiras de varias de las 500 mejores compañías que aparecen en la revista incluían:

Máxima	L 7,510	Banco de Honduras	L 7,280
Diario Plus	L 6,246	Galaxia S.A.	L 5,429
ABC Inmobiliaria	L 5,157	Lácteos de Honduras	L 4,289

Calcule las tres medidas de tendencia central.

Solución:

Media: tratando los datos como una muestra se tiene que

$$\overline{X} = \frac{7{,}510 + 7{,}280 + 6{,}246 + 5{,}429 + 5{,}157 + 4{,}289}{6} = 5{,}985$$

Mediana: primero deben colocarse los datos en una serie ordenada, si no están ordenados, la posición que se encuentre utilizando la fórmula 5.3 no tiene sentido.

4,157, 4,289, 5429, 6,246, 7,280, 7,510

$$\textit{Posición de la mediana} = \frac{6 + 1}{2} = 3.5$$

La mediana sería el promedio de los dos valores tercero y cuarto:

$$\frac{5{,}429 + 6{,}246}{2} = 5{,}837.50$$

Moda: este conjunto de datos no tiene moda debido a que todas las observaciones ocurrieron con la misma frecuencia (no se repiten).

Interpretación de datos

En esta etapa se determina qué se hace con cada uno de los resultados encontrados. Imaginemos que la Cámara de comercio dará un premio a las empresas que logren alcanzar la media de ganancias de L 5, 985,000.00, entonces solo Máxima, Diario Plus y Banco de Honduras alcanzaran este reconocimiento. Con la mediana observaremos cuál es el centro de referencia de los datos, en este caso nos sirve para encontrar algún punto de equilibrio.

Por último esta la moda, que en este ejercicio no tenemos debido a que ningún dato se repite, pero ¿qué hubiera representado si un valor anterior se repitiera? Que dos empresas tengan la misma utilidad es muy difícil, pero se puede dar y para que suceda no es necesario que sean del mismo rubro; esto indicaría que están teniendo un mismo crecimiento porcentual. Otro ejemplo de un caso de moda es el siguiente: supongamos que estamos clasificando ropa y encontramos que algunas prendas se utilizan como más frecuencia según se cambia la temporada climática; esta sería la moda y en este caso diríamos que es algo que se está usando con mucha frecuencia.

Nota: los datos son interpretados por el analista según sea el caso, llamando analista a cualquier persona que esté interesada en manipular la información. Los valores numéricos trabajados nos entregan datos que pueden cambiar las proyecciones o tendencias.

La media ponderada

$\overline{X}w$ = media ponderada

$$\overline{X} = \frac{\Sigma XW}{\Sigma W} \qquad [5.4]$$

En donde:
- Xw = media ponderada
- X = observación individual
- $\overline{W}$ = peso o ponderación asignada a cada observación

En la media se asume que cada observación es de igual importancia, sin embargo, en ciertos casos puede querer darse mayor peso a algunas de las observaciones.

Ejemplo 2:

si el último examen de Matemáticas III valdrá el doble que los otros exámenes para determinar la nota final, entonces al puntaje que se obtenga en el examen final debe dársele el doble de peso; es decir que debe contarse doble al calcular la nota. Esto es exactamente lo que hace la media ponderada al utilizar la fórmula 3.4. Asuma que obtuvo un puntaje de 89, 92 y 79 en los exámenes parciales y 94 en el examen final, estos puntajes y sus respectivas ponderaciones los construiremos en la siguiente tabla:

Nota (X)	Peso (W)	XW
89	1	89
92	1	92
79	1	79
94	2	188
	5	**448**

En esta columna llenamos los datos que previamente se nos dieron (en este caso las notas parciales).

Esta es la cantidad de veces que los datos se repiten. A diferencia de la moda, estos datos son directamente seleccionados.

Esta columna es la multiplicación de la columna X por la columna W.

Sumatoria columna W

Sumatoria columna XW

Ahora aplicamos la fórmula con los valores encontrados en la tabla:

$$\textit{Media ponderada} = \frac{448}{5} = 89.6$$

Este método es igual que sumar la nota del examen final dos veces al calcular la media:

$$\overline{X}_w = \frac{89 + 92 + 79 + 94 + 94}{5} = 89.6$$

Media ponderada: toma en cuenta la importancia relativa de las observaciones.

Ejemplo 3:

Don Luis el fontanero vende cinco tipos de limpiadores para desagües. En la tabla se muestra cada tipo junto con la utilidad por lata y el número de latas vendidas.

Limpiador	Utilidad por lata (X)	Peso (W)	XW
Burbujear	L 2.00	3	L 6.00
Poder Limpio	L 3.50	7	L 24.50
Drenajes de Ensueño	L 5.00	15	L 75.00
Limpia Más	L 7.50	12	L 90.00
Desagüe Principal	L 6.00	15	L 90.00
	L 24.00	**52**	**L 285.50**

Se puede calcular la media aritmética simple de la utilidad de don Luis como L 24.00 ÷ 5.00 = L 4.80 por lata.

Solución:

sin embargo, probablemente este no sea un buen estimado de la utilidad promedio de don Luis, debido a que vende más de algunos tipos de limpiadores que de otros. Para obtener un estado financiero más representativo del desempeño real de su negocio, don Luis debe dar más peso a los tipos más populares de limpiadores. Por tanto, el cálculo apropiado sería mediante la media ponderada, ya que la medida de peso apropiada serían las cantidades vendidas.

$$\overline{X}_w = \frac{\Sigma XW}{\Sigma W} = \frac{\text{L } 285.50}{52} = \text{L } 5.49$$

Interpretación: la media ponderada es mayor que la media aritmética simple porque don Luis vende más limpia desagües de los tipos que tienen un margen de utilidad mayor.

La media geométrica

La media geométrica puede utilizarse para mostrar los cambios porcentuales en una serie de números positivos. Como tal, tiene una amplia aplicación en los negocios y en la economía, debido a que con frecuencia se está interesado en establecer el cambio porcentual en las ventas, en el producto nacional bruto o en cualquier serie económica.

Media geométrica: proporciona una medida precisa de un cambio porcentual promedio en una serie de números.

La media geométrica (MG) se halla tomando la raíz enésima del producto de *n* números.

$$MG = \sqrt[n]{X1 * X2 * X3 \ldots Xn} \qquad [5.6]$$

La mayoría de las calculadoras manuales pueden calcular la raíz enésima de cualquier número.

La media geométrica se utiliza con más frecuencia para calcular la tasa de crecimiento porcentual promedio de algunas series dadas, a través del tiempo. Para ilustrar esta aplicación en un escenario de negocios, se consideran las cifras sobre ingresos del ejemplo que a continuación le presentamos:

Ejemplo 4:

el director ejecutivo de aerolíneas Blanco Airlines desea determinar la tasa de crecimiento promedio en los ingresos con base en las cifras dadas en la tabla. Si la tasa de crecimiento promedio es menor que el promedio industrial del 10 %, se asumirá una nueva campaña publicitaria.

Ingresos para Blanco Airlines		
Año	Ingresos	Porcentaje del año anterior
2002	L 50,000.00	-----
2003	L 55,000.00	55 ÷ 50 = 1.10
2004	L 66,000.00	66 ÷ 55 = 1.20
2005	L 60,000.00	60 ÷ 66 = 0.91
2006	L 78,000.00	78 ÷ 60 = 1.30

Solución:

Primero es necesario determinar el porcentaje que los ingresos de cada año representan respecto da los obtenidos el año anterior. En otras palabras, ¿qué porcentaje del ingreso de 2002 es el que ingresó en 2003? Esto se encuentra dividiendo los ingresos de 2002 entre los de 2003. El resultado, 1.10, revela que los ingresos de 2003 son del 110% de los ingresos de 2002. También se calculan los porcentajes para los tres años restantes. Tomando la media geométrica (MG) de estos porcentajes da:

$$MG = \sqrt[4]{(1.10)(1.2)(0.91)(1.3)} = 1.1179$$

Restando 1 para convertirlo a un incremento anual promedio da 0.1179, o un incremento promedio de 11.79% para el período de cinco años.

Por otro lado, la media aritmética simple es:

$$\overline{X} = \frac{1.1 + 1.2 + 0.91 + 1.3}{4} = \frac{4.51}{4} = 1.1275$$

O un cambio promedio de 12.75%, basado en la media aritmética simple, que se aplica a la serie que comienza con L 50,000.00 (ver tabla de ingresos pág. 16), los resultados son:

L 50,000.00	X	1.1275	=	L 56,375.00
L 56,375.00	X	1.1275	=	L 63,562.81
L 63,562.81	X	1.1275	=	L 71,667.07
L 71,667.07	X	1.1275	=	L 80,804.62

Ya que los L 80,804.62 exceden los L 78,000.00 que Blanco Airlines en realidad ganó, el incremento del 12.75% es obviamente muy alto.

Si se utiliza la tasa de crecimiento de la media geométrica del 11.79% se obtiene:

L 50,000.00	X	1.1179	=	L 55,895.00
L 56,375.00	X	1.1179	=	L 62,485.02
L 63,562.81	X	1.1179	=	L 69,852.00
L 71,667.07	X	1.1179	=	L 78,087.56

$78{,}087.56 \approx 78{,}000.00$

Esto da un valor de L 78,087.56, lo que está mucho más cerca al ingreso real

de L 78,000.00.
Interpretación: la media geométrica representa el cambio promedio con el tiempo. Debido a que la tasa de crecimiento supera el promedio de la industria del 10 %, la nueva campaña publicitaria no se llevará a cabo.

Nota: la media geométrica siempre será menor que la media aritmética, salvo en el extraño caso en el que todos los incrementos porcentuales sean iguales. Entonces las dos medias serán iguales.

Comparación entre media, mediana y moda

La media es la medida más común de tendencia central, se presta para mayor manipulación e interpretación algebraica. Desafortunadamente, la media se ve afectada por valores extremos, o valores atípicos y, a diferencia de la mediana, puede ser sesgada por las observaciones que están muy por encima o muy por debajo de esta.

El uso de la moda tiende a ser menos confuso ya que simplemente es el valor que más se repite entre los datos que se obtienen.

En general, todas dependen de la naturaleza de los datos o de la forma en que se utilizan los datos.

Actividad 1

Instrucciones: desarrolle en su cuaderno de trabajo los ejercicios que a continuación se le presentan.

1. Su empresa está introduciendo un nuevo microchip para celular. La promoción dice que el microchip hace que algunas aplicaciones funcionen mucho más rápidamente que los que actualmente se encuentran en el mercado. Se hacen veinte pruebas diferentes, produciendo los tiempos en segundos que se ven más adelante. Aunque no puede tergiversar los datos, usted desea presentar los resultados de la manera más favorable para su firma. Determine la media, la mediana y la moda. Comente los beneficios relativos de utilizar los resultados estadísticos.

Muestra de datos en segundos				
2	8	7	4	6
12	11	5	10	13
5	3	9	3	9
8	7	4	11	5

Para efectos de interpretación, los microchips actuales aceleran en promedio 8.1.

2. Como gerente de ventas de Corporación la Lechuza, usted desea calcular las medidas de tendencia central para los niveles de utilidad de su firma, durante los últimos 12 meses. Las utilidades están dadas en miles de lempiras en el siguiente cuadro:

Datos de los últimos 12 meses		
L 12.3	L 14.3	L 25.7
L 21.6	L 21.6	L -12.8
L 22.3	L 18.0	L 23.1
L -3.4	L 17.9	L 22.3

Tome los valores negativos como una resta al momento de sacar su media.

3. El director de una empresa norteamericana desea invertir en una comunidad de nuestro país y requiere que usted compare los salarios promedio en su planta de Tegucigalpa con los de la competencia que está ubicada en Comayagua.

De los 612 empleados que supervisa en Tegucigalpa, 212 ganan L 520.00 al día, a 65 se les pagan L 1,020.00, otros 150 ganan L 820.00 y otros L 687.00.

La competencia tiene 576 empleados, de los cuales 223 ganan L 522.00, a 53 se les pagan L 1,137.00, otros 135 ganan L 880.00 y, por último, 165 ganan L 721.00.

Haciendo uso de las medidas de tendencia central, escriba un pequeño informe para que el director obtenga detalladamente la información que desea.

4. Una empresa de equipos deportivos está probando el efecto de dos planes publicitarios sobre las ventas de los últimos 4 meses. Dadas las ventas que se ven en la tabla, ¿cuál programa de publicidad parece producir el crecimiento promedio más alto en ventas mensuales?

Mes	Plan 1	Plan 2
Enero	$ 1,657	$ 4,735
Febrero	$ 1,998	$ 5,012
Marzo	$ 2,267	$ 5,479
Abril	$ 3,432	$ 5,589

En plenaria, comparta las respuestas con sus compañeros y tutor

Medidas de dispersión

En nuestro esfuerzo por describir un conjunto de números hemos visto que es de utilidad ubicar el centro del conjunto de datos, pero identificar una medida de tendencia central rara vez es suficiente. Una descripción más completa del conjunto de datos puede obtenerse si se mide que tan dispersos están los datos alrededor de dicho punto central. Esto es precisamente lo que hacen las medidas de dispersión. Indican cuánto se desvían las observaciones alrededor de su media.

Medidas de dispersión: miden qué tanto se dispersan las observaciones alrededor de su media.

Se toman, por ejemplo, los tres conjuntos pequeños de datos que se observan aquí:

Conjunto de datos 1	Conjunto de datos 2	Conjunto de datos 3
0,5,10	4,5,6	5,5,5

Los tres datos tienen una media de cinco, ¿se debe por tanto concluir que los conjuntos de datos son similares? Claro que no, sin embargo, si se informan solo sus medias, sin ver las observaciones, se puede concluir que hay similitud.

Una imprecisión más notoria de los conjuntos de datos resultaría si se compara el grado en el cual se dispersaron las observaciones individuales en cada conjunto de datos o se expandieron alrededor de la media cinco. Las observaciones en el primer conjunto de datos están muy dispersas por encima y por debajo de la media, mientras que las del segundo grupo de datos están comparativamente cerca de este.

El primer conjunto de datos tiene una medida de dispersión mayor que la segunda. El tercer conjunto de datos no tiene dispersión, todas las observaciones son iguales a la media. Sabiendo esto, sería poco probable asumir de manera errónea cualquier similitud en los conjuntos de datos simplemente con base en su media. En ese sentido, las medidas de dispersión son muy útiles e informativas.

Rango o recorrido

El rango es la diferencia entre el mayor y el menor de los datos de una distribución estadística.

Su ventaja es que es fácil de calcular; su desventaja es que considera solo dos de los cientos de observaciones que hay en el conjunto de datos, el resto de las observaciones se ignoran.

Varianza y desviación estándar de una población

La desviación estándar o desviación típica es la raíz cuadrada de la varianza. Es decir, la raíz cuadrada de la media de los cuadrados de las puntuaciones de desviación.

La desviación estándar se representa por σ.

Varianza: es el promedio de las observaciones respecto a su media elevadas al cuadrado.

La varianza para una población σ^2 (se lee sigma al cuadrado) es:

Varianza poblacional

$$\sigma^2 = \frac{(x_1 - \mu)^2 + (x_2 - \mu)^2 + (x_3 - \mu)^2 + \ldots + (x_N - \mu)^2}{N}$$

$$= \frac{\Sigma(X_1 - \mu)^2}{N}$$

[5.6]

Donde: $X_1, X_2, X_3, \ldots, X_N$	Son las observaciones individuales.
μ	Es la media poblacional.
N	Es el número de observaciones.

La desviación estándar σ es:

Desviación estándar poblacional $\quad \sigma = \sqrt{\sigma^2}$

[5.7]

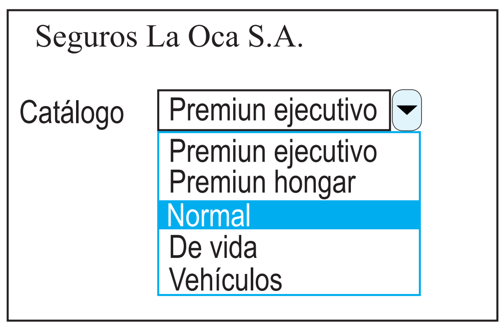
Vale la pena notar que debido a que se está trabajando con una población, la media es μ, no $\overline{X}$, como para una muestra; y el número de observaciones es N, no "n", como para una muestra.

Desviación estándar: es la raíz cuadrada de la varianza, es una medida importante de la dispersión de los datos.

A manera de ilustración:

La empresa Seguros La Oca, S.A., vende cinco pólizas de seguro diferentes del catálogo que se despliega en su página web.

Sus respectivas primas mensuales son de L 900.00, L 801.00, L 500.00, L 610.00 y, L 400.00.



La prima promedio es:

$$\mu = \frac{900 + 801 + 500 + 610 + 400}{5} = 642.2$$

$$\sigma^2 = \frac{(900 - 642.2)^2 + (801 - 642.2)^2 + (500 - 642.2)^2 + (610 - 642.2)^2 + (610 - 642.2)^2}{5}$$

$$= 34{,}319.36$$

A pesar del uso común de la varianza, esta presenta dos problemas: el primero es que se trata de un número muy grande con respecto a las observaciones. Como se puede ver, es varias veces mayor incluso que la observación más grande. Debido a su gran tamaño, con frecuencia la varianza se vuelve difícil de trabajar.

Un problema aún más angustioso es que debido a que las desviaciones son elevadas al cuadrado, la varianza siempre se expresa en términos de los datos originales elevados al cuadrado.

En el caso de Seguros La Oca, S.A., debido a que se evaluaran al cuadrado las desviaciones de la media, entonces se convierten en L 34,319.36 al cuadrado, una unidad de medida que no tiene sentido. En la mayoría de los casos la varianza se expresa en términos que no tienen significado o interpretación lógica.

Sin embargo, ambas interpretaciones pueden resolverse rápidamente, tan solo con hallar la desviación estándar:

$$\sigma = \sqrt{34{,}319.36}$$
$$= \text{L } 185.25$$

¿Por qué aplicamos raíz cuadrada?
Recordemos lo que hacemos al tener una igualdad elevada al cuadrado en ambos lados, como en el ejemplo:
$\sigma^2 = (34{,}319.36)^2$
$\sigma = \sqrt{34{,}319.36}$
$\sigma = 185.25$

Así de fácil se solucionan ambos problemas. Ahora se tiene un número más pequeño con el cual es más fácil trabajar y, más importante aun, ahora está expresado en lempiras, ya que se tomó la raíz cuadrada de los lempiras elevados al cuadrado.

El concepto de desviación estándar es muy importante en los negocios y en la economía. Por ejemplo, en finanzas la desviación estándar se utiliza como medida de riesgo relacionada con varias oportunidades de inversión.

Mediante el uso de la desviación estándar para medir la variabilidad en las tasas de rendimiento ofrecidas por diferentes inversiones, el analista puede medir el nivel de riesgo que tiene cada activo financiero. Generalmente, entre mayor sea la desviación estándar de la tasa de rendimiento de una inversión en particular, mayor será el grado de riesgo.

Veamos el siguiente ejemplo:

Varianza y desviación estándar de una población

Rodrigo Barreiro es gerente de inversiones E.G. Recientemente, Rodrigo estaba interesado en las tasas de rendimiento de los últimos cinco años de dos diferentes fondos mutuos. Corporación Mega mostró, durante un período de cinco años, tasas de rendimiento del 12, 10, 13, 9 y 11 %, mientras que Industrias Dinámicas arrojó 13, 12, 14, 10 y 6 %. Un cliente se acercó a Barreiro y expresó su interés en uno de estos fondos mutuos.

¿Cuál debería escoger Barreiro para su cliente?

Vale la pena destacar que ambos fondos tienen un rendimiento promedio del 11 % debido a que ofrecen el mismo provecho en promedio; una desviación estándar, nos ayudará a despejar dudas:

Para corporación Mega

$$\sigma^2 = \frac{(12 - 11)^2 + (10 - 11)^2 + (13 - 11)^2 + (9-11)^2 + (11 - 11)^2}{5} = 2$$

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

$$= 1.41\%$$

Para Industrias Dinámicas

$$\sigma^2 = \frac{(13 - 11)^2 + (12 - 11)^2 + (14 - 11)^2 + (10 - 11)^2 + (6 - 11)^2}{5} = 8$$

$$\sigma = \sqrt{8}$$

$$= 2.83\%$$

Interpretación: debido a que Corporación Mega presenta menos variabilidad en sus rendimientos y ofrece la misma tasa de rendimiento promedio que Industrias Dinámicas, Mega representa la más segura de las dos inversiones y, por ende, es la oportunidad de inversión preferida.

Varianza y desviación estándar para una muestra

Los ejemplos anteriores se relacionan con la varianza y la desviación estándar para una población. Los símbolos σ^2 y σ son letras griegas típicas de los parámetros.

Rara vez puede calcularse parámetros, en la mayoría de los casos más bien se estimarán tomando una muestra y calculando los datos estadísticos correspondientes. Teniendo esto presente, esta sección analiza la forma en que se calculan estas importantes medidas de dispersión según se relacionan con las muestras.

La varianza y la desviación estándar para una muestra representan medidas de dispersión alrededor de la media, se calculan de manera parecida a aquellas para una población. La varianza de muestra es s^2.

Varianza de la muestra $$s^2 = \frac{\Sigma(X_i - \overline{X})}{n - 1} \qquad [5.8]$$

Y la desviación estándar de la muestra es:

Desviación estándar de la muestra $$s^2 = \sqrt{s^2} \qquad [5.9]$$

Para ilustrar la técnica para establecer estas medidas de dispersión para una muestra, consideremos otro problema que tiene el señor Barreiro en su esfuerzo por ayudar a sus clientes a tomar decisiones sobre inversión.

Ejemplo 5:

Rodrigo Barreiro desea determinar la estabilidad del precio de una acción en particular. Decide basar su juicio en la estabilidad de la desviación estándar del precio de cierre diario de dicha acción. Al revisar las páginas financieras, Barreiro sabe que la acción ha sido transada en la bolsa durante algún tiempo y que hay muchos precios de cierre desde hace varios meses. En lugar de utilizar todos estos precios, Barreiro decide simplificar su aritmética y seleccionar una muestra aleatoria de n = 7 días (aunque 7 probablemente es una muestra muy pequeña, servirá por el momento para los propósitos que se tienen). Barreiro nota que los precios en dólares al cierre son de:

$87, $120, $54, $92, $73, $80, $63.

Solución:

entonces, $\overline{X} = \$81.29$

$$s^2 = \frac{\Sigma(X_i - \overline{X})}{n - 1}$$

$$\sigma^2 = \frac{(87 - 81.29)^2 + (120 - 81.29)^2 + (54 - 81.29)^2 + (92 - 81.29)^2 + (73 - 81.29)^2 + (80 - 81.29)^2 + (63 - 81.29)^2}{7 - 1}$$

$S^2 = 465.9$ cuadrado de dólares

$S = \sqrt{465.9}$

$ = 21.58

Interpretación:

Barreiro ha estimado que la media del precio de cierre de la acción es de $81.29, con una tendencia a variar por encima o por debajo de dicho precio en $21.58. Una explicación adicional del uso e interpretación de la desviación estándar se proporciona más tarde en este capítulo. Sin embargo, se tiene en mente que Barreiro puede interpretar siempre la desviación estándar de $21.58 como una medida de la tendencia de los precios de cierre, que tienden a fluctuar alrededor de su media de $81.29.

Actividad 2

Instrucciones: desarrolle en su cuaderno de trabajo los ejercicios que a continuación se le presentan.

1. Se utilizan dos procesos para producir discos de computadora, pero han surgido problemas respecto a las variaciones en los tamaños de tales discos. Con base en los datos de muestra aquí observados, de ocho tamaños de discos en pulgadas para cada proceso, explique cuál proceso aconsejaría usted si su objetivo es minimizar la desviación en el tamaño alrededor de la media.

Proceso 1		Proceso 2	
3.41	3.22	3.81	3.26
3.74	3.06	3.26	3.79
3.89	3.65	3.07	3.14
3.65	3.33	3.35	3.51

2. Explique con sus propias palabras qué miden la varianza y la desviación estándar. ¿Por qué su cálculo es algo diferente para las poblaciones y las muestras?

3. Un analista de inversiones le sugiere que invierta en RAV Seguridad en lugar de Conteiner Club. Dadas las tasas anuales de rendimiento que se presentan a continuación para una muestra de cada inversión, ¿qué le dice al analista si usted desea minimizar su exposición al riesgo?

RAV Seguridad		Conteiner Club	
15.5 %	3.6%	4.5%	6.2%
21.7	27.2	5.5	7.2
-7.8	2.2	3.5	4.2
-5.0	12.2	4.1	3.9

En plenaria, comparta las respuestas con sus compañeros y tutor

Organización y análisis de datos

Es el proceso de categorización, relación, diagramación, elaboración de matrices y mapas conceptuales que nos permite la reducción de los datos específicos para el análisis.

Datos estadísticos

Son números que pueden ser comparados, analizados e interpretados. El campo del cual son tomados los datos estadísticos se identifica como población o universo.

En un estudio estadístico los métodos que se aplican son:

a. **Recopilación:** de acuerdo con la localización de la información, los datos estadísticos pueden ser internos y externos.

   - Los internos: son los registros obtenidos dentro de la organización que hace un estudio estadístico.
   - Los externos: se obtienen de datos publicados y encuestas.

b. **Organización:** en la organización de los datos, el primer paso es corregir cada uno de los elementos recopilados.

c. **Representación:** hay 3 maneras de presentar un conjunto de datos: mediante enunciados, tablas estadísticas y gráficas estadísticas.

d. **Análisis:** en este proceso los datos estadísticos son estudiados e interpretados, para lo cual frecuentemente se emplean operaciones matemáticas durante el proceso.

   Si una muestra es representativa de una población, se pueden hacer importantes deducciones acerca de esta a partir del análisis de la misma.

   Una muestra es un conjunto de medidas u observaciones tomadas a partir de una población dada.

Análisis de la información

Los beneficiarios juegan el rol principal en el análisis de la información y en la determinación del método adecuado para transmitir los resultados y discutirlos con otros. La información recogida mediante herramientas deberá compilarse o analizarse de alguna manera, ya sea por un software informático o por un tabulado manual de datos.

Aplicación de la estadística en la vida cotidiana

La estadística es, hoy por hoy, una de las ciencias que están más presentes en nuestra sociedad, muchas veces sin darnos cuenta hasta qué punto es esencial el papel que juega en nuestra vida. Un ejemplo de ello podemos encontrarlo en el protagonismo que ha cobrado el profesional de la estadística en la investigación biomédica, aportando la metodología científica necesaria para comprobar, por ejemplo, si un fármaco es efectivo para el tratamiento de una enfermedad concreta. Otro ámbito en el que entender conceptos estadísticos es cada vez más necesario es el caso de los medios de comunicación. Desde ellos cada vez con mayor frecuencia se nos hace llegar información resumida con datos y cifras que en muchas ocasiones son malinterpretados, en el mejor de los casos de forma inconsciente. En el mundo financiero, muchos de los grandes profesionales se han desarrollado desde una formación inicialmente económica y han encontrado en la estadística el complemento necesario para buscar soluciones eficientes a conflictos de índole socioeconómica. Aquí veremos en detalle unos ejemplos de cómo se involucra la estadística en nuestro diario vivir.

Estadística en el lenguaje

La estadística es una rama muy cercana de las matemáticas, pero la utilizamos continuamente en nuestro lenguaje cotidiano y, a veces, darnos cuenta. Podemos reflexionar en frases como las siguientes:

Lenguaje habitual	Lenguaje estadístico
No puede suceder	Improbable
No sucede muy a menudo	No muy probable
Sucede siempre	Muy probable

Al igual que en las frases anteriores, la estadística está muy presente hoy día en la prensa y en la televisión como lenguaje común para presentar la información.

Análisis de anuncios publicitarios

Actualmente cada anuncio publicitario lleva implícitos estudios estadísticos muy específicos de las bondades de los productos, llámense volumen, precio, características, beneficios, etc. Por ejemplo, se dice que el producto "X" tiene 25 % más volumen que el de la competencia o que cuesta 10 % menos que los demás productos que están en el mercado, esto se refiere claramente al precio que tiene este producto; podríamos enumerar muchas cosas más que ofrecen algunos comerciales publicitarios, pero la idea es que usted se dé cuenta viendo algunos anuncios publicitarios cómo la estadística y la mercadotecnia son las que venden los productos.

Los presupuestos familiares

Actualmente los núcleos familiares crean presupuestos para subsistir, destinan dinero a la alimentación, vestimenta, educación, servicios públicos, vivienda, etc. Pero cómohacen este proceso muy sencillo, por medio de datos estadísticos que permiten balancear estos presupuestos y de esa forma llegar a fin de mes con lo necesario para subsistir.

Estadística versus noticias

Es común ver, leer o escuchar muchas malas noticias en los medios de comunicación, ¿pero qué son esas noticias? Son datos estadísticos colocados para generar un eco en la noticia misma, por ejemplo, la cantidad de muertos diarios o mensuales contra la tasa de mortalidad del mes anterior; es muy común encontrar en las noticias enunciados como el siguiente: 10,696 deportados de Estados Unidos en siete meses.

Las estadísticas sirven, por ejemplo para saber si el número de deportados se incrementa o disminuye mensualmente.

En este caso miramos que refleja la cantidad de migrantes que han retornado al país en los últimos 7 meses, y si lo analizamos más podríamos nosotros mismos interpretar que:

$\overline{X} = \dfrac{10{,}696}{7} = 1{,}528$ Entonces 1,528 migrantes retornan al país mensualmente.

Actividad 3

Instrucciones: en su cuaderno de trabajo, resuelva lo que se le solicita a continuación.

Ver en televisión, escuchar por la radio o leer en un diario un comercial o noticia que presente datos estadísticos y describir brevemente que observó en ellos. Recuerde colocar la fuente de donde sacó los datos. Por ejemplo:

¿Conoce la historia de la mujer que perdió la vista al consumir 28 bebidas energizantes?

Una mujer de 26 se ha quedado ciega tras beber 28 latas de bebidas energizantes diarias. Se trata de Lena Lupari, madre de tres niños, y uno de ellos con necesidades especiales. Según ha relatado Lupari en el diario MailOnline, sus tres hijos apenas le dejaban tiempo para hacer nada, por lo que comenzó a ingerir esta bebida que supone un gran aporte de cafeína. Al final esta costumbre se terminó convirtiendo en hábito y solía beber 28 latas al día; la joven asegura que no era consciente del daño que se estaba haciendo a sí misma. Además, ingería 3,000 calorías diarias, alimentándose principalmente a base de pasta y de comida basura. Lupari ha tenido que ser ingresada recientemente de urgencia a un hospital ya que su cerebro se hinchó excesivamente debido a la gran cantidad de esta bebida energética que ingería. Además, tomaba habitualmente pastillas para paliar las migrañas que sufría desde hacía 5 años, como consecuencia de todo esto la mujer ha perdido la visión. Pero los médicos le han advertido que si no comienza a cuidarse y baja de peso, sus problemas de salud serán aún más graves.

Tomado de diario *El Heraldo*

Nombre del comercial, noticia, etc.	Datos estadísticos
¿Conoce la historia de la mujer que perdió la vista al consumir 28 bebidas energizantes?	1 mujer, 26 años, madre de 3 hijos, consumo diario de 28 latas de bebidas energizantes, ingesta de 3,000 calorías por día.
Interpretación	Una mujer consume 3,000 calorías por día cuando lo ideal es que consuma 1,500 calorías diarias. El consumo máximo de estas bebidas debe ser de 2 latas por día, según el cuerpo de la persona. Crean hábito en los consumidores.

En plenaria, comparta las respuestas con sus compañeros y tutor

Conteo

En esta sección aplicaremos los principios de conteo (suma y producto) para encontrar el número total de casos.

Principios fundamentales del conteo

En lo que respecta a técnicas de conteo, tenemos dos principios importantes:

1. El principio de adición
2. El principio de multiplicación

Principio multiplicativo

Si se desea realizar una actividad que consta de **r** pasos, donde el primer paso de la actividad puede ser llevado a cabo de N_1 maneras o formas, el segundo paso de N_2 maneras o formas y el tercer paso de N_r maneras o formas, entonces esta actividad puede ser llevada a efecto de:

N_1 x N_2 x..........x N maneras o formas

El principio multiplicativo implica que cada uno de los pasos de la actividad debe ser llevado a efecto uno tras otro.

Ejemplo 6:

una persona desea construir su casa, para lo cual considera que puede construir los cimientos de cualquiera de dos maneras (concreto o bloque de cemento), mientras que las paredes las puede hacer de adobe, adobón o ladrillo, el techo puede ser de concreto o lámina galvanizada y, por último, los acabados los puede realizar de una sola manera, ¿cuántas maneras tiene esta persona de construir su casa?

Solución:

Considerando que r = 4 pasos
N1= maneras de hacer cimientos = 2
N2= maneras de construir paredes = 3
N3= maneras de hacer techos = 2
N4= maneras de hacer acabados = 1
N1 x N2 x N3 x N4 = 2 x 3 x 2 x 1 = 12 maneras de construir la casa.

El principio multiplicativo, el aditivo y las técnicas de conteo que posteriormente se tratarán nos proporcionan todas las maneras o formas posibles de llevar a cabo una actividad cualquiera.

Principio aditivo

Si se desea llevar a efecto una actividad, la cual tiene formas alternativas para ser realizada, donde la primera de esas alternativas puede ser realizada de M maneras o formas, la segunda alternativa puede realizarse de N maneras o formas... y la última de las alternativas puede ser realizada de W maneras o formas, entonces esa actividad puede ser llevada a cabo de:

$M + N + \dots\dots\dots + W$ maneras o formas

Ejemplos 7:

Una persona desea comprar una lavadora de ropa, para lo cual ha pensado que puede seleccionar entre las marcas Whirlpool, Easy y General Electric. Cuando acude a hacer la compra se encuentra con que la lavadora de la marca Whirlpool se presenta en dos tipos de carga (8 u 11 kilogramos), en cuatro colores diferentes y puede ser automática o semiautomática; mientras que la lavadora de la marca Easy, se presenta en tres tipos de carga (8, 11 o 15 kilogramos), en dos colores diferentes y puede ser automática o semiautomática, y la lavadora de la marca General Electric se presenta en solo un tipo de carga, que es de 11 kilogramos, dos colores diferentes y solo hay semiautomática. ¿Cuántas maneras tiene esta persona de comprar una lavadora?

Solución:

M = Número de maneras de seleccionar una lavadora Whirlpool
N = Número de maneras de seleccionar una lavadora de la marca Easy
W = Número de maneras de seleccionar una lavadora de la marca General Electric

$M = 2 \times 4 \times 2 = 16$ maneras
$N = 3 \times 2 \times 2 = 12$ maneras
$W = 1 \times 2 \times 1 = 2$ maneras
$M + N + W = 16 + 12 + 2 = 30$ maneras de seleccionar una lavadora

¿Cómo podemos distinguir cuándo hacer uso del principio multiplicativo y cuándo del aditivo? Es muy simple, cuando se trata de una sola actividad, la cual requiere para ser llevada a efecto de una serie de pasos, entonces haremos uso del principio multiplicativo, y si la actividad a desarrollar tiene alternativas para ser llevada a cabo, haremos uso del principio aditivo.

Actividad 4

Instrucciones: resuelva en su cuaderno de tareas lo que a continuación se le presenta.

Ejercicio 1

¿Cuántas comisiones integradas por un chico y una chica pueden formarse con 5 chicos y 8 chicas, si cierto chico rehúsa trabajar con dos chicas en particular?

Ejercicio 2

¿De cuántas maneras distintas se pueden distribuir cuatro camisas de diferente color en tres cajones?

Ejercicio 3

Rosa posee 3 blusas distintas, 2 pantalones diferentes y 4 pares de zapatos diferentes. ¿De cuántas maneras distintas puede vestirse utilizando las prendas mencionadas?

Ejercicio 4

¿Cuántas parejas de baile diferentes pueden formarse con 5 niños y 3 niñas?

En plenaria, comparta las respuestas con sus compañeros y tutor

Probabilidad

Probabilidad es la posibilidad numérica de que ocurra un evento. La probabilidad de un evento es medida por valores comprendidos entre 0 y 1, entre mayor sea la probabilidad de que ocurra un evento, su probabilidad asignada estará más próxima a 1, la probabilidad de certeza es 1, la probabilidad de una imposibilidad es 0. Esto se podría expresar así:

P (evento cierto) = 1

P (evento imposible) = 0

Por tanto, $0 \leq P(E_i) \leq 1$

En donde E_i es algún evento.

La probabilidad de que el sol salga mañana es muy alta, muy cercana a 1. La probabilidad de que se apruebe la materia de Matemáticas III sin estudiar, estando en el otro extremo, es muy cercana a cero.

Probabilidad: es la posibilidad numérica, medida entre 0 y 1, de que ocurra un evento.

El proceso que produce un evento es denominado experimento.

Un experimento es toda acción bien definida que conlleva a un resultado único bien definido. Lanzar un dado es un experimento, un resultado bien definido es un número de 1 a 6. Un experimento puede consistir en revisar un producto para determinar si cumple con ciertas especificaciones de fabricación. El resultado es (1) defectuoso o (2) no defectuoso.

El conjunto de todos los posibles resultados para un experimento es el espacio muestral.

El espacio muestral para lanzar un dado es:

SS = (1, 2, 3, 4, 5, 6)

Mientras que el espacio muestral para el experimento de lanzar una moneda al aire es:

SS = (cara, sello)

La probabilidad de que al menos uno de los eventos que están en el espacio muestral ocurra es igual a 1. Si se lanza un dado, el resultado debe ser un número entre 1 y 6, debido a que esto es una certeza puede decirse que:

$$\sum P(E_i) = 1 \qquad [5.10]$$

Introducción a la probabilidad

En la actualidad vemos que la teoría de la probabilidad ocupa un lugar importante en muchos asuntos de negocios. Los seguros y prácticas actuariales se basan firmemente en los principios de la teoría de la probabilidad. Las pólizas de seguro de vida dependen de las tablas de mortalidad, las cuales a su vez se basan en las probabilidades de muerte en edades específicas. Otras tasas de seguros, tales como el seguro de bienes raíces y de automóviles, se determinan de manera similar.

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

Existen tres formas generalmente aceptadas para enfocar la probabilidad:

1. El modelo de frecuencia relativa (a posteriori)
2. El modelo subjetivo
3. El modelo clásico (a priori)

El modelo de frecuencia relativa: utiliza datos que se han observado de forma empírica, registra la frecuencia con que ha ocurrido algún evento en el pasado y estima la probabilidad de que el evento ocurra nuevamente con base en datos históricos.

La probabilidad de un evento con base en el modelo de frecuencia relativa se determina mediante:

Frecuencia relativa $$P(E) = \frac{\textit{Número de veces que ha ocurrido el evento en el pasado}}{\textit{Número total de observaciones}} \quad [5.11]$$

Por ejemplo: si durante el año anterior hubo 50 nacimientos en un hospital local, de los cuales 32 de los recién nacidos eran niñas, el modelo de frecuencia relativa revela que la probabilidad de que el siguiente nacimiento sea una niña es:

$$P(\textit{niñas}) = \frac{\textit{Número de niñas que nacieron el año anterior}}{\textit{Número total de nacimientos}} = \frac{32}{52}$$

Como otro ejemplo tenemos que un importador de cristal recibe envíos de cajas de tres artículos, los datos para las últimas 100 cajas indicaron el número de artículos dañados que había en cada caja, por ejemplo, mostraba que en 40 de las cajas ningún artículo se había dañado mientras que en 12 de las cajas todos los tres artículos se habían roto.

Resultado (E_i) Números de defectos	Número de cajas	P (E_i)
0	40	40 ÷ 100 = 0.40
1	27	27 ÷ 100 = 0.27
2	21	21 ÷ 100 = 0.21
3	12	12 ÷ 100 = 0.12
	100	**1.00**

En el pasado, 21 de las 100 cajas contenían exactamente dos artículos dañados. Entonces el modelo de frecuencia relativa asignaría una probabilidad de que dos artículos en cualquier caja dada estuvieran dañados, así P(2) = 21 ÷ 100 = 0.21. La probabilidad para cada resultado individual se muestra en la última columna, la cual suma (1) uno.

Un problema común con el modelo de frecuencia relativa resulta cuando se hacen estimaciones con un número insuficiente de observaciones. Por ejemplo, se sabe que dos vuelos de una aerolínea en los que se hicieron registros el año pasado estuvieron retrasados al llegar a sus destinos. Por tanto, se concluye que el vuelo que se abordará el próximo mes, en la misma aerolínea, también estará retrasado. Aunque tales inferencias son comunes, no existen suficientes datos como para sacar tal conclusión y se deben evitar las decisiones basadas en tales inferencias.

En muchos casos los datos pasados no se encuentran disponibles, por tanto, no es posible calcular la probabilidad a partir del desempeño anterior. La única alternativa es estimar la probabilidad con base en nuestro mejor criterio. El modelo subjetivo requiere establecer la probabilidad de algún evento con base en la mejor evidencia disponible. En muchos casos esto puede ser apenas una conjetura hecha sobre cierta base.

El modelo subjetivo: se utiliza cuando se desea asignar una probabilidad a un evento que nunca ha ocurrido, la probabilidad de que una mujer sea elegida como presidenta de Honduras es un ejemplo, debido a que no hay datos sobre los cuales confiar, se deben analizar las opiniones y creencias para obtener una estimación subjetiva.

El modelo clásico: este método es el que se relaciona con mayor frecuencia con las apuestas y juegos de azar. La probabilidad clásica de un evento E se determina mediante:

$$P(E) = \frac{\text{Número de formas en las que puede ocurrir un evento}}{\text{Número total de posibles resultados}} \tag{5.12}$$

Aun sin conocer a fondo la probabilidad clásica, se puede estar consciente de que la probabilidad de obtener una cara en un solo lanzamiento de moneda es de la mitad. Esto puede ilustrarse utilizando la fórmula (5.12) así:

$$P(cara) = \frac{\text{Número de formas en las que puede ocurrir un evento}}{\text{Número total de posibles resultados}} = \frac{1}{2}$$

Existe solo una forma en que puede ocurrir el evento (se obtiene cara) y existen dos posibles resultados (una cara y un sello). De igual forma la probabilidad de sacar un 3 con un dado de seis caras es:

$$P(3) = \frac{\text{Número de formas en las que puede ocurrir un evento}}{\text{Número total de posibles resultados}} = \frac{1}{6}$$

Actividad 5

Instrucciones: resuelva en su cuaderno de trabajo los ejercicios de probabilidad que a continuación se le presentan:

1. ¿Cuál es la probabilidad de sacar un as en una baraja de 52 cartas? Explique su respuesta.

2. Cite tres ejemplos de negocios para cada uno de los modelos de probabilidad.

3. El año anterior las ventas semanales de petunias fueron "bajas" durante 16 semanas, "considerables" durante 27 semanas y "altas" el resto de las semanas. Cuál es la probabilidad de que las ventas de esta semana sean:
   a. Considerables
   b. Bajas
   c. Altas
   d. Por lo menos considerables (explique)

Actividad metacognitiva

Instrucciones: con base en lo aprendido, conteste las siguientes preguntas.

1. ¿Qué estrategias podemos utilizar para diferenciar las medidas de tendencia central con las medidas de dispersión?

____

____

____

2. ¿En qué situaciones de la vida cotidiana podemos hacer uso de los principios del conteo?

____

____

____

3. ¿Cuál es la importancia de conocer las probabilidades de que suceda un evento?

____

____

____

4. ¿Considera usted que adquirió nuevos aprendizajes a la estudiar los temas de esta unidad? ¿Puede mencionar cuáles?

____

____

____

5. ¿Cuáles son los contenidos de la presente unidad que le presentaron mayor dificultad? ¿Por qué?

____

____

____

Autoevaluación

I. Tipo selección única

Instrucciones: a continuación se le presentan una serie de ejercicios de evaluación del aprendizaje de esta unidad. Su trabajo consiste en desarrollar con la mayor honestidad posible, sin copiar del libro. Recuerde que esta actividad le dará información sobre su progreso educativo. Al terminar, confronte sus respuestas con las soluciones que se encuentran en el anexo 2.

1. ¿Cuál de las siguientes opciones es la medida de tendencia central que usualmente conocemos como promedio?
   a. Media
   b. Mediana
   c. Moda
   d. Rango

2. ¿Cuál de las siguientes opciones corresponde al conjunto de medidas a las que pertenece la desviación estándar?
   a. Posición
   b. Progresión
   c. Dispersión
   d. Forma

3. ¿Qué es la varianza?
   a. Es la diferencia entre el mayor y menor dato estadístico
   b. Es el cuadrado del promedio de cada observación respecto a sus medias
   c. Es la raíz cuadrada de la suma de todas las observaciones
   d. Es el cambio porcentual en una serie de números reales

4. ¿Qué tipo de datos estadísticos se obtienen en las publicaciones de encuestas?
   a. Continuos
   b. Discretos
   c. Internos
   d. Externos

5. ¿Qué es una probabilidad?
   a. Es la medida entre 0 y 1 de que ocurra un evento
   b. Es el cuadrado de una muestra observaciones
   c. Es la información que debe cumplirse y organizarse
   d. Es todo aquello que toma un valor finito dentro de una observación

6. ¿Cuál de las siguientes opciones es una representación de datos estadísticos?
   a. Números
   b. Gráficos
   c. Software
   d. Registros

7. ¿Cuál de las siguientes opciones es aquella información que puede ser comparada, analizada e interpretada?
   a. Dato
   b. Matriz
   c. Muestra
   d. Universo

8. ¿Cuál de las siguientes opciones corresponde a la frecuencia de datos en que la moda es 1 y 5?
   a. 1, 3, 3, 4, 4, 4, 5, 5
   b. 0, 1, 3, 3, 5, 5, 7, 8
   c. 1, 1, 1, 4, 4, 5, 5, 5
   d. 1, 1, 3, 3, 5, 5, 6, 6

9. Observe la imagen y conteste: ¿cuál es la probabilidad de que la señorita UVG se gane el vehículo si escoge el sobre A?

   a. P(A) = 1/1
   b. P(A) = 1/2
   c. P(A) = 2/2
   d. P(A) = 2/1

10. ¿Cuál de las opciones corresponde a la mediana de las tasas de ahorro en el siguiente caso?

Tasas de ahorro

Servicio Financiero	Tasa de ahorro (%)
Cuentas corrientes	3.01
Socios comerciales	2.96
Depósitos a plazo fijo 6 meses	3.25
Depósitos a plazo fijo 1 año	3.51
Depósitos a plazo fijo 2.5 años	4.25
Depósitos a 5 años	5.46

a. 2.73
b. 3.38
c. 3.74
d. 4.23

II. Parte práctica

Instrucciones: trabaje en forma clara y ordenada en lo que se le pide a continuación:

1. Las utilidades obtenidas por una compañía constructora en cuatro proyectos fueron de 3 %, 2 %,4 % y 6 %, respectivamente. ¿Cuál es la media geométrica de las utilidades de la compañía?

2. Calcule la mediana de las alturas de los jugadores de un equipo de baloncesto. Las alturas se le proporcionan en la siguiente tabla:

Altura (metros)	Número de jugadores
1.70 - 1.75	1
1.76 - 1.80	3
1.80 - 1.85	4
1.85 - 1.90	8
1.90 - 1.95	5
1.95 - 2.00	2

3. Calcule la varianza en este caso: un pediatra obtuvo la siguiente tabla sobre los meses de edad de 50 niños que atendió en el momento de caminar por primera vez:

Meses	Niños
9	1
10	4
11	9
12	16
13	11
14	8
15	1

4. Calcule la desviación estándar de la distribución:

9, 3, 8, 8, 9, 8, 9, 18

5. La probabilidad de que un hombre viva 20 años es ¼ y la de que su mujer viva 20 años es 1/3. Calcule la probabilidad de que ambos vivan 20 años.

6. Una pareja que se casará ha juntado dinero para la prima de su casa en la ciudad de San Pedro Sula. En la colonia Jardines del Valle les ofrecen un modelo económico de apartamento en un condominio, en la residencial Monte Fresco les ofrecen un modelo económico de residencial: un estilo californiano y un provenzal (francés). ¿Cuántas alternativas diferentes de vivienda le ofrecen a la pareja?

7. Calcule la moda de una distribución estadística que viene dada por la siguiente tabla:

Clases	F1
10 - 15	3
15 - 20	5
20 - 25	7
25 - 30	4
30 - 35	2

8. Calcule la frecuencia del siguiente caso: al lanzar una moneda de 50 centavos 100 veces, el escudo cayó al frente de nuestra vista 45 veces.

9. Calcule la media de la distribución estadística que viene dada por la siguiente tabla:

Clases	F1
10 - 20	6
20 - 30	10
30 - 40	14
40 - 50	8
50 - 60	4

10. Complete la siguiente tabla y calcule la desviación típica en el siguiente caso:

Datos	x_i	f_i	$x_i \cdot f_i$	$X_i^2 \cdot f_i$
10 - 20		1		
20 - 30		8		
30 - 40		10		
40 - 50		9		
50 - 60		8		
60 - 70		4		
60 - 70		2		