|  |  |
| --- | --- |
| Título del guion | La estadística y la probabilidad |
| Código del guion | MA\_11\_06\_CO |
| Descripción | La estadística y la probabilidad se han convertido en herramientas fundamentales para desarrollar estudios en numerosos campos de trabajo y de investigación. A partir del diseño experimental, la recolección y análisis de información, y la interpretación de los resultados obtenidos, la estadística permite tomar decisiones a nivel académico, empresarial y científico |

[SECCIÓN 1] **1 El análisis estadístico**

Los conceptos de población y muestra son muy importantes al involucrarse en el campo del análisis estadístico. De acuerdo con la claridad que se tenga de cada uno de ellos, se genera confiabilidad en las inferencias que se construyan.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Definición de población, marco muestral y muestra** |
| **Contenido** | La **población** es un conjunto de elementos sobre los que se desea  realizar una inferencia o construir una conclusión.  Un **marco muestral** es una lista de unidades de la población disponibles para obtener información.  La **muestra** corresponde a los elementos seleccionados de la población que aportan la información que permite construir inferencias acerca de la población. |

En muchas ocasiones el marco muestral no es igual a la población; en los casos en los cuales la población es infinita, muy grande o imposible de contactar es necesario recurrir a listas o bases de datos existentes de elementos de la población que faciliten el contacto para obtener la información requerida.

La muestra debe garantizar dos propiedades básicas que se reflejan en la credibilidad e imparcialidad de los resultados obtenidos:

* **Aleatoriedad**,es el proceso de seleccionar una muestra de una población garantizando imparcialidad, razón por la cual se debe recurrir a métodos en los cuales el investigador no influya en la selección de los individuos. El proceso al cual se recurre con mayor frecuencia es al de la **generación de números aleatorios** para cada uno de los elementos del marco muestral, para luego seleccionar aquellos a los cuales se les asignó los mayores o los menores de acuerdo con el criterio establecido por el investigador.
* **Representatividad**: cuando se define claramente una población objetivo se considera si existen subgrupos que son diferentes, unos de otros, en su comportamiento y que pertenecen a la misma población. La muestra que se va a seleccionar debe garantizar que todos aquellos subgrupos estén representados; esta característica se relaciona con la cobertura de la muestra sobre la población.

|  |  |
| --- | --- |
| **Recuerda** | |
| **Contenido** | Las **variables** son aquellas características que se quieren observar en la población y que por lo general se derivan de los objetivos o de las hipótesis planteadas en el trabajo de investigación. Al momento de definirlas es fundamental que se considere el **tipo de variable**.   * Si se quieren observar características relacionadas con opiniones, gustos, preferencias, etc., la variable se llama **cualitativa** * Si la variable se puede medir en una escala numérica se llama **cuantitativa**.   Estos conceptos deben ser claros al momento de diseñar el proyecto de investigación ya que los métodos para analizar cada tipo de variable son diferentes. |

|  |  |
| --- | --- |
| **Profundiza: recurso nuevo** | |
| **Código** | MA\_11\_04\_CO\_REC10 |
| **Título** | La estadística |
| **Descripción** | Interactivo que muestra conceptos básicos de estadística |

[SECCIÓN 2] **1.1 El análisis de variables cualitativas**

Una variable cualitativa se puede considerar como una pregunta en la cual sus respuestas miden cualidades, características, preferencias, gustos: género de una persona, estado civil, estrato al cual pertenece, intención de voto para determinadas elecciones, preferencia por alguna comida; o preguntas cuyas respuestas se pueden clasificar en rangos preestablecidos (nivel de satisfacción medido como: totalmente satisfecho, satisfecho, medianamente satisfecho o insatisfecho).

[SECCIÓN 3] **1.1.1 La proporción**

El primer criterio para analizar los resultados de una variable cualitativa corresponde a la proporción de la muestra que se encuentra en cada uno de los rangos de respuesta.

Para identificar la proporción es importante tener en cuenta la construcción de una tabla de frecuencias.

|  |  |
| --- | --- |
| **Recuerda** | |
| **Contenido** | Una tabla de frecuencias corresponde al conteo de las respuestas en los rangos correspondientes. Esta tabla representa el primer informe resumido del comportamiento de la variable cualitativa a analizar.  En variables cualitativas esta tabla es única, por lo cual no se genera incertidumbre en su interpretación. |

Ejemplo

El nuevo gerente de una cadena de almacenes quiere determinar si existe una tendencia en la forma de pago de sus clientes. Para ello, toma una muestra de 80 compradores que llegaron a una de las tiendas y examina si cada uno de ellos hace su pago en efectivo, con tarjeta o con bonos. La siguiente tabla muestra los resultados obtenidos por el gerente de la tienda.

|  |  |  |  |
| --- | --- | --- | --- |
| **Modalidad de pago** | ***f*** | ***fr*** | **%** |
| Efectivo | 10 | 10/80 = 0,125 | 12,5 |
| Tarjeta | 42 | 42/80 = 0,525 | 52,5 |
| Bonos | 20 | 20/80 = 0,25 | 25 |
| Otro | 8 | 8/80 = 0,1 | 10 |
| TOTAL | 80 | 80/80 = 1 | 100 |

La variable a analizar corresponde a *Forma de pago*, por lo cual los rangos de respuesta son: Efectivo, Tarjeta, Bonos, Otro.

**La frecuencia (*f* )**, corresponde al número de veces que se repite cada respuesta en la muestra estudiada. La **frecuencia relativa (*fr*)** la cual corresponde al resultado de dividir la frecuencia *f* y el tamaño de la muestra. Se puede expresar como fracción o como decimal.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Definición de proporción** |
| **Contenido** | La proporción se define como el cociente entre el conteo de uno de los rangos de respuesta y el total poblacional.  Por tanto:  MA\_11\_06\_001  donde *X* es el conteo de individuos que seleccionaron un rango de respuesta y *N* es el total de la población.  La proporción se puede expresar de forma de fracción, de decimal o de porcentaje.  En una tabla de frecuencias, la proporción corresponde a la frecuencia relativa. |

En el ejemplo, la proporción de personas que hicieron sus pagos en efectivo es de 1/8 o de 1 a 8, o de 0,125o del 12.5%.Est quiere decir que 1 de cada 8 clientes hacen sus pagos en efectivo.

Este valor, en variables cualitativas, corresponde a la **probabilidad relativa**. Es decir, la probabilidad de que un cliente pague sus compras en efectivo es de 1/8 o de 0,125*.*

[SECCIÓN 3] **1.1.2 La moda**

Una de las principales pretensiones que hay dentro del análisis estadístico corresponde a la identificación de tendencias que determinen el comportamiento de la población. La medida correspondiente a las variables cualitativas, y que busca este fin, es **la moda**.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Definición de moda** |
| **Contenido** | La moda de un conjunto de datos cualitativos es el rango de respuesta de la variable con mayor frecuencia.  Corresponde al dato con mayor frecuencia en la muestra |

Para el ejemplo del gerente de la tienda, se tiene que la moda es el pago con tarjeta. En este caso se puede afirmar que la tendencia, dentro de los clientes de la tienda, es realizar el pago con tarjeta, cuya proporción es de 42 a 80, o del 0,52, o del 52,5%.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Cuándo la moda no es representativa** |
| **Contenido** | Una moda significativamente cercana a los demás rangos de la variable no es representativa. Es decir que cuando varios valores obtenidos en la muestra tienen valores cercanos a la moda, sugieren que la moda no es significativamente diferente a las demás. |

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Cuando la moda no es única** |
| **Contenido** | Puede ser que dos valores tengan la misma proporción y corresponda a la más alta, por lo que la variable es bimodal. Es decir, es posible obtener dos valores que tengan la frecuencia mayor en la muestra. En este caso no se puede hablar de la tendencia de población hacia un valor determinado.  En los casos en los cuales la mayoría, si no todos, de los rangos de respuesta tengan el mismo valor máximo, la variable es amodal. Es decir, si todos los valores de variable son diferentes, no existe una tendencia de acumulación en la población. |

[SECCIÓN 3] **1.1.3 El análisis de gráficos**

Uno de los criterios de mayor utilidad, estadísticamente hablando para determinar tendencias, construir inferencias o pronósticos futuros del comportamiento de la población respecto a una variable corresponde al **comportamiento gráfico de la variable**.

|  |  |
| --- | --- |
| **Recuerda** | |
| **Contenido** | Los gráficos que representan mejor el comportamiento de una variable cualitativa son: el diagrama circular y el diagrama de barras. |

Para el ejemplo, a continuación se muestran varios gráficos que representan la variable Forma de pago de los clientes de la tienda.

|  |  |
| --- | --- |
| **Imagen (fotografía, gráfica o ilustración)** | |
| **Código** | MA\_11\_06\_CO\_IMG01 |
| **Descripción** | Gráfica circular de la variable cualitativa forma de pago |
| **Código Shutterstock (o URL o la ruta en AulaPlaneta)** |  |
| **Pie de imagen**  **Inferior** | Hay una tendencia de los clientes de la tienda a realizar sus pagos usando tarjetas. Esta proporción corresponde a casi el 53%.  En esta gráfica los porcentajes se han aproximado. |

|  |  |
| --- | --- |
| **Imagen (fotografía, gráfica o ilustración)** | |
| **Código** | MA\_11\_06\_CO\_IMG02 |
| **Descripción** | Diagrama de barras de la variable cualitativa forma de pago |
| **Código Shutterstock (o URL o la ruta en AulaPlaneta)** |  |
| **Pie de imagen**  **Inferior** | En estos gráficos es importante recordar que el eje horizontal no corresponde a una recta numérica razón por lo cual, las barras deber estar separadas y el orden en que se ubiquen no es específico |

|  |  |
| --- | --- |
| **Practica: recurso nuevo** | |
| **Código** | MA\_11\_04\_CO\_REC20 |
| **Título** | Analiza gráficas estadísticas |
| **Descripción** | Actividad para analizar información a partir de la representación gráfica de datos |

|  |  |
| --- | --- |
| **Practica: recurso nuevo** | |
| **Código** | MA\_11\_04\_CO\_REC30 |
| **Título** | Resuelve situaciones problema con la información del gráfico |
| **Descripción** | Actividad para resolver situaciones problema extractando información de un gráfico estadístico |

[SECCIÓN 2] **1.2 El análisis de variables cuantitativas**

Uno de los primeros resúmenes de las variables cualitativas corresponde a la tabla de frecuencias. Este tipo de análisis se conoce regularmente como de “datos agrupados”, ya que el objetivo es conformar rangos, grupos o intervalos que agrupen un cierto número de datos.

Por ejemplo, la siguiente tabla resume los resultados obtenidos por rangos de edad de una muestra de 130 aspirantes a ser gerente deportivo de un reconocido club de la ciudad.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Edad en años | *f* | *fr* | *F* | *Fr* | *mi* |
| 25 a 30 | 7 | 0,05 | 7 | 0,05 | 27,5 |
| 30 a 35 | 12 | 0,09 | 19 | 0,14 | 32,5 |
| 35 a 40 | 25 | **0,19** | 44 | 0,33 | 37,5 |
| 40 a 45 | 32 | 0,25 | 76 | **0,58** | 42,5 |
| 45 a 50 | 26 | 0,20 | 102 | 0,78 | 47,5 |
| 50 a 55 | 18 | 0,14 | 120 | 0,92 | 52,5 |
| 55 a 60 | 10 | 0,08 | 130 | 1,00 | 57,5 |
| Total | 130 | 1,00 |  |  |  |

De la tabla se puede ver que: el 19% de los aspirantes tiene entre 35 y 40 años; el 58 % de los aspirantes al cargo tiene 45 años o menos. Vale la pena aclara que si un gerente tiene 45 años, debe ser incluido en el intervalo de 40 a 45 años.

[SECCIÓN 3] **1.2.1 Las medidas de tendencia central: media, mediana y moda**

Las medidas de tendencia central pretenden encontrar un punto centro o punto de acumulación de los datos. Las principales medidas de tendencia central son: el promedio aritmético o media, la mediana y la moda.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Definición de la media aritmética** |
| **Contenido** | La **media** **aritmética** de un conjunto de datos se define como el promedio aritmético de ellos. Generalmente se denota por  MA\_11\_06\_002    Para calcular la media: si  *x*1*, x*2*, . . . , xn*  es un conjunto de datos, entonces  MA\_11\_06\_003 |

El promedio se puede definir como el valor que representa al conjunto de datos respecto a una variable.

Se considera la media como el valor que se espera tenga un individuo de la población. Es por esto que se usa como elemento fundamental de pronóstico.

Ejemplo

Los tiempos de reacción, en minutos, de 15 personas al habérseles aplicado un estímulo externo son:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 27 | 26 | 8 | 21 | 21 |
| 14 | 6 | 18 | 17 | 19 |
| 3 | 21 | 10 | 10 | 5 |

La media de los tiempos es:

MA\_11\_06\_004

Es decir que el tiempo medio de reacción de una persona de dicha población es 15 minutos.

Se puede afirmar que, si una persona se somete al mismo estímulo, el tiempo que tardará en reaccionar estará alrededor de 15 minutos.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Definición de la mediana** |
| **Contenido** | La **mediana** se define como el dato que divide a un conjunto de datos en dos partes porcentualmente iguales. Se nota como  MA\_11\_06\_005    Para calcular la mediana es necesario ordenar el conjunto de datos de menor a mayor  MA\_11\_06\_006    Luego, ubicar el punto o valor que está en el centro de ellos; para encontrar la mediana se tienen dos casos:  Caso 1. Si el número de datos, *n,* es impar  MA\_11\_06\_007    Caso 2. Si el número de datos, *n,* es par  MA\_11\_06\_008    En este caso, la mediana es el punto medio entre las dos observaciones que se ubican en el centro. |

Para el ejemplo de los tiempos de reacción al estímulo se considera el caso 1, ya que el número de datos es impar, entonces se tiene:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 3 | 5 | 6 | 8 | 10 | 10 | 14 | **17** | 18 | 19 | 21 | 21 | 21 | 26 | 27 |

El valor de la mediana corresponde al dato ubicado en la posición

MA\_11\_06\_009

que para el caso corresponde a 17 minutos.

Es decir que el 50% de las personas tuvieron tiempos de reacción menores que 17 minutos, mientras que el restante 50% de las personas tuvieron tiempos de reacción mayores que 17 minutos.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Definición de la moda** |
| **Contenido** | La **moda** se define como el dato que más se repite dentro del conjunto. Se denota  MA\_11\_06\_010  Cuando se tiene un gran número de datos se puede presentar con mayor frecuencia la existencia de más de una moda. En algunos casos se tiene que la moda no existe. |

|  |  |
| --- | --- |
| **Practica: recurso aprovechado** | |
| **Código** | MA\_11\_04\_CO\_REC40 |
| **Título** | Compara dos series de datos con las medidas estadísticas |
| **Descripción** | Interactivo que permite practicar el cálculo de las medidas de tendencia central y de posición |

|  |  |
| --- | --- |
| **Practica: recurso nuevo** | |
| **Código** | MA\_11\_04\_CO\_REC50 |
| **Título** | Halla las medidas de tendencia central |
| **Descripción** | Actividad para resolver situaciones problema que involucran medidas de tendencia central |

[SECCIÓN 2] **1.3 Las medidas de dispersión: rango, varianza y desviación estándar**

Las medidas de dispersión son valores numéricos que miden la dispersión o variabilidad entre los datos. Estas medidas se consideran como un criterio para determinar la cercanía de las observaciones.

* Si los datos están relativamente cerca unos de otros, respecto a la escala en la cual se midieron, las medidas de dispersión toman valores pequeños.
* Si por el contrario los datos están relativamente lejanos unos de otros, las medidas de dispersión toman valores numéricos grandes.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Definición de rango** |
| **Contenido** | El **rango** de un conjunto de datos es la diferencia numérica entre el dato mayor del conjunto y el dato menor.  Rango = *DM* – *Dm*  El rango se puede considerar como el intervalo cerrado en el que se encuentra la totalidad de los datos; entre más pequeño sea su valor, más cercanos estarán los datos. |

La **varianza** es una medida que pretende establecer la cercanía de cada uno de los datos respecto a la media. Se utiliza como medida de confiabilidad del promedio.

Si su valor es pequeño, la mayoría de los datos está cerca entre sí, por lo cual el promedio es un buen individuo típico del conjunto y se constituye en un muy buen pronóstico para la variable.

Si su valor es grande, los datos no son tan cercanos, existe mucha dispersión entre ellos, por lo cual el promedio no es un buen individuo típico. Si se usará como pronóstico estaría cargado de incertidumbre.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Definición de desviación** |
| **Contenido** | Primero, es necesario calcular las desviaciones de cada uno de los datos.  Una **desviación** *d*1se define como la distancia que hay de cada dato al valor de la media.  MA\_11\_06\_011  Se puede concluir que si el valor de la desviación es negativo, entonces el dato correspondiente es menor que la media, y si la desviación es positiva, entonces el dato idóneo es mayor que la media. |

Ya que la media o promedio es un punto de equilibrio de los datos, la suma de las desviaciones siempre es cero.

Es por ello que para determinar una distancia promedio, se usa el valor de las desviaciones cuadradas, de modo que la suma no sea cero.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Definición de Varianza muestral** |
| **Contenido** | La **varianza muestral** de un conjunto de datos *x1* , *x2*, *. . .* , *xn*, notada como *S2*, se define como la media aritmética o promedio de los cuadrados de las desviaciones.  Se calcula así:  MA\_11\_06\_012 |

Ejemplo

Se midió el tiempo, en minutos, que tardaron dos medicamentos diferentes en aliviar una cierta dolencia. En cada uno de los siguientes casos se registraron los tiempos para 5 personas.

**Medicamento 1:** 1, 2, 1, 0,5, 0,5.

**Medicamento 2:** 3, 2, 0,5, 0,2, 1,3.

El tiempo medio de reacción es:

**Medicamento 1:**

MA\_11\_06\_013

**Medicamento 2:**

MA\_11\_06\_014

Para el caso del medicamento 1 se espera que el tiempo de reacción de un paciente al que se le administre sea de 1 minuto, mientras que para el caso del medicamento 2 sea de 1,4 minutos.

Las varianzas para cada caso son:

**Medicamento 1**

MA\_11\_06\_015

**Medicamento 2**

MA\_11\_06\_016

De los resultados se puede afirmar que el medicamento 1 tiene menor variabilidad, es decir que el tiempo de reacción del individuo frente a este estará cercano a un minuto mientras que para el medicamento 2 la variabilidad es mayor, por tanto, el tiempo de reacción no estará tan cerca de 1,4 minutos.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Definición de desviación estándar** |
| **Contenido** | La desviación estándar muestral, notada como *S*, es la raíz cuadrada positiva de la varianza:  MA\_11\_06\_017 |

Esta medida es la que usualmente se utiliza en la interpretación de los datos, sin embargo, la medida que permite generar análisis estadísticos formales es la varianza.

Para el caso de los medicamentos, se tiene que:

**Medicamento 1**

MA\_11\_06\_018

**Medicamento 2**

MA\_11\_06\_019

|  |  |
| --- | --- |
| **Practica: recurso nuevo** | |
| **Código** | MA\_11\_04\_CO\_REC60 |
| **Título** | Resuelve situaciones problema que involucran medidas de variabilidad |
| **Descripción** | Actividad para resolver situaciones problema que involucran medidas de dispersión |

[SECCIÓN 3] **1.3.1 El diagrama de cajas y bigotes. Puntos inusuales**

El diagrama de cajas es un resumen gráfico en el que se describen varias de las características más destacadas de un conjunto de datos.

Algunas de las características que se muestran en un diagrama de cajas son:

* La identificación de datos inusuales.
* La dispersión de los datos con respecto a la mediana.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Definición de cuartíles** |
| **Contenido** | Los cuartiles son valores de la variable que dividen el conjunto ordenado de datos en cuatro partes iguales. Cada una de estas partes contiene el 25% del total de los datos.  El primer cuartil, *Q*1 es el número que a lo sumo el 25% de los datos es menor que él.  El segundo cuartil, *Q*2, es la mediana.  El tercer cuartil, *Q*3, es un número que a lo más el 75% de los datos es menor que él.  Para calcular los cuartiles se debe encontrar la mediana, es decir *Q*2; con los datos menores o iguales a *Q*2 se calcula una nueva mediana que corresponde a *Q*1 y con los datos mayores o iguales a *Q*2 se calcula otra nueva mediana que corresponde a *Q*3. |

El procedimiento para construir un diagrama de cajas es:

* Calcular el valor de los cuartiles. Luego, ubicarlos en una recta numérica y construir rectángulos cuyas bases están formadas por las distancias que hay entre cuartiles
* Después se define una medida de dispersión, llamada *rango intercuartílico*, relacionando los cuartiles uno y tres.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Definición de rango intercuartílico** |
| **Contenido** | El **rango intercuartílico** o cuarta dispersión es la diferencia entre el cuartil tres y el cuartil uno  Rango intercuartílico = *Q*3 – *Q*1  Este rango contiene el 50% de los datos de la muestra. |

* Para la construcción de los bigotes se toman varios criterios de acuerdo con el tipo de variable o con las intenciones del investigador. Sin embargo, en la mayoría de los casos la longitud de los bigotes corresponde a:

1,5 × (rango intercuartílico) = 1,5 × (*Q*3 – *Q*1)

* Una vez determinada esta medida de dispersión se construye una línea desde el valor del primer cuartil hasta el valor

*Q*1 – 1,5 × (*Q*3 – *Q*1)

y de la misma forma, desde el tercer cuartil hasta

*Q*3 + 1,5 × (*Q*3 – *Q*1)

Los datos que están incluidos en esta longitud son los que se consideran dentro de un rango permitido, ya que están cubiertos dentro del rango calculado.

Finalmente, se construyen los segundos bigotes conservando la longitud calculada en el paso anterior.

Es decir, el segundo bigote izquierdo se traza desde

*Q*1 – 1,5 × (*Q*3 – *Q*1) hasta *Q*1 – 2 × 1,5 × (*Q*3 – *Q*1)

Mientras que el segundo bigote derecho se traza desde

*Q*3 + 1,5 × (*Q*3 – *Q*1) hasta *Q*3 + 2 × 1,5 × (*Q*3 – *Q*1)

|  |  |
| --- | --- |
| **Imagen (fotografía, gráfica o ilustración)** | |
| **Código** | MA\_11\_06\_CO\_IMG03 |
| **Descripción** | Diagrama de cajas y bigotes |
| **Código Shutterstock (o URL o la ruta en AulaPlaneta)** | OJO ILUSTRADOR LAS Q van todas en mayuscula itálica |
| **Pie de imagen**  **Inferior** | En algunas ocasiones puede suceder que coincidan dos de los valores de los cuartíles, por lo cual, existirá tan solo una caja.  Es posible que el valor de los tres cuartiles sea el mismo. Es decir que no existirán cajas. Solamente una línea que corresponde a la mediana. En este caso la varianza es muy pequeña |

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Definición de datos o puntos inusuales** |
| **Contenido** | Un dato se llama **inusual o atípico** si está fuera del diagrama de cajas y bigotes.  Son valores que influyen de manera significativa en el cálculo de las medidas de tendencia central y de dispersión que caracterizan la variable. Estos datos atípicos corresponden a valores mal tomados en el momento de la obtención de la información, o a individuos dentro de la muestra o la población que no son representativos y que por el contrario se alejan mucho de la característica que se está midiendo. |

Ejemplo

El reporte del número de llamadas que recibió un centro de atención al cliente de un banco, durante los últimos 20 días es:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 194 | 181 | 153 | 159 | 217 | 172 | 168 | 220 | 150 | 156 |
| 123 | 176 | 212 | 202 | 219 | 205 | 638 | 224 | 180 | 176 |

Elaborar el diagrama de cajas y bigotes para esta situación.

Los valores necesarios para la construcción del diagrama son:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Cuartil 1 | Mediana | Cuartil 3 | R. Intercuartílico | Longitud del bigote |
| 165,75 | 180,5 | 213,25 | 47,5 | 71,25 |

El valor mínimo del diagrama es 38 y el valor máximo es 355.75. El diagrama correspondiente es:

|  |  |
| --- | --- |
| **Imagen (fotografía, gráfica o ilustración)** | |
| **Código** | MA\_11\_06\_CO\_IMG04 |
| **Descripción** | Diagrama de cajas y bigotes para el número de llamadas |
| **Código Shutterstock (o URL o la ruta en AulaPlaneta)** | 50  100  150  200  250  300  350  400  450  500 |
| **Pie de imagen**  **Inferior** | Ya que existe un valor mayor que 355,75, entonces, existe un dato inusual en la muestra. Este valor corresponde a 638 |

|  |  |
| --- | --- |
| **Practica: recurso nuevo** | |
| **Código** | MA\_11\_04\_CO\_REC70 |
| **Título** | Relaciona el diagrama de cajas con la muestra correspondiente |
| **Descripción** | Actividad para relacionar diagramas de cajas y de puntos con el conjunto de datos correspondiente |

[SECCIÓN 3] **1.3.2 El coeficiente de variación**

El coeficiente de variación se puede considerar como una medida de dispersión o de variabilidad. Su propósito es comparar la media con la variabilidad que presenta la variable.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Definición de coeficiente de variación** |
| **Contenido** | El **coeficiente de variación** se define como la relación porcentual que existe entre la desviación estándar y la media.  MA\_11\_06\_020 |

El coeficiente de variación no tiene unidades; se utiliza expresado como porcentaje y puede ser un valor negativo, siempre y cuando el valor de la media sea negativo.

Una forma de interpretar el coeficiente de variación corresponde a su valor. A mayor valor, se evidencia que hay mayor heterogeneidad de los valores de la variable; a menor valor significa que hay menor heterogeneidad de los valores de la variable.

Ejemplo

Un edil de la zona quiere conocer el comportamiento del gasto en telefonía celular de los habitantes de uno de los barrios. Para tal fin, preguntó a 15 padres cabeza de familia por la cantidad de dinero, en miles de pesos, que pagó por concepto del uso de la línea celular. Los resultados son:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 21 | 18,5 | 50 | 13,5 | 23 | 19,5 | 45,5 | 37 | 25 | 12,3 | 125 | 9,5 | 16 | 14 | 32,5 |

Los valores de la media y la desviación estándar son:

MA\_11\_06\_021

,y,

Por lo cual, el coeficiente de variación es:

MA\_11\_06\_022

[SECCIÓN 2] **1.4 Análisis de tendencias**

Uno de los aspectos más importantes en el análisis estadístico está en construir inferencias relacionadas con los resultados obtenidos de la muestra. Sin embargo, es necesario complementarlos con el uso de técnicas combinadas que permitan un análisis en conjunto.

A continuación se relacionan algunos criterios que cumplen con este propósito.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Relación entre el diagrama de cajas y las medidas de tendencia central** |
| **Contenido** | * Si las medidas de tendencia central son iguales, media, mediana y moda, el diagrama de cajas es simétrico respecto a la mediana. * Si la mediana es mayor que la media, entonces la distribución de los datos está cargada hacia la derecha. En este caso se dice que la distribución está sesgada hacia la derecha. * Si la mediana es menor que la media, entonces la distribución de los datos está cargada hacia la izquierda. En este caso se dice que la distribución está sesgada hacia la izquierda. |

|  |  |
| --- | --- |
| **Practica: recurso nuevo** | |
| **Código** | MA\_11\_04\_CO\_REC80 |
| **Título** | Resuelve situaciones problema que involucran coeficientes de variación |
| **Descripción** | Actividad para solucionar situaciones problema que involucran coeficientes de variación |

|  |  |
| --- | --- |
| **Profundiza: recurso nuevo** | |
| **Código** | MA\_11\_04\_CO\_REC90 |
| **Título** | El análisis de tendencias |
| **Descripción** | Interactivo que muestra ejemplos sobre análisis de tendencias |

[SECCIÓN 2] **1.5 Consolidación**

Actividades para afianzar lo que has aprendido en esta sección.

|  |  |
| --- | --- |
| **Practica: recurso nuevo** | |
| **Código** | MA\_11\_04\_CO\_REC100 |
| **Título** | Refuerza tu aprendizaje: El análisis estadístico |
| **Descripción** | Actividades sobre el análisis estadístico |

[SECCIÓN 1] **2 La probabilidad**

La palabra probabilidad la usa la gran mayoría de las personas en su comunicación. Sin embargo, frecuentemente es empleada en términos de posibilidad y no de probabilidad. Para iniciar su estudio es necesario conocer unos conceptos previos.

[SECCIÓN 2] **2.1 Experimentos aleatorios y espacios muestrales**

Un **experimento aleatorio** es cualquier acción o proceso en el cual no se tiene certeza de su resultado final hasta que no se ejecute.

Es un experimento en el cual existe incertidumbre en el resultado, sin embargo, sí es posible determinar los posibles resultados finales.

Una de las características que se debe tener en cuenta al hablar de probabilidad es la de **aleatoriedad**, la cual garantiza que los posibles resultados del experimento no pueden ser, de ninguna manera, manipulados por alguien que desarrolle el experimento. La aleatoriedad garantiza que hay imparcialidad a la hora de realizar el experimento.

Un experimento aleatorio común es el de lanzar una moneda al aire. En este caso, no es posible determinar el resultado de la caída de la moneda antes de lanzarse, pero sí es factible determinar que el resultado puede ser cara o sello. En el caso del experimento que consiste en lanzar un dado de seis caras al aire, no es posible determinar el resultado antes de ejecutarse el lanzamiento, sin embargo sí es viable determinar que debe ser un número entre 1 y 6.

Si se considera el caso de incluir en una bolsa oscura el nombre de seis personas y se escoger aleatoriamente 3 de ellas, el experimento consiste en seleccionar 3 personas de seis. Se sabe con certeza que los seleccionados deben ser algunos de los seis nombres incluidos al inicio, pero no se puede afirmar quiénes son antes de ejecutar el experimento.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Definición de espacio muestral** |
| **Contenido** | El **espacio muestral** de un experimento aleatorio, notado como *S*, es el conjunto de todos los posibles resultados que se pueden obtener al realizar el experimento. |

Es importante hacer énfasis en que la definición de probabilidad se basa en la teoría de conjuntos clásica, por este motivo es que el lenguaje conjuntista es común en este tipo de contextos; por consiguiente, es necesario resaltar que el espacio muestral se define y se nota en términos de conjuntos.

**Ejemplo**

El gerente de una empresa quiere conformar, aleatoriamente, una comisión de dos personas para un trabajo especial. Él cuenta con un grupo de cuatro personas: Mario, Carlos, Lucía y Betty.

¿Cuál es el espacio muestral en esta situación?

El espacio muestral estará formado por todas las posibles parejas que se pueden conformar con los cuatro candidatos. En este caso el espacio muestral es:

*S* = {(Mario, Carlos), (Mario, Lucía), (Mario, Betty), (Carlos, Lucía), (Carlos , Betty), (Lucía, Betty)}

El espacio muestral se considera como el **conjunto universal** del experimento aleatorio.

Se llamará **evento** a cada subconjunto que se pueda conformar con los elementos del espacio muestral. Los eventos suelen denotarse con letras mayúsculas. En caso de que el evento esté formado por un elemento, se llamará **evento simple**.

Para el ejemplo del experimento aleatorio de seleccionar las dos personas, se muestran algunos posibles eventos:

* Sea *A* el evento que consiste en que dentro de las dos personas seleccionadas, Mario sea una de ellas; entonces:

*A* = {(Mario, Carlos), (Mario, Lucía), (Mario, Betty)}

* Si *B* es el evento que consiste en que las dos personas seleccionadas sean del mismo género, entonces:

*B* = {(Mario, Carlos), (Lucía, Betty)}

* Si se considera como el evento que las dos personas seleccionadas sean hombres, se tiene que:

*C* = {(Mario, Carlos)

Entonces, *C* es un evento simple.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Definición de probabilidad** |
| **Contenido** | La probabilidad es una medida de incertidumbre que mide la ocurrencia de un evento dentro de un experimento aleatorio.  La probabilidad de ocurrencia de un evento *A*, denotada *P*(*A*) se define como:  MA\_11\_06\_023  En donde, #(*A*) corresponde al número de elementos que contiene el evento *A*, y #(*S*) es el número de elementos que contiene el espacio muestral. |

Para el caso de la selección de las dos personas, es posible calcular algunas probabilidades.

Por ejemplo, la probabilidad de que las dos personas escogidas sean de diferente género.

Sea *A* el evento que consiste en que las dos personas seleccionadas son de diferente género, entonces:

*A* = {(Mario, Lucía), (Mario, Betty), (Carlos, Lucía), (Carlos , Betty)}

Por lo cual:

MA\_11\_06\_024

|  |  |
| --- | --- |
| **Recuerda** | |
| **Contenido** | La probabilidad de ocurrencia de un evento puede ser expresada como fracción, proporción, decimal o porcentaje.  De acuerdo con el contexto, la fracción puede no ser necesariamente expresada en su forma más simple. |

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Propiedades de la definición de probabilidad** |
| **Contenido** | De la anterior definición, es importante mencionar las siguientes características:   * La probabilidad se mide como un cociente entre dos cantidades positivas. Luego, la probabilidad de ocurrencia de un evento es un número positivo   MA\_11\_06\_025     * La probabilidad de un evento simple *C*, es:   MA\_11\_06\_026   * Ya que cualquier evento *B* es un subconjunto del espacio muestral S, entonces:   MA\_11\_06\_027  , por lo tanto   * Si *P*(*B*) = 0, *B* se llama evento imposible; y si *P*(*B*) = 1, *B* se llama evento seguro |

|  |  |
| --- | --- |
| **Profundiza: recurso nuevo** | |
| **Código** | MA\_11\_06\_CO\_REC110 |
| **Título** | Los experimentos aleatorios |
| **Descripción** | Interactivo para repasar el concepto de experimentos aleatorios |

|  |  |
| --- | --- |
| **Practica: recurso nuevo** | |
| **Código** | MA\_11\_06\_CO\_REC120 |
| **Título** | Halla el espacio muestral |
| **Descripción** | Actividad para hallar el espacio muestral de un experimento aleatorio |

[SECCIÓN 2] **2.2 La probabilidad y los conjuntos**

Al definir la probabilidad desde el punto de vista de los conjuntos, es importante establecer la relación existente entre algunas operaciones entre conjuntos y el cálculo de probabilidades. Las operaciones entre conjuntos usadas con mayor frecuencia en el cálculo de probabilidades son:

* La unión de dos eventos *A* y *B*, notada

MA\_11\_06\_028



se lee *“A o B”*, es el evento formado por todos los resultados que están en *A* o en *B* o en ambos eventos.

* La intersección de dos eventos *A* y *B*, notada

MA\_11\_06\_029



se lee *“A y B”*, es el evento formado por los elementos que están en *A* y en *B*.

* El complemento de un evento *A*, notado

MA\_11\_06\_030



se lee “complemento de *A*”, es el conjunto de todos los resultados del espacio muestral *S*, que no están en el evento *A*.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Propiedades para el cálculo de probabilidades** |
| **Contenido** | Dos eventos *A* y *B* son mutuamente excluyentes o disjuntos si su intersección es el conjunto vacío  MA\_11\_06\_031  .   * Dados dos eventos *A* y *B*, mutuamente excluyentes, se tiene que:   MA\_11\_06\_032     * En general, si los dos eventos no son disjuntos entonces   MA\_11\_06\_033     * De la propiedad anterior se puede deducir la siguiente fórmula:   MA\_11\_06\_034     * Para cualquier evento *A*, se tiene que:   MA\_11\_06\_035 |

Ejemplo

Un estudio realizado entre los nuevos clientes de una cadena de distribución de tecnología, determinó que de los 157 clientes que compraron artículos con descuento, 85 compraron celulares, 100 de los clientes eran personas menores de 25 años y 75 de las personas compraron celulares y eran menores de 25 años.

Si a la sección de tecnología con descuento llega un cliente nuevo, responder las siguientes preguntas:

a. ¿Cuál es la probabilidad de que compre un celular o sea menor de 25 años o ambos?

b. ¿Cuál es la probabilidad de que el cliente tenga 25 años o más?

c. ¿Cuál es la probabilidad de que no compre un celular y tenga 25 años o más?

El diagrama de Venn que describe la situación se muestra a continuación.

|  |  |
| --- | --- |
| **Imagen (fotografía, gráfica o ilustración)** | |
| **Código** | MA\_11\_06\_CO\_IMG05 |
| **Descripción** | Representación en diagrama de Venn del estudio realizado entre los clientes de la tienda de tecnología |
| **Código Shutterstock (o URL o la ruta en AulaPlaneta)** | 10  75  25  **A**  **B**  47  **S** |
| **Pie de imagen**  **Inferior** | Si se definen los eventos  *A*: La persona compra un celular.  *B*: la persona es menor de 25 años. |

Del diagrama anterior se puede ver que:

MA\_11\_06\_036



MA\_11\_06\_037



a. ¿Cuál es la probabilidad de que compre un celular o sea menor de 25 años o ambos?

MA\_11\_06\_038



b. ¿Cuál es la probabilidad de que el cliente tenga 25 años o más?

MA\_11\_06\_039



c. ¿Cuál es la probabilidad de que no compre un celular y tenga 25 años o más?

MA\_11\_06\_040



|  |  |
| --- | --- |
| **Practica: recurso nuevo** | |
| **Código** | MA\_11\_06\_CO\_REC130 |
| **Título** | Aplica las reglas de la unión y la intersección para hallar probabilidades |
| **Descripción** | Actividad para hallar probabilidades a partir de diagramas de Venn |

[SECCIÓN 2] **2.3 La probabilidad y el conteo**

En un experimento aleatorio es necesario identificar dos elementos fundamentales a la hora de construir el espacio muestral.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Definición de población y muestra de un experimento aleatorio** |
| **Contenido** | Al considerar un experimento aleatorio, se define la **población** como el conjunto formado por los elementos disponibles para construir un posible resultado del experimento aleatorio.  La **muestra** está conformada por el número de elementos que contiene un evento unitario  Generalmente, se denota con *N* al número de elementos que hay en la población y *n* al número de elementos que hay en la muestra. |

Por ejemplo, si se lanza una moneda, no cargada, cinco veces al aire la población estará formada por dos elementos *cara* y *sello*. Son los dos únicos elementos posibles para conformar un resultado del experimento aleatorio. La muestra deberá conformarse con cinco elementos, ya que se consideran los cinco lanzamientos de la moneda. Es decir, que un posible elemento del espacio muestral es *CSSSC,* en donde el primer lanzamiento fue cara, el segundo sello, y así sucesivamente hasta que el último lanzamiento fue cara. En este caso *N* = 2 y *n* = 5.

Existen dos características al seleccionar o construir la muestra que son determinantes al momento de hallar el espacio muestral: **la repetición y el orden**.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Definición de muestra con repetición** |
| **Contenido** | Dado un experimento aleatorio, se dice que la muestra tiene repetición cuando para conformarla se pueden usar varias veces el mismo elemento de la población. |

Ejemplo

Una empresa tiene disponibles tres cargos nuevos: gerente, director de campaña y asesor de imagen. Para dichas ofertas se postularon seis personas capacitadas para desempeñar cualquiera de los tres cargos. Si los candidatos son: Tom, Fred, Sid, Emma, Sara y Leo y la selección de los tres candidatos se hace de forma aleatoria, se tiene que:

* La población la conforman las seis personas, *N* = 6.
* La muestra la componen las tres personas seleccionadas, *n* = 3.
* En este caso vale la pena mencionar que, al seleccionar una persona para un cargo, esta no puede volver a considerarse para la elección del cargo siguiente, por lo cual no existe repetición en la muestra.

El caso de una muestra con repetición se da en el ejemplo del lanzamiento de la moneda cinco veces. En esta ocasión, la población está conformada por dos elementos: cara y sello, y la muestra por cinco elementos, es decir que es necesario repetir elementos de la población para conformar la muestra. Esta es una muestra con repetición.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Definición de muestra ordenada** |
| **Contenido** | Dado un experimento aleatorio, se dice que la muestra es **ordenada** si al conformarla es importante el orden en que se ubiquen los elementos de la población. |

Ejemplo

Para la semifinal del torneo femenino intercolegiado de atletismo se clasificaron Marcos, Juan, Elías, Lucas y José. Se realiza la prueba y los tres primeros en llegar a la meta se clasifican.

Se tiene que *N* = 5 y *n* = 3; además, en el momento de la llegada de los competidores lo único que importa es que arriben dentro de los tres primeros lugares.

En esta oportunidad, no es importante si el atleta llega de primero, segundo o tercero, en los tres casos ha clasificado, es decir la muestra no es ordenada.

Si para el mismo ejemplo, se decide que el primer atleta en pasar la meta recibirá un premio económico, el segundo un reloj y el tercero solamente la clasificación, el orden de llegada sí importa, ya que no significa lo mismo llegar de primero que de segundo. En este caso, la muestra es ordenada.

Una vez establecidos los criterios de orden y repetición, así como los conceptos de población y muestra, es posible definir algunas técnicas de conteo que permiten calcular el número de elementos del espacio muestral o del evento de forma más rápida y efectiva.

[SECCIÓN 3] **2.3.1 El principio de multiplicación**

Esta técnica se aplica en aquellos experimentos aleatorios en los cuales existe orden y repetición en la muestra

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Definición del principio de multiplicación** |
| **Contenido** | Si en un experimento aleatorio se tiene una población de tamaño *N*, se debe tomar una muestra de tamaño *n* y además, la muestra tiene orden y repetición, entonces:  #(S) = *Nn*  donde #(S) corresponde al número de elementos del espacio muestral. |

Ejemplo

Un programador de computadores está generando un nuevo programa que le permita construir aleatoriamente un número para la lotería, el cual consta de cuatro cifras y una serie de dos. Para conformar el número se considera que sí existe repetición, ya que el número puede estar formado por dígitos iguales, por ejemplo 2233. En forma similar, existe el orden en el número ya que 4333 es diferente de 3343. Para conformar el primer número se tienen 10 posibilidades correspondientes a los dígitos, por lo tanto, la población es 10; además, el número está formado por seis dígitos, luego la muestra es 6.

Entonces:

#(*S*) = 106 = 1 000 000

Existe 1 000 000 de posibles billetes de lotería distintos.

Ejemplo

A la final de un torneo de béisbol llegan dos equipos: Leones y Tigres. Se van a enfrentar en una serie de cuatro juegos. Si los cuatro partidos se deben jugar, ¿cuál es la probabilidad de que Tigres no gane ninguno de los juegos?

En este experimento aleatorio la población está conformada por dos elementos, *N* = 2, ya que en cada partido se tienen dos opciones: que gane el equipo *A* o que gane el equipo *B* (en un partido de béisbol no existen los empates). Además, *n* = 4, ya que la serie consta de 4 partidos.

En este caso la muestra es ordenada, ya que no es lo mismo que Leones gane el primer partido y pierda el segundo, a que pierda el primero y gane el segundo. Igualmente, existe repetición ya que si uno de los equipos gana el primer partido, es posible que lo vuelva a hacer en el segundo. Por lo tanto:

#(*S*) = *Nn* = 24 = 16

Si se considera el evento *A*: Tigres no gana ningún partido, entonces el equipo Leones ganaría todos los juegos es decir que *N* = 1 y *n* = 4. Por lo tanto:

#(*A*) = *Nn* = 14 = 1

Esto es que:

MA\_11\_06\_041

Luego, la probabilidad de que Tigres no gane un partido es de 1 entre 16, o de 0,0625 o del 6,25%.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Definición de principio de multiplicación para poblaciones diferentes** |
| **Contenido** | Dado un experimento aleatorio, si existen *k* poblaciones distintas para conformar una muestra de tamaño *n*, entonces,  MA\_11\_06\_042 |

Ejemplo

Una agencia de viajes ofrece a los turistas de una determinada ciudad un programa turístico para 3 días. Para el primer día se ofrece paseo por la ciudad o caminata por la sabana; para el segundo día se propone visita a museos, *tour* por el centro de la ciudad o cabalgata por los alrededores del centro de la ciudad; para el tercer día se presenta un *tour* nocturno por los bares del centro o visita a la casa de poesía de la ciudad.

El tiempo que se requiere en cada actividad hace que el viajero pueda escoger solamente una actividad por día. El principio de multiplicación se aplica como se muestra en la imagen.

|  |  |
| --- | --- |
| **Imagen (fotografía, gráfica o ilustración)** | |
| **Código** | MA\_11\_06\_CO\_IMG06 |
| **Descripción** | Diagrama para formas de escoger un tour |
| **Código Shutterstock (o URL o la ruta en AulaPlaneta)** | 2 rutas  Día 2  Día 1  3 rutas  Día 3  2 rutas  2x3x2=12 |
| **Pie de imagen**  **Inferior** | Existen tres poblaciones: una para el primer día, otra para el segundo y finalmente una para el tercero. |

En ocasiones es útil encontrar el número de elementos del espacio muestral a partir de una representación gráfica en la cual se muestran las diferentes posibilidades de terminar un experimento aleatorio.

**El diagrama de árbol** es la representación en la cual se describe gráficamente el espacio muestral, y en la que cada ramificación constituye un elemento de la muestra y cada camino corresponde a un resultado posible del experimento.

Ejemplo

Un estudiante de primer semestre de la universidad puede inscribir dos materias del ciclo básico: Matemáticas básicas e Introducción a la Ingeniería. Al consultar los horarios disponibles se da cuenta de que cada una de las materias las dictan tres profesores diferentes: Dra. Perdomo, Dr. Romero y Dr. Peña. Además, cada uno de los profesores tiene disponibles horarios en la mañana y en la tarde.

Adicionalmente, el estudiante puede hacer una selección sin restricciones de materia, profesor y horario; el número de distintas opciones se presenta en el siguiente diagrama de árbol.

|  |  |
| --- | --- |
| **magen (fotografía, gráfica o ilustración)** | |
| **Código** | MA\_11\_06\_CO\_IMG07 |
| **Descripción** | Diagrama de árbol para las opciones de escoger materia, profesor y horario |
| **Código Shutterstock (o URL o la ruta en AulaPlaneta)** | Matemáticas  Romero  Peña  Perdomo  Mañana  Tarde  Mañana  Tarde  Mañana  Tarde  Introducción  Romero  Peña  Perdomo  Mañana  Tarde  Mañana  Tarde  Mañana  Tarde |
| **Pie de imagen**  **Inferior** | Cada una de las ramas corresponde a un posible resultado del experimento aleatorio, por tanto, existen 12 posibilidades para inscribir. |

[SECCIÓN 3] **2.3.2 La permutación**

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Definición de factorial** |
| **Contenido** | El operador factorial se define sobre los números naturales incluyendo el cero. Se representa mediante el número natural y el símbolo !.  El **factorial de un número** se define como el producto del número con todos sus naturales anteriores hasta 1.  MA\_11\_06\_043    Para que la operación esté bien definida en algunas aplicaciones futuras es necesario definir que 0! = 1. |

La permutación es la técnica que se aplica en aquellos experimentos aleatorios en los cuales existe orden y no hay repetición en la muestra

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Definición de permutación** |
| **Contenido** | Dado un experimento aleatorio con una población *N* y una muestra *n*, si en la muestra existe orden pero no repetición, el número de elementos del espacio muestral corresponde a la permutación de *n* en *N*, la cual se simboliza *nPN* y se define:  MA\_11\_06\_044 |

El resultado de una permutación es un número natural.

Para el caso en el que el tamaño de la población *N* sea igual al tamaño de la muestra *n* se tiene que:

MA\_11\_06\_045

**

Ejemplo

Para la elección de la junta directiva del consejo escolar se han postulado siete candidatos. En los estatutos del colegio se ha estipulado que una vez realizada la elección, el candidato con mayor votación será el presidente, el segundo en número de votos será el tesorero y el tercero en votos será el secretario.

El experimento aleatorio consiste en conformar el consejo estudiantil de tres miembros con siete candidatos disponibles. Por tanto *N* = 7 y *n* = 3.

En este caso existe el orden ya que no es lo mismo ser el presidente que el secretario. Además, no existe repetición, puesto que una persona no puede ocupar dos cargos.

Por tanto, el número de elementos del espacio muestral es:

MA\_11\_06\_046



Hay 210 formas distintas de conformar el consejo con los siete candidatos disponibles.

Ejemplo

En un torneo de póquer hay seis finalistas: Martín, MacMilan, Wills, Llano, García y Rozo; ellos deben jugar hasta que uno se quede con todo el dinero que hay en juego. La organización del torneo premiará a los dos finalistas: al primero un permio en dinero y un reloj de lujo y al segundo un reconocimiento en dinero. Si los seis finalistas tienen la misma capacidad de ser ganadores, ¿cuál es la probabilidad de que Wills sea el campeón?

En este experimento aleatorio se tiene una población de seis finalistas *N* = 6 y una muestra de dos finalistas *n* = 2. Existe orden y no hay repetición, por tanto:

MA\_11\_06\_047

****

Sea *C* el evento que consiste en que Wills es el campeón. Se tiene que:

MA\_11\_06\_048

.

Por tanto:

MA\_11\_06\_049



Es decir, la probabilidad de que Wills quede campeón es de 5 en 30.

[SECCIÓN 3] **2.3.3 La combinatoria**

La combinatoria es la técnica que se aplica en aquellos experimentos aleatorios en los cuales existe orden y no hay repetición en la muestra.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Definición de combinatoria** |
| **Contenido** | Dado un experimento aleatorio con una población *N* y una muestra *n*, si en la muestra no existe orden ni repetición, el número de elementos del espacio muestral corresponde a la combinatoria de *n* en *N*, la cual se simboliza  MA\_11\_06\_050    y se define como:  MA\_11\_06\_051 |

Ejemplo

La persona encargada del departamento de desarrollo humano de una empresa debe programar cinco entrevistas para el nuevo cargo de operario, al cual se han presentado 8 candidatos.

El encargado decide seleccionar de forma aleatoria las cinco personas que va a entrevistar.

El experimento aleatorio consiste en seleccionar 5 personas de un grupo de 8 disponibles. *N* = 8 y *n* = 5.

En este caso no existe orden en la muestra ya que no importa si se selecciona primero o último, la persona va a ser entrevistada; igualmente, no existe repetición ya que una persona no puede ser entrevistada dos veces. Por tanto:

MA\_11\_06\_052



Existen 56 formas distintas de escoger 5 personas de un grupo de 8 disponibles.

Ejemplo

Se tienen seis personas disponibles para realizar el turno de la tarde, en una de las tiendas del centro; cuatro de ellas son mujeres y dos son hombres. Si se quiere escoger aleatoriamente tres de ellas para trabajar en la tarde del día siguiente, ¿cuál es la probabilidad de que las tres personas seleccionadas sean mujeres?

En este caso se tiene una población, *N* = 6, de seis trabajadores disponibles para elegir una muestra *n* = 3. No existe repetición ya que ninguna persona puede ser seleccionada dos veces y no hay orden en la elección ya que si son seleccionadas van a realizar el mismo trabajo.

Por lo tanto:

MA\_11\_06\_053

Además, sea *B* el evento que consiste en que tres personas seleccionadas son mujeres, entonces:

MA\_11\_06\_054

ya que de cuatro mujeres disponibles se deben seleccionar tres y de dos hombres se debe seleccionar ninguno. En conclusión:

MA\_11\_06\_055

Es decir, la probabilidad de que las tres personas seleccionadas sean mujeres es de 1 a 5 o del 20%

|  |  |
| --- | --- |
| **Practica: recurso nuevo** | |
| **Código** | MA\_11\_06\_CO\_REC140 |
| **Título** | Halla la cantidad de elementos del espacio muestral |
| **Descripción** | Actividad para hallar el número de elementos del espacio muestral utilizando técnicas de conteo |

|  |  |
| --- | --- |
| **Practica: recurso nuevo** | |
| **Código** | MA\_11\_06\_CO\_REC150 |
| **Título** | Resuelve situaciones problema que involucran probabilidades |
| **Descripción** | Actividad para resolver situaciones problema que involucran probabilidades |

[SECCIÓN 2] **2.4 La probabilidad condicional**

En ocasiones es necesario calcular la probabilidad de un evento sujeto a alguna condición o a algún otro evento que ya ocurrió.

La **probabilidad condicional** se define para dos eventos *A* y *B*, donde uno de ellos hace el papel de condición sobre el otro. Es decir, se debe considerar que uno de los dos eventos ya ocurrió y se quiere saber qué pasa con el otro evento.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Definición de probabilidad condicional** |
| **Contenido** | Dados dos eventos *A* y *B*, se llama probabilidad condicional de *A* sobre *B*, la cual se simboliza  MA\_11\_06\_056    y se lee “la probabilidad de *A* dado *B*”; se calcula así:  MA\_11\_06\_057    en donde el evento *B* ya ocurrió y se quiere calcular la probabilidad de *A*, o *A* ya ocurrió y se quiere calcular la probabilidad de *B*. |

Ejemplo

Un automóvil debe seguir una vía en la que hay dos semáforos. Para este caso, el experimento aleatorio consiste en pasar los dos semáforos y el espacio muestral es:

*S* = {(*D*, *D*), (*D*, *N*), (*N*, *D*), (*N*, *N*)}

donde *D* es detenerse y *N* es no detenerse.

Se consideran los eventos

*A*: el auto se detuvo en el primer semáforo y

*B*: el auto se detuvo en el primer semáforo

Es posible calcular la probabilidad de que el auto se detenga en el segundo semáforo, si se sabe que se detuvo en el primero. En este caso al evento *A* se le llama la condición para la ocurrencia de *B*.

Se tiene que

MA\_11\_06\_058

entonces:

MA\_11\_06\_059

Es decir, la probabilidad de que el auto se detenga en el segundo semáforo si se sabe que paró en el primero es 1 de 2.

[SECCIÓN 3] **2.4.1 La probabilidad condicional en diagramas de Venn**

Es posible utilizar los diagramas de Ven para representar información en la cual se involucra una condición. El siguiente ejemplo muestra esta situación.

Ejemplo

En una ruta de una vía central de la ciudad que tiene tres semáforos, se calculan las probabilidades de que un automóvil se detenga en alguno de ellos.

Sean *A* el evento que consiste en que el automóvil se detiene el en primer semáforo, *B* el evento que consiste en que se detenga en el segundo y *C* el evento que consiste en que se detenga en el tercero.

Las probabilidades asociadas a los eventos se relacionan en el siguiente diagrama.

|  |  |
| --- | --- |
| **magen (fotografía, gráfica o ilustración)** | |
| **Código** | MA\_11\_06\_CO\_IMG08 | |
| **Descripción** | Diagrama de Venn para el caso de los tres semáforos |
| **Código Shutterstock (o URL o la ruta en AulaPlaneta)** | A  0.5  0.01  0.05  0.1  0.05  0.05  0.1  0.05  C  B |
| **Pie de imagen**  **Inferior** | Del diagrama se pueden calcular las probabilidades de cada uno de los eventos y de las intersecciones. |

Se puede calcular la probabilidad de que el auto se detenga en el tercer semáforo dado que se detuvo en los dos primeros.

MA\_11\_06\_060

[SECCIÓN 3] **2.4.2 El diagrama de árbol para la probabilidad condicional**

En algunos contextos es necesario organizar los eventos relacionados en diagramas que permitan construir todas las probabilidades asociadas a estos eventos. Esta situación se ilustra en el siguiente ejemplo.

Ejemplo

El profesor Cortés sabe, por experiencia, que la probabilidad de que un estudiante apruebe la asignatura que imparte es de 0,35. Además, sabe que la probabilidad de que un estudiante que aprobó la asignatura haya asistido a las sesiones de monitoria es de 0,75 y que la probabilidad de que un estudiante que no aprobó la asignatura no haya asistido a la monitoria es de 0,82.

En este caso se tienen dos eventos:

*A*: el estudiante aprueba la asignatura

*B*: eL estudiante asiste a las sesiones de monitoria

Se sabe que:

MA\_11\_06\_061

.

Las probabilidades asociadas a estos dos eventos se pueden representar mediante el siguiente diagrama:

|  |  |
| --- | --- |
| **Imagen (fotografía, gráfica o ilustración)** | |
| **Código** | MA\_11\_06\_CO\_IMG09 |
| **Descripción** | Diagrama árbol de probabilidades condicionales |
| **Código Shutterstock (o URL o la ruta en AulaPlaneta)** | P(A)=0.35 |
| **Pie de imagen**  **Inferior** | Las probabilidades asociadas se calculan de acuerdo con las propiedades para calcular la probabilidad del complemento de un conjunto. |

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **La probabilidad de la intersección de conjuntos** |
| **Contenido** | De acuerdo con la definición de probabilidad condicional, se tiene:  MA\_11\_06\_062  Entonces  MA\_11\_06\_063 |

Por lo cual, el diagrama se puede complementar como sigue:

|  |  |
| --- | --- |
| **Imagen (fotografía, gráfica o ilustración)** | |
| **Código** | MA\_11\_06\_CO\_IMG010 |
| **Descripción** | Diagrama de árbol de probabilidades condicionales |
| **Código Shutterstock (o URL o la ruta en AulaPlaneta)** | P(A)=0.35 |
| **Pie de imagen**  **Inferior** | El diagrama se completa con el cálculo de las intersecciones para cada una de las ramas del árbol. |

Así, cualquier pregunta relacionada con los dos eventos o sus complementos se puede resolver muy fácilmente mirando las ramas del diagrama.

|  |  |
| --- | --- |
| **Practica: recurso nuevo** | |
| **Código** | MA\_11\_06\_CO\_REC160 |
| **Título** | Halla la probabilidad de los eventos |
| **Descripción** | Actividad para hallar la probabilidad de eventos compuestos utilizando el diagrama de árbol |

[SECCIÓN 3] **2.4.3 Probabilidad total**

Las siguientes condiciones permiten determinar la probabilidad total de un evento cuando depende de la ocurrencia de otros eventos.

|  |  |
| --- | --- |
| **Imagen (fotografía, gráfica o ilustración)** | |
| **Código** | MA\_11\_06\_CO\_IMG11 |
| **Descripción** | Diagrama de probabilidad total |
| **Código Shutterstock (o URL o la ruta en AulaPlaneta)** | A  B |
| **Pie de imagen**  **Inferior** | Si *A* es un subconjunto de *B* y además *B* se puede expresar como la unión de varios conjuntos. |

Dado que:

MA\_11\_06\_064

Además:

MA\_11\_06\_065

Como todas las intersecciones planteadas son disjuntas se tiene que:

**MA\_11\_06\_066**

Por tanto:

MA\_11\_06\_067

Es decir que, la probabilidad de ocurrencia de un evento se puede calcular a partir de las probabilidades condicionales asociadas a la ocurrencia de un evento condición.

En el ejemplo del profesor Cortés, la probabilidad total se puede aplicar para calcular la probabilidad de que un estudiante asista a la monitoria.

Es decir:

MA\_11\_06\_068

MA\_11\_06\_069

Es decir, la probabilidad de que un estudiante del profesor Cortés asista a la monitoria es de 37,95%

[SECCIÓN 2] **2.5. Independencia**

Para el caso de dos eventos relacionados mediante la probabilidad condicional, es posible determinar si la condición no influye sobre los resultados de ocurrencia del evento. En tal caso un evento no condiciona al otro.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **La definición de independencia** |
| **Contenido** | Dados dos eventos *A* y *B*, se dice que *A* y *B* son **independientes** si:  MA\_11\_06\_070  En otras palabras:  MA\_11\_06\_071  MA\_11\_06\_072  MA\_11\_06\_073  Si la probabilidad de la intersección de dos eventos es igual a la multiplicación de las probabilidades de cada uno, entonces *A* y *B* son independientes |

Ejemplo

La probabilidad de que una persona viva más de 90 años es de 0,35 y la probabilidad de que una persona consuma un suplemento vitamínico es de 0,2. Además, la probabilidad de que una persona viva más de 90 años y consuma el suplemento vitamínico es de 0,07.

La situación se representa en el siguiente diagrama de Venn.

|  |  |
| --- | --- |
| **Imagen (fotografía, gráfica o ilustración)** | |
| **Código** | MA\_11\_06\_CO\_IMG12 |
| **Descripción** | Diagrama de probabilidad total |
| **Código Shutterstock (o URL o la ruta en AulaPlaneta)** | 0.28  0.07  0.13  **A**  **B**  0.52  **S** |
| **Pie de imagen**  **Inferior** | Donde *A*: la persona vive más de 90 años y *B*: la persona consume el suplemento vitamínico. |

Se tiene que *P*(*A*) = 0,35 y *P*(*B*) = 0,2.

Como *P*(*A*) *P*(*B*) = 0,35 × 0,2 = 0,07, entonces *A* y *B* son independientes.

En otras palabras, el hecho de que una persona consuma el suplemento vitamínico no influye en que viva más de 90 años o viceversa.

|  |  |
| --- | --- |
| **Practica: recurso nuevo** | |
| **Código** | MA\_11\_06\_CO\_REC180 |
| **Título** | Indica si el evento es dependiente o independiente |
| **Descripción** | Actividad para reconocer dependencia o independencia de eventos |

|  |  |
| --- | --- |
| **Profundiza: recurso nuevo** | |
| **Código** | MA\_11\_06\_CO\_REC170 |
| **Título** | El teorema de la probabilidad total y el Teorema de Bayes |
| **Descripción** | Interactivo que muestra el Teorema de la probabilidad total y el teorema de Bayes |

|  |  |
| --- | --- |
| **Practica: recurso nuevo** | |
| **Código** | MA\_11\_06\_CO\_REC190 |
| **Título** | Halla la probabilidad teniendo en cuenta la tabla de contingencia |
| **Descripción** | Actividad para hallar la probabilidad de ocurrencia de eventos condicionales a partir de la información dada en tablas de contingencia |

[SECCIÓN 2] **2.6 Consolidación**

Actividades para afianzar lo que has aprendido en esta sección.

|  |  |
| --- | --- |
| **Practica: recurso nuevo** | |
| **Código** | MA\_11\_06\_CO\_REC200 |
| **Título** | Refuerza tu aprendizaje: La probabilidad |
| **Descripción** | Actividades sobre la probabilidad |

[SECCIÓN 1] **3 Las variables aleatorias**

La modelización de los experimentos aleatorios es una parte fundamental para poder generalizar el análisis del cálculo de probabilidades. Es por ello que es necesario establecer criterios que permitan construir este tipo de modelos.

|  |  |
| --- | --- |
| **Profundiza: recurso aprovechado** | |
| **Código** | MA\_11\_06\_CO\_REC210 |
| **Título** | Las variables aleatorias |
| **Descripción** | Interactivo que muestra las variables aleatorias discretas y continuas |

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Definición de variable aleatoria** |
| **Contenido** | Una **variable aleatoria** establece un conteo específico sobre los elementos del espacio muestral.  Generalmente, las variables aleatorias se denotan con letras mayúsculas equivalentes al manejo algebraico |

Ejemplo

Al considerar el experimento aleatorio que consiste en seleccionar dos personas dentro de un grupo de cinco candidatos (Hugo, Martín, Camila, Sara y Olga), el espacio muestral es:

*S* = {(*H*, *M*), (*H*, *C*), (*H*, *S*), (*H*, *O*), (*M*, *C*), (*M*, *S*), (*M*, *O*), (*C*, *S*), (*C*, *O*), (*S*, *O*)}

Si se define *X* como el número de mujeres que hay en la muestra seleccionada, es posible asignar a cada elemento del espacio muestral un valor numérico.

Se puede asignar a cada elemento del espacio muestral una imagen correspondiente:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Evento | Imagen |  | Evento | Imagen |
| (*H, M*) | *X*(*H, M*) *= 0* |  | (*M, S*) | 1 |
| (*H, C*) | 1 |  | (*M, O*) | 1 |
| (*H, S*) | 1 |  | (*C, S*) | 2 |
| (*H, O*) | 1 |  | (*C, O*) | 2 |
| (*M, C*) | 1 |  | (*S, O*) | 2 |

Se puede ver que la variable aleatoria *X* toma valores 0, 1 y 2.

Las variables aleatorias pueden ser de dos tipos: **discretas** y **continuas**.

* Para el caso del experimento aleatorio que consiste en lanzar una moneda hasta que salga cara, el espacio muestral es:

*S* = {*C*, *SC*, *SSC*, *SSSC*, *SSSSC*, *SSSSSC*,…}

Si se define la variable aleatoria *Z* como el número de lanzamientos necesario para obtener cara, para este caso, la variable *Z* toma valores en los números naturales. Es decir *Z* = 0, 1, 2, 3,…

En este caso se dice que la variable aleatoria es **discreta**.

* En otro caso, se selecciona aleatoriamente una persona de una muestra determinada y se mide el tiempo de reacción ante un estímulo. La variable aleatoria *X* mide el tiempo que tarda en reaccionar la persona seleccionada una vez se ha aplicado el estímulo.

En este caso la variable *X* se mide en un intervalo cerrado [0, *t*] donde *t* es el máximo tiempo esperado para reaccionar al estímulo.

En este caso se dice que la variable aleatoria es **continua**.

|  |  |
| --- | --- |
| **Profundiza: recurso aprovechado** | |
| **Código** | MA\_11\_06\_CO\_REC220 |
| **Título** | Caracteriza variables aleatorias discretas |
| **Descripción** | Actividad para caracterizar variables aleatorias discretas |

[SECCIÓN 2] **3.1 La función de distribución de probabilidades**

Una vez se ha definido una variable aleatoria discreta sobre el espacio muestral, es posible definir una función que permita calcular las probabilidades de ocurrencia en el experimento aleatorio.

Ejemplo

La probabilidad de que una batería continúe funcionando después de un año de uso es de 0,3. Se inspecciona un lote de cuatro baterías que han tenido un año de uso y se prueban una a una para determinar si continúa funcionando o no.

El espacio muestral, si se considera *B*: buena y *D*: defectuosa, es:

*S* = {*BBBB*, *BBBD*, *BBDB*, *BDBB*, *DBBB*, *BBDD*, *BDDB*, *BDBD*, *DDBB*, *DBDB*, *DBBD*, *BDDD*, *DBDD*, *DDBD*, *DDDB*, *DDDD*}

Sea *X* la variable aleatoria que mide el número de baterías defectuosas en la muestra. Por lo cual *X* = 0, 1, 2, 3, 4.

Para el caso *BBBB*, la variable aleatoria toma el valor 0.

Se puede calcular *P*(*X* = 0), es decir, la probabilidad de que no hayan baterías defectuosas en el lote.

Ya que *P*(*B*) = 0,3, entonces *P*(*D*) = 0,7.

Además, si se asume el principio de independencia, entonces *P*(BB) = (0,3)(0,3).

En consecuencia

*P*(*X* = 0) = *P*(*BBBB*) = (0,3)(0,3)(0,3)(0,3) = (0,3)4

ya que solo hay un resultado posible para *X* = 0 en el espacio muestral.

Si se considera el caso *X* = 1, se tienen 4 posibles resultados en el espacio muestral:

*BBBD*, *BBDB*, *BDBB*, *DBBB*

Por tanto:

P(X = 1) = P(BBBD) + P(BBDB) + P(BDBB) + P(DBBB)

= (0.3)(0.3)(0.3)(0.7) + (0.3)(0.3)(0.7)(0.3) + (0.3)(0.7)(0.3)(0.3) + (0.7)(0.3)(0.3)(0.3)

=4(0.3)3(0.7)

De igual forma, es posible calcular el de la probabilidad para cada uno de los valores de la variable.

|  |  |
| --- | --- |
| Si *X =* 0 | P(*X* = 0) = P(*BBBB*) = (0,3)(0,3)(0,3)(0,3) = (0,3)4 |
| Si *X =* 1 | P(*X* = 1) = 4(0,3)3(0,7) |
| Si *X* = 2 | P(*X* = 2) = 6(0,3)2(0,7)2 |
| Si *X* = 3 | P(*X* = 3) = 4(0,3)1(0,7)3 |
| Si *X* = 4 | P(*X* = 1) = (0,7)4 |

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Definición de función de distribución** |
| **Contenido** | Sea *X* una variable aleatoria discreta, se define la **función de distribución de probabilidades, *f(X)*** como:  *f*(*x*) = *P*(*X* = *xi*)  Para todos los valores *i* que toma la variable aleatoria. |

Para el caso del ejemplo anterior,

MA\_11\_06\_074

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Introducción a la notación** |
| **Contenido** | En el campo de las variables aleatorias es común encontrar la notación en términos funcionales, es por ello que, la notación:  MA\_11\_06\_075  que significa  MA\_11\_06\_076  se pueda relacionar con el lenguaje de las desigualdades, por ejemplo: al menos, a lo más, a lo sumo, máximo, no mayor que, entre. |

En el ejercicio anterior, es posible hallar la probabilidad de que en la muestra de cuatro baterías, al menos tres sean defectuosas. Para

tal fin se debe calcular:

MA\_11\_06\_077

En este caso, también es posible hacer el cálculo como:

MA\_11\_06\_078

si se usa la propiedad del cálculo de la probabilidad del complemento de un conjunto.

[SECCIÓN 2] **3.2 La función de distribución acumulada de probabilidades**

Es usual que para el cálculo de probabilidades de una variable aleatoria sea más eficiente el uso de los complementos, ya que facilita el proceso. Por esta razón es que se construye la función acumulada de probabilidades.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Definición de función de distribución acumulada de probabilidades** |
| **Contenido** | Dada una variable aleatoria discreta con *X* = 0, 1, 2,…,*n*, con función de distribución de distribución *f*(*x*), se define la función de distribución acumulada de probabilidades de *X*, *F*(*x*) como:  MA\_11\_06\_079 |

Para el ejemplo de las baterías, la función acumulada de probabilidades se construye de la siguiente forma:

Si x ≤ 0 *F*(0) = *f*(0) = 0,0081

Si x ≤ 1 F(1) = *f*(0) + *f*(1) = 0,0081 + 0,0756 = 0,0837

Si x ≤ 2 F(2) = *f*(0) + *f*(1) + *f*(2) = 0,0081 + 0,0756 + 0,2646 = 0,3483

Si x ≤ 3 F(3) = *f*(0) + *f*(1) + *f*(2) + *f*(3) = 0,0081 + 0,0756 + 0,2646 + 0,4116 = 0,7599

Si x ≤ 4 F(4)= *f*(0) + *f*(1) + *f*(2) + *f*(3) + *f*(4) = 0,0081 + 0,0756 + 0,2646 + 0,4116 + 0,2401 = 1

Una de las principales características de la función acumulada de probabilidades es que en el último valor de la variable el resultado debe ser 1.

En términos funcionales, la función *F*(*x*) se expresa:

MA\_11\_06\_080

En consecuencia, la función *F*(*x*) se define para todos los valores reales de *x*. Esta función permite generalizar el modelo del cálculo de probabilidades para cualquier valor real. Además, se trata de una función definida a trozos.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Cómo usar *F*(*x*) para calcular probabilidades** |
| **Contenido** | El uso de la función acumulada de probabilidades para calcular valores de la variable debe fundamentarse en la siguiente relación:  MA\_11\_06\_081  Para cualquier par de valores *a* y *b* que estén en el rango de la variable aleatoria. |

Para el ejemplo, la probabilidad de que entre una y tres baterías estén defectuosas es:

MA\_11\_06\_082

[SECCIÓN 2] **3.3 El valor esperado**

Otro de los componentes de una variable aleatoria es el **valor esperado**, el cual pretende calcular el promedio ponderado de la variable de acuerdo con las probabilidades asociadas a sus valores.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Definición de valor esperado** |
| **Contenido** | Dada una variable aleatoria *X* = 0, 1, 2,…, *n*, con función de distribución *f*(*x*)*,* se define el valor esperado de *X*, como:  MA\_11\_06\_083 |

Para el caso del ejemplo de las baterías, la función de distribución:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *x* | 0 | 1 | 2 | 3 | 4 |
| *f*(*x*) | 0,0081 | 0,0756 | 0,2646 | 0,4116 | 0,2407 |

El valor esperado es:

MA\_11\_06\_084

Es necesario aclarar que el valor esperado tiene unidades y corresponde a las mismas en las cuales se mide la variable.

[SECCIÓN 2] **3.4 La varianza**

Ya que se definió la **media poblacional** o **valor esperado**, es posible definir una medida de dispersión que permite determinar la validez de futuras conclusiones.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Definición de varianza** |
| **Contenido** | Dada una variable aleatoria *X* = 0, 1, 2,…,n, con función de distribución *f*(*x*)y valor esperado *E*(*x*), se define la varianza de *x*  MA\_11\_06\_085  como:  MA\_11\_06\_086  Es posible mostrar que  MA\_11\_06\_087  por lo cual *V*(*x*) siempre es un valor positivo. |

Para el caso de las baterías, si:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *x* | 0 | 1 | 2 | 3 | 4 |
| *x*2 | 0 | 1 | 4 | 9 | 16 |
| *f*(*x*) | 0,0081 | 0,0756 | 0,2646 | 0,4116 | 0,2407 |

Como:

*E*(*x*2) = (0 × 0,0081 + 1 × 0,0756 + 4 × 0,2646 + 9 × 0,4116 + 16 × 0,2407)

Entonces:

MA\_11\_06\_088

La varianza poblacional no tiene interpretación y sus unidades son cuadradas.

[SECCIÓN 2] **3.5 La distribución binomial**

Es posible crear modelos de variables aleatorias que se ajusten a diferentes contextos y que faciliten el cálculo de probabilidades. Una de las variables aleatorias más comunes es la distribución binomial.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Características de una variable aleatoria binomial** |
| **Contenido** | Una variable aleatoria se llama binomial si cumple las siguientes características:   * El experimento aleatorio se llama dicotómico o de Bernoulli, es decir que la variable solo puede tomar dos valores. Para este contexto los denominaremos éxito o fracaso.   Vale la pena mencionar que el fracaso no tiene una interpretación ética que implica algo negativo.   * Se conoce la probabilidad de éxito denotada *p*. * Si se define *q* como la probabilidad de fracaso, se tiene: * El experimento aleatorio se repite un número finito de veces. Generalmente se denota el número de repeticiones como *n.* Se tiene que *x* = 0, 1, 2,…, *n* * Las repeticiones del experimento son independientes entre sí. * La variable *X* mide el número de éxitos que hay en *n* repeticiones del experimento aleatorio |

Si en un determinado contexto es posible establecer que las cinco características se cumplen, es posible afirmar que el experimento aleatorio es binomial.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **La distribución binomial** |
| **Contenido** | Sea *X* una variable aleatoria binomial, entonces la función de distribución de probabilidades es:  MA\_11\_06\_089  donde:  *n* es el número de repeticiones  *p* es la probabilidad de éxito  *q* es la probabilidad de fracaso |

Ejemplo

La probabilidad de que una persona adquiera cierto virus cuando se expone mucho tiempo al Sol es de 0,25. Si se seleccionan 8 personas que se han expuesto al Sol, ¿cuál es la probabilidad de que tres de ellas contraigan el virus?

En este caso la variable aleatoria mide si la persona contrae o el virus o no razón por lo cual es una variable dicotómica.

Como se conoce la probabilidad de que contraiga el virus entonces *p* = 0,25.

En consecuencia, es posible hallar *q*:

*p* + *q* = 1, entonces, *q* = 1 – *p* = 1 – 0,25 = 0,75

El experimento se repite 8 veces.

Se supone la independencia, es decir, el hecho de que una persona contraiga el virus o no lo contraiga, no influye sobre el hecho que la otra persona lo haga o no.

En consecuencia, *X* es una variable aleatoria binomial, donde:

*p* = 0,25 *q* = 0,75 *n* = 8

Entonces, la probabilidad de que tres personas de las ocho contraigan el virus es:

MA\_11\_06\_090

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Valor esperado y varianza de la distribución binomial** |
| **Contenido** | Sea *X* una variable aleatoria binomial con función de distribución  MA\_11\_06\_091  Entonces:  *E*(*x*) = *np* *V*(*x*) = *npq* |

En el ejemplo de las personas que se exponen al sol se tiene que:

*E*(*x*) = 8 × 0,25 = 2 personas, *V*(*x*) = 8 × 0,25 × 0,75 = 1,5

Es decir que, se espera que entre las 8 personas dos de ellas contraigan el virus.

|  |  |
| --- | --- |
| **Profundiza: recurso aprovechado** | |
| **Código** | MA\_11\_06\_CO\_REC230 |
| **Título** | Aplica la distribución binomial |
| **Descripción** | Actividad para resolver ejercicios y problemas con la distribución binomial |

[SECCIÓN 2] **3.6 Las variables aleatorias continuas**

Las variables aleatorias continuas están directamente relacionadas con las variables cuantitativas en un estudio estadístico. Su aplicación es de gran utilidad ya que su estudio se fundamenta en las herramientas construidas en el cálculo infinitesimal.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **Definición de variable aleatoria continua** |
| **Contenido** | Sea *X* definida en un intervalo cerrado [*a*, *b*], si existe una función *f*(*x*) que cumple con las siguientes condiciones:   * *f*(*x*) es positiva en el intervalo [*a*, *b*] * El área bajo la curva en el intervalo [*a*, *b*] es 1   Entonces:  *X* es una variable aleatoria continua y *f*(*x*) se llama la función de distribución de probabilidades de *X*. |

De la definición es importante aclarar que, a diferencia de las variables aleatorias discretas, es necesario que se defina una función real que cumpla con los dos requisitos para definir la variable aleatoria.

[SECCIÓN 3] **3.6.1 La distribución uniforme**

Una de las variables aleatorias más sencillas y que frecuentemente se usa en el campo del muestreo, es la distribución uniforme.

|  |  |
| --- | --- |
| **Destacado** | |
| **Título** | **La distribución uniforme** |
| **Contenido** | Sea *x* definida en el intervalo [*a*, *b*], si  MA\_11\_06\_092  Entonces *x* es una variable aleatoria continua con función de distribución uniforme.  Además:  MA\_11\_06\_093 |

La gráfica de la distribución uniforme es:

|  |  |
| --- | --- |
| **Imagen (fotografía, gráfica o ilustración)** | |
| **Código** | MA\_11\_06\_CO\_IMG13 |
| **Descripción** | Gráfica de la distribución uniforme |
| **Código Shutterstock (o URL o la ruta en AulaPlaneta)** | http://pro.arcgis.com/es/pro-app/tool-reference/data-management/GUID-3CC0E179-1226-49E7-8D1F-0AC476ED8C23-web.gif |
| **Pie de imagen**  **Inferior** | La distribución uniforme corresponde a la función lineal constante, definida en el intervalo [*a*, *b*]. |

Por ejemplo, para x que pertenece al intervalo [1, 6] una variable aleatoria con función de distribución uniforme, se tendría que:

MA\_11\_06\_094

Así, el cálculo de probabilidades se reduce a calcular el área de los rectángulos que se forman.

Es decir que:

MA\_11\_06\_095

Para valores *c* y *d* que pertenecen al intervalo [*a*, *b*]

En el ejemplo:

MA\_11\_06\_096

|  |  |
| --- | --- |
| **Practica: recurso aprovechado** | |
| **Código** | MA\_11\_06\_CO\_REC240 |
| **Título** | Hallar probabilidades para variables aleatorias continuas |
| **Descripción** | Actividad para hallar probabilidades y funciones de distribución para variables aleatorias continuas |

|  |  |
| --- | --- |
| **Practica: recurso aprovechado** | |
| **Código** | MA\_11\_06\_CO\_REC250 |
| **Título** | Caracteriza variables aleatorias continuas |
| **Descripción** | Actividad para caracterizar variables aleatorias continuas |

|  |  |
| --- | --- |
| **Practica: recurso aprovechado** | |
| **Código** | MA\_11\_06\_CO\_REC260 |
| **Título** | Obtener la probabilidad de la variable aleatoria |
| **Descripción** | Actividad para obtener probabilidades de variables aleatorias teniendo en cuenta su histograma |

|  |  |
| --- | --- |
| **Practica: recurso aprovechado** | |
| **Código** | MA\_11\_06\_CO\_REC270 |
| **Título** | Halla probabilidades para variables aleatorias con distribución normal estándar |
| **Descripción** | Actividad para resolver ejercicios y problemas para variables a aleatorias con distribución normal |

|  |  |
| --- | --- |
| **Practica: recurso aprovechado** | |
| **Código** | MA\_11\_06\_CO\_REC280 |
| **Título** | Hallar el área bajo la curva utilizando tabla |
| **Descripción** | Actividad para resolver ejercicios de áreas bajo la curva para distribuciones normales |

[SECCIÓN 2] **3.7 Consolidación**

Actividades para afianzar lo que has aprendido en esta sección.

|  |  |
| --- | --- |
| **Practica: recurso nuevo** | |
| **Código** | MA\_11\_06\_CO\_REC290 |
| **Título** | Refuerza tu aprendizaje: Las variables aleatorias |
| **Descripción** | Actividad sobre las variables aleatorias |

[SECCIÓN 1] **4** **Competencias**

Ahora que se han revisado y estudiado con detenimiento los conceptos y herramientas, es hora de ejercitar y demostrar las habilidades adquiridas.

|  |  |
| --- | --- |
| **Profundiza (recurso de ejercitación)** | |
| **Código** | MA\_11\_06\_REC300 |
| **Título** | Competencias: Probabilidades con los grupos sanguíneos |
| **Descripción** | Actividad para realizar el cálculos de combinaciones con los grupos sanguíneos |

|  |  |
| --- | --- |
| **Profundiza (recurso de ejercitación)** | |
| **Código** | MA\_11\_06\_REC310 |
| **Título** | Proyecto: La curva normal |
| **Descripción** | Interactivo que propone un proyecto para usar la curva normal para el análisis de procesos |

[SECCIÓN 1]**Fin de unidad**

|  |  |
| --- | --- |
| **Mapa conceptual** | |
| **Código** | MA\_11\_06\_REC320 |
| **Título** | Mapa conceptual |
| **Descripción** | Mapa conceptual sobre La estadística y la probabilidad |

|  |  |
| --- | --- |
| **Evaluación: recurso nuevo** | |
| **Código** | MA\_11\_04\_REC330 |
| **Título** | Evaluación |
| **Descripción** | Actividades para evaluar La estadística y la probabilidad |

|  |  |  |
| --- | --- | --- |
| **Webs de referencia** | | |
| **Código** | MA\_11\_06\_REC340 | |
| **Web 01** | Ejemplo de aplicación de la distribución binomial | https://www.youtube.com/watch?v=EisaSQ1j\_Kk |
| **Web 02** | Aspectos importantes de las variables aleatorias | https://www.youtube.com/watch?v=n0T\_HcJ7oak |
| **Web 03** | Problemas resueltos sobre variables aleatorias | http://www.ugr.es/~mvargas/PTema3.pdf |