INTRODUCCIÓN A LA TEORÍA DE PROBABILIDADES CON PYTHON

Marcia M. Bayas, Ronald H. Rovira,
Lídice V. Haz, Marllelis Guitiérrez

Introducción a la teoría de probabilidades

Marcia M. Bayas
Ronald H. Rovira
Lídice V. Haz
Marllelis Guitiérrez

Diciembre, 2022

La presente obra fue revisada por pares académicos externos ciegos conforme al proceso editorial de Editorial CILADI.

ISBN: 978-9942-7078-2-6

Primera edición: Diciembre 2022
Ediciones
ISBN: 978-9942-7078-2-6
Este texto ha sido sometido
a evaluación de pares externos
con base en la normativa de la editorial.

Editor general:
Antonio Poveda G.

Autores
Marcia M. Bayas
Ronald H. Rovira
Lídice V. Haz
Marllelis Guitiérrez

Dedicado a nuestros
familiares y a nuestros
estudiantes

Contenido

Prólogo

La teoría de la probabilidad y la estadística ocupan un lugar importante en la mayoría de las ciencias. Sus aplicaciones van desde el modelado de fenómenos naturales y sociales de carácter aleatorio y sus conclusiones son una base fundamental para la toma de decisiones en situaciones de conflicto. Por lo tanto, en el plan de estudios de ingeniería y específicamente en tecnología de la información, electrónica y automatización y telecomunicaciones se debe poner énfasis en que los estudiantes aprendan y dominen los conceptos de esta asignatura. En este libro se presentan de manera concisa y práctica los conceptos fundamentales de la teoría de la probabilidad. El objetivo de este trabajo es proporcionar a los estudiantes una guía para la resolución de problemas típicos. En esta guía se incluyen ejemplos que van acompañados de soluciones algorítmicas en Python con lo que esperamos un incremento de la aceptación e interés de la asignatura por parte de los estudiantes.

Capítulo 1

Eventos aleatorios

La teoría de las probabilidades es el conjunto de leyes que explican los fenómenos aleatorios: eventos aleatorios, variables aleatorias, sus funciones, propiedades y operaciones sobre ellos. La teoría de la probabilidad surgió y se desarrolló por los aportes de:

- Blaise Pascal[1];
- Pierre Fermat;
- Christiaan Huygens;
- Jacob Bernoulli;

[1]Se atribuye a Pascal y Fermat el origen de la teoría de la probabilidad, sin embargo, Gerolamo Cardano escribió un manual de juegos de azar un siglo antes [1].

- Pierre-Simon Laplace;
- Pafnuti Chebyshov;
- Andrei Kolmogorov, entre otros.

1.1 Conceptos básicos

Un experimento aleatorio es un experimento cuyos posibles resultados son conocidos, pero no se puede establecer de antemano cuál de los posibles valores toma al ser ejecutado. Por ejemplo, el lanzamiento de una moneda. Sabemos que puede salir 'cara' o 'cruz' pero no sabemos si caerá de cara o cruz. Otros ejemplos de experimentos aleatorios pueden ser:

- el lanzamiento de un dado;
- la extracción de una carta;
- la selección de un objeto entre otros.

Al conjunto de los posibles resultados de un experimento aleatorio se le denomina espacio muestral. Por lo general, al espacio muestral se lo denota con la letra griega omega Ω[1].

[1] En algunos textos se usa la letra mayúscula S

Si se trata del lanzamiento de una moneda el espacio muestral es:

$$\Omega = \{\text{cara}, \text{cruz}\}$$

En el caso del lanzamiento de un dado el espacio muestral es:

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

La cardinalidad[1] de un conjunto es el número de elementos de un conjunto. En el ejemplo anterior la cardinalidad del espacio muestral es

$$N(\Omega) = 6$$

Un evento aleatorio es un subconjunto del espacio muestral que obedece a determinadas condiciones. En el lanzamiento de una moneda un ejemplo de evento podría ser que la moneda caiga de cara.

$$E = \{\text{la moneda cae de cara}\}$$

Obviamente este subconjunto tiene un solo elemento:

$$E = \{\text{cara}\}$$

y su cardinalidad es:

[1] En la teoría de conjuntos la cardinalidad de un conjunto se denota por dos barras $|A|$.

$$N(E) = 1$$

En el lanzamiento de un dado podríamos tomar como evento que salga un número impar:

$$E = \{1{,}3{,}5\}$$

y su carinalidad es:

$$N\left(E\right) = 3$$

La cardinalidad de un evento siempre es menor o igual a la cardinalidad del espacio muestral:

$$N\left(E\right) \leq N\left(\Omega\right)$$

La probabilidad de un evento es una medida de certidumbre de que ocurra el evento y se determina por el cociente entre las cardinalidades del evento y el espacio muestral:

$$p(E) = \frac{N(E)}{N(\Omega)} \tag{1.1}$$

De esta definición se deduce que la probabilidad de un evento es un número entre 0 y 1. Si el evento siempre ocurre la probabilidad es 1 mientras que si el evento nunca ocurre la probabilidad es 0.

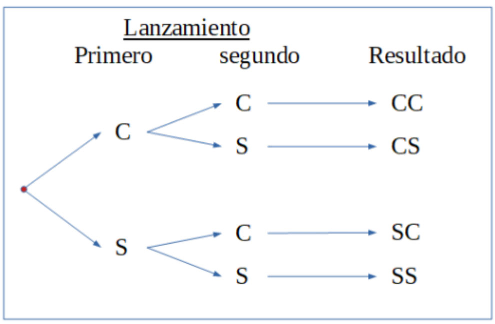
Ejemplo 1. Un experimento aleatorio consiste en lanzar sucesivamente dos veces una moneda y registrar la salida en ambas ocasiones, determine por enumeración el espacio muestral, y la probabilidad de que la moneda caiga las dos veces de cara.

Resolución

Como la moneda se lanza dos veces los registros de la salida del experimento son duplas, por ejemplo, 'cara-cara'. Por lo tanto el espacio muestral es:

$$\Omega = \{\text{cara}-\text{cara},\ \text{cara}-\text{cruz},\ \text{cruz}-\text{cara},\ \text{cruz}-\text{cruz}\}$$

En forma gráfica[1] se puede obtener de la siguiente manera:



El evento E está definido por:

$E = \{\text{cara}-\text{cara}\}$,

[1]A esta forma gráfica se le llama diagrama de árbol; permite determinar el espacio muestral de experimentos aleatorios complejos.

y su probabilidad $p\,(E)$ es entonces:

$$p(E) = \frac{N(E)}{N(\Omega)} = \frac{1}{4} = 0.25$$

1.2 Python

Python es un lenguaje de programación de propósito general de alto nivel. Para acceder a este programa se puede optar por una de las siguientes maneras:

- Python en internet: https: //www.programiz.com/python-programming/online-compiler/

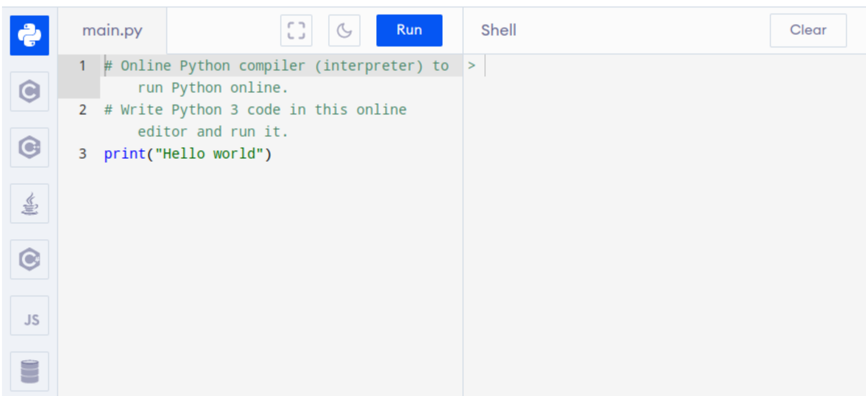


Figura 1.1: Programiz Python Online Compiler

- Descarga de PyCharm: //www.jetbrains.com/es-es /pycharm/download

- Descarga de Anaconda: //www.anaconda.com/products/distribution
- Spyder IDE: https://www.spyder-ide.org

Algoritmo 1.1 Ejemplo del editor de Spyder

```
#!/usr/bin/env python3
# -*- coding: utf-8 -*- """
Created on Mon May 30 08:02:24 2022

@author: PyE

""" print('Bienvenidos')
```

Sintaxis básica de Python

En la escritura de programas usualmente tenemos que emplear: comentarios, operaciones aritméticas, operaciones lógicas y de comparación, estructuras condicionales, estructuras repetitivas, funciones y una diversidad de tipos de datos.

Los comentarios permiten explicar las distintas partes del código. En Python los comentarios se inician con el símbolo #.

Las operaciones aritméticas en Python se expresan de la siguiente forma:

- adición A+B;
- substracción A-B;
- multiplicacón A*B;
- división A/B;
- exponenciación A**B.

Operaciones lógicas y de comparación.

Las operaciones lógicas y de comparación dan como resultado los valores booleanos: True (la expresión es verdadera) y False (la expresión es falsa).

Los operadores de comparación en Python se expresan de la siguiente manera:

- igual que $a == b$;
- no es igual que $a! = b$;
- menor que $a < b$;
- menor o igual que $a <= b$;
- mayor que $a > b$;
- mayor o igual que $a >= b$.

Los operadores lógicos:

- conjunción *and*;
- disyunción *or*;
- negación *not.*

Los bloques de código en las estructuras de control se escriben con sangría. La estructura condicional se usa en las operaciones de decisión. En Python tiene la forma:

```
if condición:
    grupo de ordenes
else:
    grupo de ordenes
```

Ejemplo 2. Escriba un código para elegir el menor de 2 números.

Resolución

Algoritmo 1.2 Estructura condicional

```
#asignación de valores
x = 2
y = 4
if x < y:
    print('El número menor es: ', x)
else:
    print('El número menor es: ', y)
```

A veces es necesario ejecutar un bloque de código si una de varias condiciones se cumple. En este caso se usan las estructuras condicionales en cascada las cuales tienen la siguiente forma:

```
if condición 1:
    grupo de ordenes
elif condición 2:
        grupo de ordenes
elif condición 3:
        grupo de ordenes
else:
        grupo de ordenes
```

Ejemplo 3. Escriba un código para determinar en que cuadrante se encuentra un punto ingresado por teclado.

Resolución

Algoritmo 1.3 *Estructura condicional en cascada*

```
x, y = int(input()), int(input())
if x > 0 and y > 0:
    print("primer cuadrante")
elif x < 0 and y > 0:
    print("segundo cuadrante")
elif x < 0 and y < 0:
    print("tercer cuadrante")
elif x > 0 and y < 0:
    print("cuarto cuadrante")
else:
    print("P se encuentra en los ejes...
 o en el centro de coordenadas.")
```

Las estructuras de control de repetición ejecutan ciclicamente un grupo de órdenes según ciertas condiciones. La estructura de repetición *while* tiene la forma:

```
while condición:
        grupo de ordenes
```

Ejemplo 4. Escriba un código para sumar los primeros 6 números naturales.

Resolución

Algoritmo 1.4 *Estructura de repetición 'while'*

```
número, suma = 1, 1
while número <= 5:
    número += 1
    suma = suma + número
    print (f'La suma de los {número}...
     primeros números es {suma}')
```

El resultado de la ejecución de este código es:

```
La suma de los 2 primeros números es 3
La suma de los 3 primeros números es 6
La suma de los 4 primeros números es 10
La suma de los 5 primeros números es 15
La suma de los 6 primeros números es 21
```

Las repeticiones también se puede implementar con los ciclos *for*. La diferencia con un ciclo *while* consiste en que en el ciclo *for* se recorre los elementos de un objeto iterable[1]. Una estructura de repetición 'for' tiene la forma:

```
for 'elemento' in 'iterable':
        grupo de ordenes
```

Ejemplo 5. Escriba un código para imprimir los elementos de la lista ['Ana', 'Ben', 'Ruth', 'Tom'].

[1] En python un objeto iterable puede ser una lista, tupla o un diccionario

Resolución

Algoritmo 1.5 *Estructura de repetición 'for'*

```
Estudiantes = ['Ana', 'Ben', 'Ruth', 'Tom']
for estudiante in Estudiantes:
    print(estudiante)
```

El resultado de la ejecución de este código es:

Ana
Ben
Ruth
Tom

Determinación del espacio muestral y los eventos

El espacio muestral y los eventos como cualquier conjunto se pueden determinar de dos maneras:

- Por extensión o
- por comprensión.

Un conjunto está definido por extensión si se enumeran sus elementos. Por ejemplo el conjunto de los colores primarios[1]:

[1]Existen varios modelos de colores primarios: RYB, CMY, RGB.

$$A = \{'rojo', 'verde', 'azul'\}$$

Para enumerar los elementos de un espacio muestral dado por experimentos de tipo lanzamiento, por ejemplo monedas, dados etc., se puede usar una serie de ciclos *for* anidados.

Ejemplo 6. Escriba un código para determinar por enumeración el espacio muestral del experimento que consiste en lanzar una moneda dos veces.

Resolución

Algoritmo 1.6 Enumeración de conjuntos mediante ciclos *for*

```
posibles = {'cara','cruz'}
espacio_muestral = set()
for tiro1 in posibles:
    for tiro2 in posibles:
        salida = (tiro1 , tiro2)
        espacio_muestral.add(salida)
print(espacio_muestral)
```

El resultado de la ejecución de este código es:

```
{('cara', 'cara'),
 ('cara', 'cruz'),
 ('cruz', 'cruz'),
 ('cruz', 'cara')}
```

Análisis del código

- {e1, e2, ...} - define un conjunto por enumeración;
- set() - crea un conjunto vacío;
- for x in X: - x toma los valores de X iterativamente;
- conjunto.add - añade un elemento a conjunto.

El espacio muestral en este ejemplo corresponde al producto cartesiano[1] del conjunto original por si mismo. Por ejemplo, en el lanzamiento de dos dados el producto cartesiano es el siguiente:

Primer lanzamiento \ Segundo lanzamiento	1	2	3	4	5	6
1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Cuadro 1.1: Producto cartesiano de un conjunto por si mismo

Si se quisiera lanzar la moneda tres veces o más veces se podría usar más ciclos *for*

```
for a in X:
    for b in X:
        .
        .
            for w in X:
```

[1] El producto cartesiano es el conjunto de todos los pares ordenados que se pueden formar a partir de dos conjuntos.

Sin embargo, no sería muy eficiente. Una forma más conveniente es usar el módulo itertools.

Ejemplo 7. Escriba un código para determinar por enumeración el espacio muestral correspondiente al lanzamiento triple de una pirinola de tres caras, cuyos rótulos son: 'A', 'B', 'C'.

Resolución

Algoritmo 1.7 Enumeración de conjuntos con el módulo *itertools*

```
from itertools import product
posibles = {'A','B','C'}
espacio_muestral = set(product(posibles,
    repeat = 3))
print(espacio_muestral)
```

El resultado de la ejecución de este código es:

```
{('B', 'B', 'B'), ('B', 'C', 'B'),
 ('C', 'A', 'A'), ('C', 'B', 'A'),
 ('A', 'A', 'A'), ('C', 'C', 'A'),
 ('A', 'C', 'A'), ('C', 'A', 'C'),
 ('C', 'B', 'C'), ('A', 'B', 'A'),
 ('B', 'A', 'A'), ('A', 'A', 'C'),
 ('B', 'B', 'A'), ('C', 'C', 'C'),
 ('B', 'C', 'A'), ('A', 'C', 'C'),
```

```
('A', 'B', 'C'), ('B', 'A', 'C'),
('B', 'B', 'C'), ('B', 'C', 'C'),
('C', 'A', 'B'), ('C', 'B', 'B'),
('A', 'A', 'B'), ('C', 'C', 'B'),
('A', 'C', 'B'), ('B', 'A', 'B'),
('A', 'B', 'B')}
```

Análisis del código

- from módulo[1] import función - importa solo una función de un módulo;
- itertools - módulo que contiene funciones iterativas;
- product - función que implementa el producto cartesiano;
- repeat = n - establece las veces que se ejecuta 'product'.

Ejemplo 8. Escriba un código para determinar por enumeración el espacio muestral del lanzamiento triple de un dado.

[1]En Python existen muchos módulos. Los de uso más frecuente son Numpy, itertools, matplotlib, etc.

Resolución

Algoritmo 1.8 Enumeración de conjuntos con el módulo *itertools*

```
from itertools import product
posibles = list(range(1,7))
espacio_muestral = set(product(posibles,
  repeat = 3))
print(espacio_muestral)
```

El resultado (fragmento) de la ejecución de este código es:

```
{(5, 1, 6), (5, 3, 3), (5, 4, 2),
 (2, 1, 6), (1, 6, 6), (2, 2, 5),
 (6, 6, 4), ...
```

Análisis del código

- list() - crea una lista[1] de elementos indexados;
- range(n) - crea una secuencia de números de 0 a n-1.

Un conjunto está definido por compresión, si sus elementos se describen a través de propiedades que tienen en común.

[1]En una lista se almacenan múltiples elementos indexados.

Los elementos que integran a un evento son aquellos elementos del espacio muestral que satisfacen una condición C determinada, que en notación formal se escribe[1]:

$$E = \{\forall x \in \Omega \,/\, x \text{ satisface } C\} \tag{1.2}$$

Por ejemplo, en el lanzamiento de un dado podría ser de interés que el puntaje de salida sea menor que 4. Entonces la expresión formal o definición por comprensión de este evento es:

$$E = \{\forall x \in \Omega \,/\, x < 4\}$$

Ejemplo 9. Escriba un código en Python para enumerar el evento $E = \{\text{puntaje menor que } 4\}$ a partir de su descripción formal en el lanzamiento de un dado.

Resolución

Algoritmo 1.9 Determinación de eventos por comprensión

```
espacio_muestral = set(range(1,7))
E = set([x for x in espacio_muestral
  if x < 4])
print(f"E = {E}\n")
```

[1]En la descripición de conjuntos se usan los cuantificadores: $\forall$ se lee 'para todo'.

El resultado de la ejecución de este código es:

```
E = {1, 2, 3}
```

Análisis del código

- print(f"E = {E}\n") - imprime la salida con formato;
- {E} es el valor de la variable a mostrar y
- \n es un salto de renglón.

Ejemplo 10. Escriba un código en Python para determinar los elementos de un evento si debe cumplir la condición que la suma de los puntos en dos lanzamientos de un dado no sea mayor que 7.

Resolución

Algoritmo 1.10 Determinación de eventos por comprensión

```
from itertools import product
posibles = list(range(1,7))
espacio_muestral = set(product(posibles,..
  repeat = 2))
E = set([x for x in espacio_muestral...
  if sum(x) < 7])
print(E)
```

El resultado de la ejecución de este código es:

```
{(2, 4), (1, 2), (2, 1), (1, 5), (3, 1),
 (4, 1), (1, 1), (5, 1), (4, 2), (1, 4),
 (2, 3), (3, 3), (2, 2), (3, 2), (1, 3)}
```

Análisis del código

- sum(x) - suma los elementos de un conjunto

Ejemplo 11. Determine los elementos de un evento si debe cumplir la condición que la suma de los puntos en dos tiros sea múltiplo de 3.

Resolución

Algoritmo 1.11 Determinación de eventos por comprensión

```
from itertools import product
posibles = list(range(1,7))
espacio_muestral = set(product(posibles,..
   repeat = 2))
E = set([x for x in espacio_muestral...
   if sum(x)%3 == 0])
print(E)
```

El resultado de la ejecución de este código es:

```
{(2, 4), (1, 2), (2, 1), (1, 5), (5, 4),
 (5, 1), (4, 2), (4, 5), (3, 3), (3, 6),
 (6, 6), (6, 3)}
```

Análisis del código

- a%b - es el operador módulo que da como resultado el residuo del cociente de a y b.

Ejemplo 12. Escriba un programa que calcule la probabilidad de que el producto de los puntos de tres lanzamientos de los dados sea menor que 50.

Resolución

Algoritmo 1.12 Determinación de eventos por comprensión

```
import numpy as np
from itertools import product
posibles = list(range(1,7))
espacio_muestral = set(product(posibles,...
   repeat = 3))
E = set([x for x in espacio_muestral...
  if np.prod(x) < 50])
p = len(E)/len(espacio_muestral)
print(f"La probabilidad de que ocurra...
  E es {p:.4f}")
```

El resultado de la ejecución de este código es:

```
La probabilidad de que ocurra E es 0.6944
```

Análisis del código

- numpy[1] - módulo de Python que permite realizar cálculos científicos;
- np.prod - calcula el producto de los elementos de un conjunto;
- len(E) - función que entrega el número de elementos de un objeto;
- {p:.4f} - en la impresión con formato muestra la salida con cuatro cifras decimales.

1.3 Definiciones de probabilidad

Existen varias definiciones de probabilidad:

- clásica;
- geométrica;

[1]En Python es común referirse a los módulos importados con un alias: numpy - np, itertools - it, matplotlib.pyplot - plt, etc.

- subjetiva;
- frecuencial o frecuentista.

Probabilidad clásica

En los ejemplos resueltos hasta el momento se ha utilizado la definición clásica[1]:

La probabilidad de que un evento ocurra es igual a la razón entre el número de casos favorables al evento y el número total de los casos posibles.

Las siguientes propiedades de la probabilidad se derivan de su definición:

1. la probabilidad de un evento cierto es igual a uno;
2. la probabilidad de un evento imposible es cero;
3. la probabilidad de un evento aleatorio es un número positivo entre cero y uno;
4. la probabilidad de que ocurran eventos que formen un grupo completo[2] es igual a uno.

[1] La definición clásica de probabilidad se asocia con los trabajos de Jacob Bernoulli y Pierre-Simon Laplace.

[2] Se dice que un grupo es completo o colectivamente exhaustivo si por lo menos uno de los eventos ocurre cuando se realiza un experimento.

Probabilidad geométrica

Es la probabilidad de que un punto golpee una región determinada, como en el lanzamiento de un dardo.

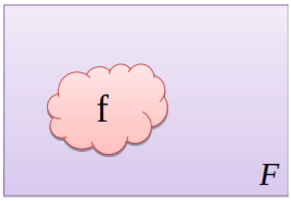


Figura 1.2: Probabilidad geométrica

La probabilidad geométrica se calcula de la siguiente manera:

$$p(E) = \frac{S(f)}{S(F)} \tag{1.3}$$

Donde $S(f)$ es el área a acertar y $S(F)$ el área de la región total.

Ejemplo 13. Se lanza desde lo lejos una piedra, que debe caer en un círculo. Un cuadrado de $200\,cm^2$ encierra al círculo, cuya área es de $30\,cm^2$. Determine la probabilidad de acertar en el círculo.

Resolución

La probabilidad de acertar en el círculo es el cociente entre las áreas.

$$p(E) = \frac{30}{200} = 0.15$$

Probabilidad subjetiva

La probabilidad subjetiva es el grado de creencia personal de que ocurra algún evento. El concepto de probabilidad subjetiva se aplica en la teoría de toma de decisión.

Ejemplo 14. Podemos mencionar los siguientes:

- la probabilidad de que Pedro llegue a ser presidente;
- la probabilidad de que llueva;
- la probabilidad de que la selección gane el mundial.

En cualquiera de estos casos si se pregunta a varias personas se obtiene distintas respuestas que están basadas en criterios subjetivos como gusto o disgusto.

Probabilidad frecuencial

La probabilidad frecuencial, frecuentista o empírica es la probabilidad de que un evento ocurra en un número grande de ensayos. Se calcula como el cociente entre el número de ocurrencias de un evento, n_e y el número total de ensayos[1] N en un experimento real.

$$p(E) = \lim_{N \to \infty} \frac{n_e}{N} \tag{1.4}$$

Ejemplo 15. Se lanza mil veces una moneda. En 300 lanzamientos sale cara. ¿Cuál es la probabilidad de que salga cruz?

Resolución

Definición de evento:

$$E = \{\text{sale 'cruz'}\}$$

El número de ocurrencias de que salga 'cruz' es:

$$N(E) = 1000 - 300$$

Por lo tanto,

$$p(E) = \frac{700}{1000} = 0.7$$

Ejemplo 16. Cálculo de probabilidad frecuentista de una moneda con sesgo[2].

[1] La probabilidad frecuentista es el límite de la frecuencia relativa cuando el número de experimentos tiende al infinito.

[2] Se dice que una moneda está segada (con truco) si la probabilidad de que caiga de un lado es mayor que la probabilidad que caiga del otro lado.

Resolución

Algoritmo 1.13 Cálculo de la probabilidad frecuentista

```
import random as rd
d =['cara','cara','cara','cara','cruz']
N_e = 0
n = 10000
for i in range(n):
    if rd.choice(d) == 'cara':
        N_e += 1
print(f'p(cara) es {N_e/n:.4f}')
```

El resultado de la ejecución de este código es:

```
p(cara) es 0.8017
```

Análisis del código

- random - módulo de Python que permite generar números pseudoaleatorios y realizar acciones aleatorias, como elegir un valor aleatorio de una lista.
- rd.choice() - selecciona aleatoriamente un elemento de una secuencia especificada.

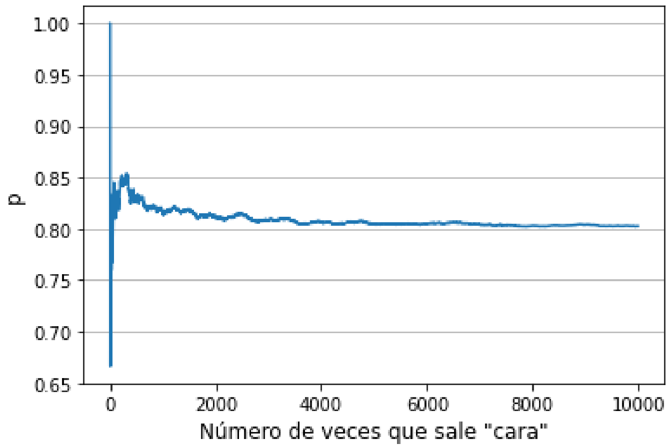


Figura 1.3: Probabilidad frecuentista

1.4 Técnicas de conteo

El cálculo de la probabilidad de que ocurra un evento es en esencia un conteo. En los casos hasta ahora vistos estas cuentas no eran muy complejas y se podía enumerar tanto el espacio muestral como los eventos. Sin embargo, en la mayoría de casos reales enumerar los eventos resulta muy complejo. Las técnicas de conteo son un conjunto de métodos que permiten contar el número posible de ordenamientos de los elementos de un conjunto y se basan en dos principios:

- multiplicativo y
- aditivo.

El principio multiplicativo

Si una actividad involucra varios pasos, donde el primer paso se puede realizar de N_1 formas, el segundo paso de N_2 y el "r"-ésimo paso de N_r formas entonces la actividad se puede realizar de N maneras:

$$N = N_1 \times N_2 \times \cdots \times N_r \tag{1.5}$$

Ejemplo 17. Determine el número de formas en las que se puede vestir un joven si dispone de dos pares de zapatos, dos pantalones y tres camisas.

Resolución

El número de formas se puede calcular mediante la regla de la multiplicación

$$N = 2 \times 2 \times 3 = 12$$

El joven se puede vestir de 12 maneras diferentes.

Ejemplo 18. Escriba el código en Python para crear una lista de las combinaciones posibles en las que se puede vestir el joven del ejemplo anterior.

Resolución

Algoritmo 1.14 Determinación de eventos por comprensión

```
from itertools import product
camisas = {'c1','c2','c3'}
pantalones = {'p1','p2'}
zapatos = {'z1','z2'}
espacio_muestral = set(product(camisas,...
 pantalones, zapatos))
for combinacion in espacio_muestral:
    print(combinacion)
```

El resultado de la ejecución de este código es:

```
('c3','p1','z2')  ('c1','p1','z1')
('c1','p2','z1')  ('c2','p1','z2')
('c3','p2','z1')  ('c1','p2','z2')
('c1','p1','z2')  ('c2','p2','z1')
('c3','p2','z2')  ('c2','p2','z2')
('c3','p1','z1')  ('c2','p1','z1')
```

El principio aditivo

Si existen varias alternativas A_i de elegir un objeto y en cada una de estas hay a su vez $N(A_i)$ opciones el número total de formas de elegir el objeto es la suma del número de posibilidades de cada alternativa.

$$N = \sum_{i=1}^{k} N(A_i) \tag{1.6}$$

.

Ejemplo 19. Determine el número de formas en las que se puede comprar una computadora. Existen tres marcas disponibles en el mercado. La primera marca ofrece computadoras de 14" con disco duro o disco sólido. La segunda marca ofrece computadoras de 11", 14" y 16" con y sin pantalla táctil y la tercera marca ofrece todas las variantes mencionadas.

Resolución

El número de formas posibles en la compra de la computadora de la primera marca se puede calcular mediante la regla de la multiplicación

$$N(A_1) = 1 \times 2 = 2$$

Si se compra de la segunda marca

$$N(A_2) = 3 \times 2 = 6$$

y si se compra de la tercera marca

$$N(A_3) = 3 \times 2 \times 2 = 12$$

El número total de alternativas es

$$N = 2 + 6 + 12 = 20$$

Combinaciones y permutaciones

En muchos casos es necesario reordenar aleatoriamente los elementos de un conjunto, como el caso del barajado de cartas antes de un juego de azar. Supongamos que se debe reordenar solo tres cartas de corazones no 'barajadas'

$$\{J, Q, K\}$$

Ordenar estas cartas es como colocarlas en tres casillas vacías. Para llenar la primera casilla se tiene tres alternativas. Para llenar la segunda casilla se tiene 2 alternativas, en vista de que ya se usó una carta. Finalmente para la última casilla solmente se tiene una alternativa. De acuerdo al principio de multiplicación el número total de alternativas es:

$$N = 3 \times 2 \times 1 = 6$$

En términos generales la expresión anterior se puede escribir como:

$$n! = n \times (n-1) \times (n-2) \times \cdots \times 2 \times 1 \times 0! \qquad (1.7)$$

$n!$ es el factorial de un número[1].

Ejemplo 20. Escriba un código para barajar tres cartas de corazones. Las cartas son $\{J, Q, K\}$

Resolución

Algoritmo 1.15 Ordenamiento de elementos de un conjunto

```
import itertools as it
elementos =['J', 'Q', 'K']
for permutacion in ...
 list(it.permutations(elementos)):
    print (permutacion)
```

El resultado de la ejecución de este código es:

```
('J', 'Q', 'K')
('J', 'K', 'Q')
('Q', 'J', 'K')
('Q', 'K', 'J')
('K', 'J', 'Q')
('K', 'Q', 'J')
```

[1] Por definición el factorial de 0 es igual a 1.

Análisis del código

- permutations() - forma los ordenamientos posibles, sin elementos repetidos.

Una permutación[1] es un arreglo ordenado de elementos tomados de un conjunto. El número posible de permutaciones de r elementos tomados de un conjunto de n elementos se halla mediante la expresión matemática:

$$P_r^n = \frac{n!}{(n-r)!} \tag{1.8}$$

Si el número de elementos a ordenar coincide con el número de elementos del conjunto original entonces el número de permutaciones es igual al factorial.

Ejemplo 21. Determine el número posible de permutaciones de 3 elementos tomados de un conjunto de 7 elementos.

Resolución

El número de permutaciones posibles se obtiene directamente de la ecuación 1.8

$$P_3^7 = \frac{7!}{(7-3)!}$$

$$P_3^7 = \frac{7 \times 6 \times 5 \times 4!}{4!} = 210$$

[1]Existen registros de que 1000 a. C, en China, se conocía la permutación.

Ejemplo 22. Determine el número posible de códigos de seguridad de cuatro dígitos que se pueden formar con los dígitos del 0 al 9 sin que estos se repitan.

Resolución

Como los dígitos no se pueden repetir entonces la tarea es encontrar el número posible de formas de seleccionar 4 elementos de entre 10.

$$P_4^{10} = \frac{10!}{(10-4)!}$$

$$P_4^{10} = \frac{10 \times 9 \times 8 \times 7 \times 6!}{6!}$$

Se pueden formar 5040 códigos diferentes.

Ejemplo 23. El horario de un colegio tiene 5 horas académicas. Determine la cantidad de alternativas posibles de elaboración de este horario si un curso tiene 11 asignaturas.

Resolución

El número de alternativas posibles corresponde al número de permutaciones de 5 elementos tomados de 11.

$$P_5^{11} = \frac{11!}{(11-5)!}$$

$$P_5^{11} = \frac{11 * 10 * 9 * 8 * 7 * 6!}{6!}$$

El horario se puede elaborar de 55440 formas diferentes.

Ejemplo 24. Determine el número de formas que se pueden sentar 4 personas en una mesa de cuatro puestos.

Resolución

En este caso se toman todos los elementos del conjunto original. Por lo tanto, el número de alternativas posibles corresponde al factorial del número.

$$P_4^4 = \frac{4!}{(4-4)!} = \frac{4 \times 3 \times 2 \times 1 \times 0!}{0!}$$

Son 24 alternativas posibles.

Las permutaciones son arreglos ordenados. Ahora veámos el caso de que el arreglo no sea ordenado. Llenemos una bolsa opaca con tres letras tomadas del conjunto $\{A, B, C, D\}$. ¿De cuántas formas diferentes se puede llenar la bolsa?

ABC	ABD	ACD	BCD

Cuadro 1.2: Alternativas de llenado de la bolsa

Como el arreglo no es ordenado da lo mismo en que posición vaya una letra. Por lo tanto, todas las formas siguientes son equivalentes.

$$\text{Formas equivalentes} \begin{cases} ABC \\ ACB \\ BAC \\ BCA \\ CBA \\ CAB \end{cases}$$

Cuando en el arreglo no importa el orden se llama combinación. El número de combinaciones posibles de n objetos tomados en grupos de r elementos es:

$$C_r^n = \frac{n!}{r!(n-r)!} \tag{1.9}$$

Ejemplo 25. Determine el número posible de combinaciones de 8 elementos tomados en grupos de 4.

Resolución

El número posible de combinaciones se obtiene aplicando la ecuación 1.9:

$$C_4^8 = \frac{8!}{4!(8-4)!}$$

$$C_4^8 = \frac{8 \times 7 \times 6 \times 5 \times 4!}{4 \times 3 \times 2 \times 1! \times 4!}$$

El número posible de combinaciones es 70.

Ejemplo 26. La estantería de una biblioteca tiene 23 libros: 10 de ciencia ficción y 13 de arte contemporáneo. ¿De cuántas maneras se pueden tomar dos libros del mismo género literario?

Resolución

Al elegir los libros solamente interesa el género; el orden no es relevante y por lo tanto, el número de alternativas de acuerdo al principio de adición es:

$$N = C_2^{10} + C_2^{13}$$

C_2^{10} es el número de alternativas de elegir dos libros de ciencia ficción,

C_2^{13} es el número de alternativas de elegir dos libros de arte contemporáneo.

$$N = \frac{10!}{2!\,(10-2)!} + \frac{13!}{2!\,(13-2)!}$$

$$N = \frac{10 \times 9 \times 8!}{2!\,(8)!} + \frac{13 \times 12 \times 11!}{2!\,(11)!}$$

$$N = 45 + 78 = 123$$

Son 123 alternativas posibles.

Ejemplo 27. Escriba un código que genere las permutaciones del conjunto $A = \{1, 2, 3\}$ y las combinaciones tomando dos elementos a la vez.

Resolución

Algoritmo 1.16 Combinaciones y permutaciones

```
import itertools as it
A = [1, 2, 3]
perm = it.permutations(A)
comb = it.combinations(A, 2)
print('Las permutaciones de', A,'son:')
for i in list(perm):
    print (i)
print('Las combinaciones de', A,...
 'tomando 2 elementos son:')
for i in list(comb): print(i)
```

El resultado de la ejecución de este código es:

```
Las permutaciones de [1, 2, 3] son:
(1, 2, 3)
(1, 3, 2)
(2, 1, 3)
(2, 3, 1)
(3, 1, 2)
(3, 2, 1)
Las combinaciones de [1, 2, 3] tomando
```

```
2 elementos son:
(1, 2) (1, 3) (2, 3)
```

Ejemplo 28. Escriba un código en Python para calcular el número de permutaciones y combinaciones de 2 elementos tomados de un conjunto de 10 elementos.

Resolución

Algoritmo 1.17 Combinaciones y permutaciones

```
import math
p = math.perm(10,2)
c = math.comb(10,2)
print(f'El número de permutaciones de 2...
 elementos tomados de 10 es {p}')
print(f'El número de combinaciones de 2...
 elementos tomados de 10 es {c}')
```

El resultado de la ejecución de este código es:

```
El número de permutaciones de 2 elementos
tomados de 10 es 90
El número de combinaciones de 2 elementos
tomados de 10 es 45
```

Ejemplo 29. Se requieren estudiantes destacados para dictar clases de tutorías: 4 estudiantes de física, 6 estudiantes de química y 3 para clases generales. Entre los cursos avanzados hay 6 estudiantes de física y 8 de química. ¿De cuántas formas se pueden seleccionar a los tutores?

Resolución

El número de alternativas para elegir a 4 tutores de física es:

$$N_1 = C_4^6 = \frac{6!}{4!\,(6-4)!} = 15$$

El número de alternativas para elegir a 6 tutores de química es:

$$N_2 = C_6^8 = \frac{8!}{6!\,(8-6)!} = 28$$

Para las clases generales se seleccionan 3 estudiantes de 4 disponibles, por lo tanto el número de alternativas es:

$$N_3 = C_3^4 = \frac{4!}{3!\,(4-3)!} = 4$$

De acuerdo al principio de la multiplicación el número total de alternativas es:

$$N = N_1 \times N_2 \times N_3$$

El cálculo se puede hacer en Python

Algoritmo 1.18 Ejemplo 29

```
from math import comb
c = comb(6,4)*comb(8,6)*comb(4,3)
print(f'El número de combinaciones...
 posibles es {c}')
```

El resultado de la ejecución de este código es:

```
El número de combinaciones posibles es 1680
```

Ejemplo 30. Una urna contiene 8 bolas blancas, 4 negras y 7 rojas. Se extraen 3 bolas al azar. Encuentre la probabilidad de sacar al menos 1 bola roja.

Resolución

La condición es que se saque al menos una bola roja entonces las siguientes son alternativas válidas:

$$RCC \quad RRC \quad RRR$$

R - bola roja

C - bola no roja

De acuerdo al principio de adición el número total de alternativas es

$$N(E) = N(RCC) + N(RRC) + N(RRR)$$

Encontrar el número de alternativas de sacar una bola roja y dos de otro color es equivalente a seleccionar una bola roja entre 7 y dos bolas de otro color entre 12 y luego por el principio de la multiplicación se tiene:

$$N\left(RCC\right) = C_1^7 \times C_2^{12}$$

$$N\left(RCC\right) = \frac{7!}{1!\left(7-1\right)!} \times \frac{12!}{2!\left(12-2\right)!} = 462$$

Las otras variantes son:

$$N\left(RRC\right) = C_2^7 \times C_1^{12}$$

$$N\left(RRC\right) = \frac{7!}{2!\left(7-2\right)!} \times \frac{12!}{1!\left(12-1\right)!} = 252$$

$$N\left(RRR\right) = C_3^7 \times C_0^{12}$$

$$N\left(RRR\right) = \frac{7!}{3!\left(7-3\right)!} \times \frac{12!}{0!\left(12-0\right)!} = 35$$

El número total de alternativas es:

$$N\left(E\right) = 462 + 252 + 35 = 749$$

El número de alternativas de sacar tres bolas de la urna es:

$$N\left(\Omega\right) = C_3^{19} = \frac{19!}{3!\left(19-3\right)!} = 969$$

La probabilidad de sacar al menos una bola roja de la urna es:

$$p(E) = \frac{N(E)}{N(\Omega)}$$

$$p(E) = \frac{749}{969} = 0.7730$$

Este resultado se lo puede verificar en Python

Algoritmo 1.19 Ejemplo 30

```
from math import comb
c1 = comb(7,1)*comb(12,2)
c2 = comb(7,2)*comb(12,1)
c3 = comb(7,3)*comb(12,0)
c = c1+c2+c3
N = comb(19,3)
print(f'p(E) es {c/N:.4f}')
```

El resultado de la ejecución de este código es:

```
p(E) es 0.7730
```

Ejemplo 31. Determine la probabilidad de sacar no más de dos bolas rojas de la urna mencionada en el ejemplo anterior.

Resolución

En este caso las alternativas válidas son las siguientes:

$$CCC \quad RCC \quad RRC$$

$$N\left(CCC\right) = C_0^7 \times C_3^{12}$$

$$N\left(CCC\right) = \frac{7!}{0! \times 7!} \times \frac{12 \times 11 \times 10 \times 9!}{3!\left(12-3\right)!} = 220$$

$$N\left(RCC\right) = C_1^7 \times C_2^{12}$$

$$N\left(RCC\right) = \frac{7!}{1!\left(7-1\right)!} \times \frac{12!}{2!\left(12-2\right)!} = 462$$

$$N\left(RRC\right) = C_2^7 \times C_1^{12}$$

$$N\left(RRC\right) = \frac{7!}{2!\left(7-2\right)!} \times \frac{12!}{1!\left(12-1\right)!} = 252$$

El número total de alternativas es:

$$N\left(E\right) = 220 + 462 + 252 = 934$$

Por lo tanto, la probabilidad buscada es:

$$p\left(E\right) = \frac{934}{969} = 0.964$$

Ejemplo 32. Escriba un código para calcular la probabilidad frecuentista del ejemplo 30.

Resolución

Algoritmo 1.20 El problema de las urnas

```
import numpy as np, random as rd
urna = {}
for i in range(19):
    if i < 8:
        urna[i] = 'blanca'
    elif i > 7 and i < 12:
        urna[i] = 'negra'
    else:
            urna[i] = 'roja'
N, N_E = 100000, 0
for i in range(N):
    lst = []
    for i in range(3):
        lst.append(urna[rd.randint(0, 18)])
    lst = np.array(lst)
    blanca = sum(lst == 'blanca')
    negra = sum(lst == 'negra')
    roja = sum(lst == 'roja')
    if roja > 0:
        N_E += 1
print('La probabilidad de E es', N_E / N)
```

El resultado de la ejecución de este código es:

```
La probabilidad de E es 0.74709
```

Análisis del código

- urna = {} - crea una lista vacía para almacenar los elementos de la urna.
- el primer ciclo *for* - llena la urna con las bolas de distintos colores.
- N = 100000 - establece el número de simulaciones a realizar.
- N_E = 0 - inicia el contador de ocurrencias del evento E.
- segundo ciclo *for* - cuenta el número de ocurrencias de E.
- lst.append(urna[rd.randint(0, 18)]) - realiza la selección aleatoria de las bolas de la urna.

Permutación y combinación con repetición

En los ejemplos anteriores se analizó el caso en que los elementos del conjunto a ordenar son distintos. ¿que pasaría con las permutaciones y combinaciones en el caso de que se repitan los elementos?

Ejemplo 33. Escriba un código para investigar si las permutaciones del conjunto $A = \{'a','a','a','b','b'\}$ se repiten y en caso de hacerlo indique cuántas veces.

Resolución

Algoritmo 1.21 Permutación con elementos repetidos

```
import itertools as it
from collections import defaultdict
i=0
E = ['a','a','a','b','b']
ponderado = defaultdict(int)
for permutacion in...
 it.permutations(E):
    ponderado[permutacion] += 1
print(f"Ocurrencia de las permutaciones:")
for permutacion, count in...
 ponderado.items():
    print(f"La permutacion {permutacion}...
 ocurre {count} veces")
    i +=1
print(f"Se repiten {i} permutaciones")
```

El resultado de la ejecución de este código es:

```
Ocurrencia de las permutaciones:
('a','a','a','b','b') ocurre 12 veces
('a','a','b','a','b') ocurre 12 veces
('a','a','b','b','a') ocurre 12 veces
('a','b','a','a','b') ocurre 12 veces
('a','b','a','b','a') ocurre 12 veces
('a','b','b','a','a') ocurre 12 veces
('b','a','a','a','b') ocurre 12 veces
('b','a','a','b','a') ocurre 12 veces
('b','a','b','a','a') ocurre 12 veces
('b','b','a','a','a') ocurre 12 veces
Se repiten 10 permutaciones
```

Análisis del código

- collections - es un módulo que proporciona estructuras alternativas de agrupamiento.
- diccionario - es una estructura de agrupamiento que permite expresar de forma concisa los datos. Los elementos del diccionario se presentan en pares. Por ejemplo, $E = \{'a' : 3, 'b' : 2\}$ donde el primer elemeno $'a'$ es la clave y el segundo elemento 3 es el valor.
- defaultdict - es un contenedor del módulo colecciones. A diferencia de los diccionarios defaultdict no genera un mensaje de error si la clave no está presente.

- primer ciclo *for* - crea la permutación de E; la usa como clave para *ponderado* e incrementa el respectivo valor cada vez que ocurre una permutación específica.

- segundo ciclo *for* - cuenta las ocurrencias de las permutaciones.

- items() - se usa para devolver una lista con todas las claves del diccionario con valores.

- count() - devuelve el número de elementos de un valor especificado.

Si algunos de los objetos son idénticos, la permutación y la combinación se transforman en un problema de selección con repetición.

Si el primer elemento del conjunto a ordenar se repite n_1 veces, el segundo n_2 veces, ... y el último se repite n_m veces entonces el número de permutaciones de los elementos es:

$$P^n_{n_1,n_2,\cdots,n_m} = \frac{P^n_n}{n_1! \times n_2! \times \cdots \times n_m!} \tag{1.10}$$

El número de total de elementos es

$$n = \sum n_i$$

Es importante notar que en la ecuación 1.10 se deben tomar todos los elementos del conjunto original.

Ejemplo 34. Determine la cantidad de números de 9 cifras que se pueden formar a partir del conjunto

$$A = \{2, 2, 2, 3, 3, 3, 3, 4, 4\}$$

.

Resolución
El primer elemento se repite 3 veces;
el segundo 4 veces y
el tercero 2 veces.
El número de permutaciones sin repetición es:

$$P^9_{3,4,2} = \frac{9^9_9}{3! \times 4! \times 2!}$$

$$P^9_{3,4,2} = \frac{9 \times 8 \times 7 \times 6 \times 5 \times 4!}{6 \times 4! \times 2} = 1260$$

El número de combinaciones de un conjunto A de n elementos tomados en grupos de k elementos $(k < n)$ cuando se puede repetir los elementos de A, se calcula mediante la expresión:

$$CR^n_k = \frac{(n + k - 1)!}{k!\,(n - 1)!} \tag{1.11}$$

Dos combinaciones con repetición se consideran idénticas si tienen los mismos elementos repetidos el mismo número de veces, independientemente de su orden.

Ejemplo 35. Determine el número de variantes posibles de helados de cono de tres sabores. El heladero dispone de tres sabores de helado: chocolate, limón y vainilla y los clientes pueden repetir los sabores.

Resolución

Este es un problema de combinación con repetición $n = 3$ y $k = 3$.

El número de alternativas es entonces:

$$CR_k^n = \frac{(3+3-1)!}{3!\,(3-1)!}$$

$$CR_k^n = \frac{5 \times 4 \times 3!}{3! \times 2!} = 10$$

Ejemplo 36. Escriba un código para listar las distintas alternativas de seleccionar un cono de helado de tres bolas si hay disponibles tres sabores $\{'chocolate', 'limon', 'vainilla'\}$.

Resolución

Algoritmo 1.22 Listado de alternativas de conos de helado

```
import itertools as it
sabores = ['chocolate','limon','vainilla']
for i in...
 it.combinations_with_replacement(sabores,3):
    print(i)
```

El resultado de la ejecución de este código es:

```
('chocolate', 'chocolate', 'chocolate')
('chocolate', 'chocolate', 'limon')
('chocolate', 'chocolate', 'vainilla')
('chocolate', 'limon', 'limon')
('chocolate', 'limon', 'vainilla')
('chocolate', 'vainilla', 'vainilla')
('limon', 'limon', 'limon')
('limon', 'limon', 'vainilla')
('limon', 'vainilla', 'vainilla')
('vainilla', 'vainilla', 'vainilla')
```

Capítulo 2

Probabilidad condicional y teorema de Bayes

2.1 Operaciones con eventos

El espacio muestral, los eventos y sus relaciones se pueden representar convenientemente mediante los diagramas de Venn. En estos diagramas el espacio muestral Ω está representado por el conjunto de puntos, que se encuentran en el interior de una determinada figura geométrica, por lo general un rectángulo. Los eventos E_i se representan como figuras de menor o igual tamaño que Ω.

Dado los eventos E_1, E_2 y el espacio muestral Ω se pueden realizar las siguientes operaciones:

- la diferencia de dos eventos es el subconjunto del espacio muestral que contiene a los elementos de E_1 que no se encuentran en E_2:

$$E_1 - E_2 = \{x/x \in E_1 \wedge x \notin E_2\}$$

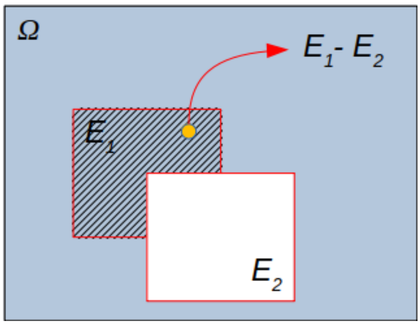


Figura 2.1: Diferencia de eventos

Si $E_1 = \{a, b, c, d\}$ y $E_2 = \{a, e, i, o\}$ entonces
$E_1 - E_2 = \{b, c, d\}$;

- el complemento de un evento es el subconjunto de todos los elementos del espacio muestral Ω, que no pertenecen a E:

$$E^c = \Omega - E = \{x/x \in \Omega \wedge x \notin E\}$$

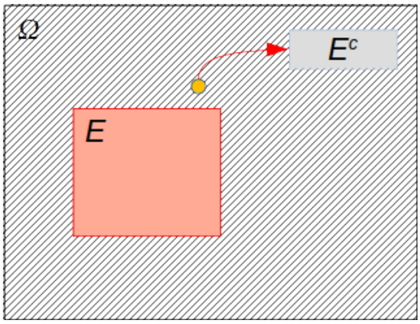


Figura 2.2: Complemento de un evento

Si $\Omega = \{a, e, i, o, b, c, d\}$ y $E = \{a, e, i, o\}$ entonces

$E^c = \{b, c, d\}$;

- la unión de dos eventos es el subconjunto del espacio muestral que contiene a los elementos de E_1 y E_2, pero sin repetir los elementos:

$E_1 \cup E_2 = \{x/x \in E_1 \vee E_2\}$

Si $E_1 = \{a, b, c, d\}$ y $E_2 = \{a, e, i, o\}$ entonces

$E_1 \cup E_2 = \{a, b, c, d, e, i, o\}$;

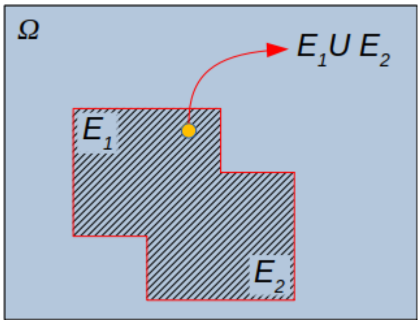


Figura 2.3: Unión de dos eventos

- la intersección de dos eventos es el subconjunto del espacio muestral que contiene a aquellos elementos que pertenecen tanto a E_1 y a E_2:

$$E_1 \cap E_2 = \{x/x \in E_1 \wedge E_2\}$$

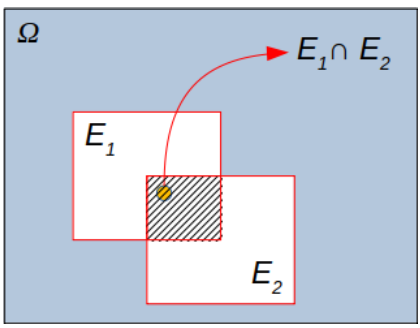


Figura 2.4: Intersección de dos eventos

Si $E_1 = \{a, b, c, d\}$ y $E_2 = \{a, e, i, o\}$ entonces

$$E_1 \cap E_2 = \{a\};$$

Las siguientes definiciones son aplicables a los eventos:

- E_1 y E_2 son mutuamente excluyentes si no pueden ocurrir al mismo tiempo. Si uno de los eventos ocurre, el otro no puede ocurrir, y viceversa. En el lanzamiento de una moneda solo puede salir cara o cruz, pero no ambas;

- Un conjunto de eventos son colectivamente exhaustivo si su unión da como resultado el espacio muestral, $E_1 \cup E_2 \cup \cdots \cup E_n = \Omega$.

- E_1 y E_2 son independientes si al ocurrir uno no afecta la ocurrencia del otro.

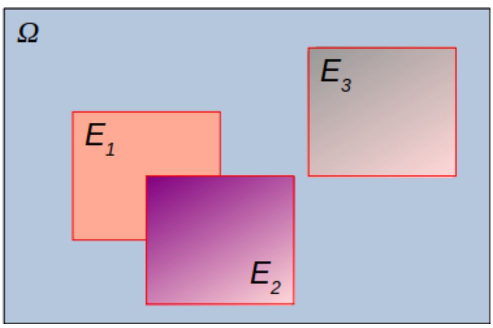


Figura 2.5: E_1 y E_3 son mutuamente excluyentes, pero no E_1 y E_2.

Ejemplo 37. Escriba un código para enumerar los elementos del complemento de $E = \{\text{los múltiplos de 3}\}$ si el espacio muestral son los números entre 1 y 17.

Resolución

Algoritmo 2.1 El complemento de un evento

```
e_muestral= set([x for x in range(1,18)])
E = set([x for x in e_muestral if x%3==0])
print(f"El espacio muestral es:...
 {e_muestral}\n El evento E es: {E}\n...
 El complemento de E es:{ e_muestral-E}")
```

El resultado de la ejecución de este código es:

```
El espacio muestral es:
{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12,
13, 14, 15, 16, 17}
El evento E es:
{3, 6, 9, 12, 15}
El complemento de E es:
{1, 2, 4, 5, 7, 8, 10, 11, 13, 14, 16, 17}
```

Ejemplo 38. Escriba un código para enumerar los elementos de la unión de los eventos $E_1 = \{\text{los múltiplos de 3}\}$ y $E_2 = \{\text{los múltiplos de 2}\}$ si el espacio muestral son los números entre 1 y 10.

Resolución

Algoritmo 2.2 Unión de eventos

```
E_1 = set([x for x in range(1,11)...
  if x%3==0])
E_2 = set([x for x in range(1,11)...
  if x%2==0])
print(f'E_1 = {E_1}')
print(f'E_2 = {E_2}')
print(f'E_1 U E_2 = {E_1|E_2 }')
```

El resultado de la ejecución de este código es:

```
E_1 = {9, 3, 6}
E_2 = {2, 4, 6, 8, 10}
E_1 U E_2 = {2, 3, 4, 6, 8, 9, 10}
```

Ejemplo 39. Escriba un código para enumerar los elementos de la intersección de los eventos $E_1 = \{\text{los múltiplos de } 3\}$ y $E_2 = \{\text{los múltiplos de } 2\}$ si el espacio muestral son los números entre 1 y 10.

Resolución

Algoritmo 2.3 Intersección de eventos

```
E_1 = set([x for x in range(1,11)...
 if x%3==0])
E_2 = set([x for x in range(1,11)...
 if x%2==0])
print(f'E_1 = {E_1}')
print(f'E_2 = {E_2}')
print(f'La interseccion de E_1 y E_2 =...
 {E_1&E_2 }')
```

El resultado de la ejecución de este código es:

```
E_1 = {9, 3, 6}
E_2 = {2, 4, 6, 8, 10}
La interseccion de E_1 y E_2 = {6}
```

Propiedades de las operaciones con eventos.

Propiedad conmutativa

- $E_1 \cup E_2 = E_2 \cup E_1$
- $E_1 \cap E_2 = E_2 \cap E_1$

Propiedad asociativa

- $(E_1 \cup E_2) \cup E_3 = E_1 \cup (E_2 \cup E_3)$
- $(E_1 \cap E_2) \cap E_3 = E_1 \cap (E_2 \cap E_3)$

Propiedad distributiva

- $E_1 \cap (E_2 \cup E_3) = (E_1 \cap E_2) \cup (E_1 \cap E_3)$
- $E_1 \cup (E_2 \cap E_3) = (E_1 \cup E_2) \cap (E_1 \cup E_3)$

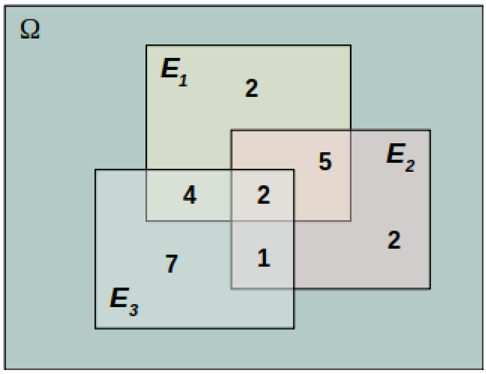
Propiedades de la diferencia de conjuntos

- $E_1 - E_2 = E_1 \cap E_2^c$
- $(E_1 \cap E_2) - E_3 = (E_1 - E_3) \cap (E_2 - E_3)$
- $E_1 \cap (E_2 - E_3) = (E_1 \cap E_2) - (E_1 \cap E_3)$

Las leyes de Morgan

- $(E_1 \cup E_2)^c = E_1^c \cap E_2^c$
- $(E_1 \cap E_2)^c = E_1^c \cup E_2^c$

Ejemplo 40. Un conjunto de personas son simpatizantes de los equipos A, B y C de la localidad. A partir del siguiente diagrama de Venn determine lo siguiente: el número total de aficionados, el número de simpatizantes de cada equipo, el número de personas que siguen a los equipos A y B, B y C; y el número de personas que siguen a A y B pero no a C. Se denota con E_1, E_2 y E_3 a los equipos A, B y C respectivamente.



Resolución

El número total de aficionados es la suma de las personas que muestra el diagrama en cada uno de los subconjuntos:

$$N(\Omega) = 2 + 4 + 2 + 5 + 7 + 1 + 2 = 23$$

El número de aficionados del equipo A es:

$$N(E_1) = 2 + 4 + 2 + 5 = 13$$

El número de aficionados del equipo B es:

$$N(E_2) = 2 + 5 + 1 + 2 = 10$$

El número de aficionados del equipo C es:

$$N\left(E_3\right) = 4 + 2 + 7 + 1 = 14$$

El número de personas que siguen a los equipos A y B es:

$$N\left(E_1 \cap E_2\right) = 5 + 2 = 7$$

El número de personas que siguen a los equipos B y C es:

$$N\left(E_2 \cap E_3\right) = 2 + 1 = 3$$

El número de personas que siguen a los equipos A y B pero no a C es:

$$N\left(\left(E_1 \cap E_2\right) - E_3\right) = 7 - 2 = 5$$

En el ejemplo anterior se observa que si dos eventos E_1 y E_2 no son mutuamente excluyentes la cardinalidad de su unión es:

$$N(E_1 \cup E_2) = N(E_1) + N(E_2) - N(E_1 \cap E_2) \qquad (2.1)$$

Ejemplo 41. Diecisiete jóvenes practican deportes después de sus clases. Se sabe que 8 de ellos juegan voleibol, 3 juegan fútbol y voleibol, 12 juegan fútbol y basketbol, 6 juegan basketbol y voleibol; y 3 juegan los tres deportes. Use un diagrama de Venn para determinar el número de estudiantes que juega cada deporte.

Resolución

Denotemos el hecho de que los estudiantes practican un tipo de deporte de la siguiente manera:

$E_1 = \{\text{juega fútbol}\}$

$E_2 = \{\text{juega voleibol}\}$

$E_3 = \{\text{juega basketbol}\}$

De las condiciones del ejemplo se puede escribir que:

$N(\Omega) = 17$

$N(E_2) = 8$

$N(E_1 \cap E_2) = 3$

$N(E_1 \cap E_3) = 12$

$N(E_2 \cap E_3) = 6$

$N(E_1 \cap E_2 \cap E_3) = 3$

Entonces se puede deducir que

$$N((E_1 \cap E_3) - (E_1 \cap E_2 \cap E_3)) = 9$$

y que

$$N((E_2 \cap E_3) - (E_1 \cap E_2 \cap E_3)) = 3$$

y que

$$N((E_1 \cap E_2) - (E_1 \cap E_2 \cap E_3)) = 0$$

Con esta información se completa el diagrama de Venn

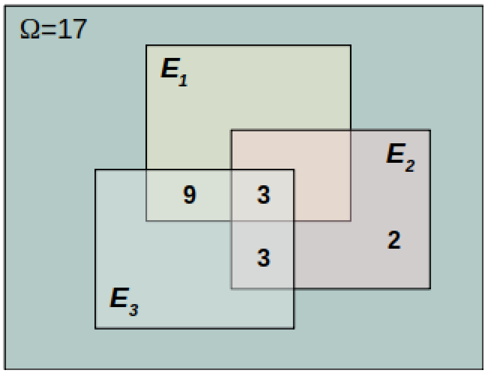


Figura 2.6: Ilustración del ejemplo 41

Entonces:
el número de estudiantes que juegan fútbol es 12;
el número de estudiantes que juegan voleibol es 8 y
el número de estudiantes que juegan basketbol es 15.

Ejemplo 42. Se sabe que de 53 turistas 25 han visitado Cuenca, 35; Baños de Agua Santa y 10 no fueron a ninguna de las dos ciudades. Determine el número de personas que han visitado ambas ciudades.

Resolución

Definamos los eventos
$E_1 = \{\text{han visitado Cuenca}\}$;
$E_2 = \{\text{han visitado Baños de agua santa}\}$;
$E_1 \cup E_2 = \{\text{asistieron a alguna de las dos ciudades}\}$;
$N\left(E_1 \cup E_2\right) = 53 - 10 = 43.$
Los datos los podemos expresar en forma tabular

	E_1	E_1^c	$\sum$
E_2	a	b	c
E_2^c	d	e	f
$\sum$	g	h	i

La casilla a - corresponde al número de turistas que han visitado ambas ciudades: $E_1 \cap E_2$;

La casilla b - los que han visitado Baños, pero no Cuenca: $E_1^c \cap E_2$;

La casilla c - los que han visitado Baños: E_2;

La casilla d - los que han visitado Cuenca, pero no Baños: $E_1 \cap E_2^c$;

La casilla e - los que no han visitado ni Cuenca ni Baños: $E_1^c \cap E_2^c$;

La casilla f - los que no han visitado Baños: E_2^c;

La casilla g - los que han visitado Cuenca: E_1;

La casilla h - los que no han visitado Cuenca: E_1^c y

La casilla i - el número total de turistas.

Después de estas definiciones se puede llenar el cuadro

	E_1	E_1^c	$\sum$
E_2	17	18	35
E_2^c	8	10	18
$\sum$	25	28	53

El número de personas que asistieron a ambas ciudades es 17. Esto se puede corroborar aplicando la ecuación 2.1

$$N(E_1 \cup E_2) = N(E_1) + N(E_2) - N(E_1 \cap E_2)$$

$$43 = 25 + 35 - N(E_1 \cap E_2)$$
$$N(E_1 \cap E_2) = 17$$

2.2 La probabilidad marginal

La probabilidad marginal es la probabilidad de que ocurra un solo evento independientemente de otros eventos. En el ejemplo anterior la probabilidad marginal de que algún turista tomado al azar haya visitado Cuenca es $p(E_1) = \frac{25}{53} = 0.714$.

Ejemplo 43. De un grupo de 7 estudiantes que asisten a tutorías: 5 asisten a clases de física, 4 a química y 2 a ambas. Determine las probabilidades marginales de que un estudiante estudie física y de que estudie química.

Resolución

Definimos los eventos:

$E_1 = \{\text{estudia física}\}$;

$E_2 = \{\text{estudia química}\}$;

Se presentan los datos en forma tabular

	E_1	E_1^c	$\sum$
E_2	2	2	4
E_2^c	3	0	3
$\sum$	5	2	7

Del cuadro se deduce que las probabilidades marginales son:

$$p\,(E_1) = \frac{5}{7} = 0.714$$
$$p\,(E_2) = \frac{4}{7} = 0.571$$

2.3 La probabilidad condicional

La probabilidad condicional es la probabilidad de que ocurra un evento cuando se sabe que ocurre otro evento específico. La notación de la probabilidad condicional es:

$$p\,(E_1|E_2)$$

y se lee la probabilidad del evento 1 dado el evento 2, es decir que ocurre el evento 2.

En el ejemplo anterior la probabilidad de seleccionar un estudiante que estudie física sabiendo que estudia química es:

$$p\left(E_1|E_2\right) = \frac{2}{4} = 0.5$$

debido a que los estudiantes que estudian química son 4 y de entre ellos solo 2 estudian física. Por otro lado la probabilidad de seleccionar un joven que estudie química dado que estudia física es:

$$p\left(E_2|E_1\right) = \frac{2}{5} = 0.4$$

2.4 La probabilidad conjunta

La probabilidad conjunta es la probabilidad de que dos eventos sucedan al mismo tiempo. Matemáticamente se denota por $p\left(E_1 \cap E_2\right)$. En el ejemplo 43 la probabilidad de seleccionar un joven que estudie ambas asignaturas es:

$$p\left(E_1 \cap E_2\right) = \frac{2}{7} = 0.286$$

La probabilidad conjunta de que ocurran dos eventos se puede calcular a partir de las probabilidades condicionales y marginales de la siguiente forma:

$$p\left(E_1 \cap E_2\right) = p\left(E_1|E_2\right) p\left(E_2\right) \tag{2.2}$$

o también como:

$$p\left(E_1 \cap E_2\right) = p\left(E_2 | E_1\right) p\left(E_1\right) \tag{2.3}$$

Ejemplo 44. Se fabrican juguetes en 2 líneas de producción L1 y L2. De 8 juguetes que produce L1 dos son defectuosos, mientras que de 10 que fabrica L2 uno es defectuoso. Determine la probabilidad de que al tomar un juguete defectuoso de una bodega este sea fabricado por L1.

Resolución

Definamos los eventos:

$E_1 = \{\text{se fabrica en L1}\}$;

$E_2 = \{\text{se fabrica en L2}\}$;

$E_3 = \{\text{es defectuoso}\}$.

Supongamos que se fabrican solo 18 juguetes para completar el cuadro de datos.

	E_1	E_2	$\sum$
E_3	2	1	3
E_3^c	6	9	15
$\sum$	8	10	18

En el problema se pide calcular la probabilidad conjunta $p\left(E_1 \cap E_3\right)$.

De la tabla de datos se tiene que:

$$p\left(E_1 \cap E_3\right) = \frac{2}{18} = 0.111$$

De acuerdo a la ecuación 2.2

$$p(E_1 \cap E_3) = p(E_1|E_3)\, p(E_3)$$

La probabilidad marginal $p(E_3) = \frac{3}{18} = 0.167$
La probabilidad condicional $p(E_1|E_3) = \frac{2}{3} = 0.667$
Por lo tanto:

$$p(E_1 \cap E_3) = \frac{2}{3} \times \frac{3}{18} = \frac{2}{18} = 0.111$$

Ejemplo 45. En una heladería se han agotado todos los sabores menos vainilla y limón. De 100 niños que se acercan a comprar: 40 piden vainilla, 25 limón y 10 ambos sabores. Si se elige un niño al azar determine la probabilidad de que haya pedido: limón, vainilla o ambos, y la probabilidad de haya pedido limón dado que ha pedido vainilla y al contrario que haya pedido limón dado que ha pedido vainilla.

Resolución

Definamos los eventos:
$E_1 = \{\text{pide vainilla}\}$;
$E_2 = \{\text{pide limón}\}$.
Los datos los podemos expresar en forma tabular

	E_1	E_1^c	$\sum$
E_2	10	15	25
E_2^c	30	45	75
$\sum$	40	60	100

Probabilidades marginales:

$$p\,(E_1) = \frac{40}{100} = \frac{2}{5} = 0.4$$

$$p\,(E_2) = \frac{25}{100} = \frac{1}{4} = 0.25$$

Probabilidades condicionales:

$$p\,(E_1|E_2) = \frac{10}{25} = \frac{2}{5} = 0.4$$

$$p\,(E_2|E_1) = \frac{10}{40} = \frac{1}{4} = 0.25$$

Probabilidad conjunta:

$$p\,(E_1 \cap E_2) = \frac{10}{100} = \frac{1}{10} = 0.1$$

Ejemplo 46. Considerando el ejemplo anterior escriba un código para determinar la probabilidad frecuentista de que si un niño pide limón el segundo sabor sea vainilla.

Resolución

Algoritmo 2.4 Probabilidad condicional

```
import random as rd
muestra, N_l, N_v, N_vl = 1000, 0, 0, 0
for i in range(muestra):
    op1 = rd.choice(['L', 'V'])
    op2 = rd.choice(['L', 'V'])
    if op1 == 'L' or op2 == 'L':
        N_l += 1
    if (op1 == 'L' and op2 == 'V') or...
     (op1 == 'V' and op2 == 'L'):
        N_vl += 1
p = round(N_vl / N_l, 2)
print(f"La probabilidad de elegir...
 vainilla dado que elige limón es {p}.")
```

El resultado de la ejecución de este código es:

```
La probabilidad de elegir vainilla dado
que elige limón es 0.68.
```

2.5 El teorema de la probabilidad total

Sea que el espacio muestral se puede descomponer en distintos eventos mutuamente excluyentes E_i y además en el espacio muestral está definido un evento A como lo muestra la siguiente figura 2.7 entonces la cardinalidad del evento A está dada por:

$$N\left(A\right) = N\left(E_1 \cap A\right) + N\left(E_2 \cap A\right) + N\left(E_3 \cap A\right)$$

y su respectiva probabilidad esta dada por:

$$p\left(A\right) = \frac{N\left(E_1 \cap A\right) + N\left(E_2 \cap A\right) + N\left(E_3 \cap A\right)}{N\left(\Omega\right)}$$

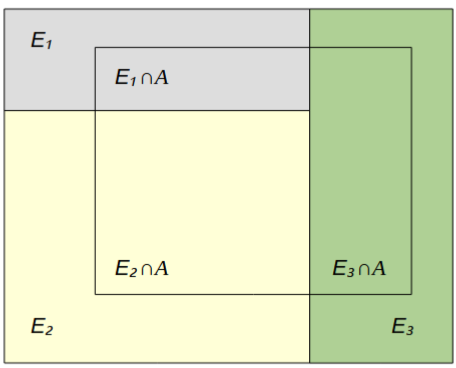


Figura 2.7: Probabilidad total

A la probabilidad de ocurrencia del evento A considerando todos los demás eventos se le denomina probabilidad total y en términos generales se escribe como:

$$p\left(A\right) = \sum_{i=1}^{n} p\left(E_i \cap A\right) \tag{2.4}$$

La probabilidad total relaciona las probabilidades marginales con las probabilidades condicionales.

$$p\left(E_i \cap A\right) = p\left(A|E_i\right) p\left(E_i\right)$$

Por lo tanto,

$$p\left(A\right) = \sum_{i=1}^{n} p\left(A|E_i\right) p\left(E_i\right) \tag{2.5}$$

Ejemplo 47. Se inspecciona tres grupos de puertas y se encuentra que de 380 puertas del grupo A el 4% no cumplen con las medidas nominales, del grupo B el 3% de 270 no cumplen con las medidas nominales y del grupo C el 6% de 350 son defectuosas. Si se elige una puerta al azar. Determine la probabilidad de que esta sea defectuosa.

Resolución

Definamos los eventos:

$E_1 = \{$pertenece al grupo A$\}$;

$E_2 = \{$pertenece al grupo B$\}$;

$E_3 = \{$pertenece al grupo C$\}$;

$F = \{$la puerta es defectuosa$\}$.

El número total de puertas inspeccionadas es:

$$N = 380 + 270 + 350 = 1000$$

Las probabilidades marginales son:

$$p\left(E_1\right) = \frac{380}{1000} = 0.380$$

$$p\left(E_2\right) = \frac{270}{1000} = 0.270$$

$$p\left(E_3\right) = \frac{350}{1000} = 0.350$$

Las probabilidades condicionales son:

$$p\left(F|E_1\right) = 0.040$$

$$p\left(F|E_2\right) = 0.030$$

$$p\left(F|E_3\right) = 0.060$$

De acuerdo al teorema de la probabilidad total la probabilidad de que la puerta seleccionada sea defectuosa es:

$$p\left(F\right) = \sum_{i=1}^{n} p\left(F|E_i\right) p\left(E_i\right)$$

$$p\left(F\right) = 0.040 \times 0.380 + 0.030 \times 0.270 + 0.060 \times 0.350$$

$$p\left(F\right) = 0.044$$

2.6 El teorema de Bayes

Thomas Bayes propuso un método para calcular la probabilidad condicional de un evento E respecto a otro evento F si se conoce las probabilidades de los eventos y la probabilidad condicional de F sabiendo que ha ocurrido E.

De la probabilidad conjunta de dos eventos se tiene:

$$p\left(E \cap F\right) = p\left(E|F\right) p\left(F\right)$$

y

$$p\left(E \cap F\right) = p\left(F|E\right) p\left(E\right)$$

Por lo tanto,

$$p\left(E|F\right) p\left(F\right) = p\left(F|E\right) p\left(E\right)$$

y de aquí se obtiene:

$$p\left(E|F\right) = \frac{p\left(F|E\right) p\left(E\right)}{p\left(F\right)} \tag{2.6}$$

En términos generales, dado un número de eventos que forman un grupo exhaustivo y considerando el teorema de la probabilidad total, el teorema de Bayes se escribe de la siguiente forma:

$$p\left(E_i|F\right) = \frac{p\left(F|E_i\right) p\left(E_i\right)}{\sum_{i=1}^{n} p\left(F|E_i\right) p\left(E_i\right)} \tag{2.7}$$

Ejemplo 48. Una fábrica de papas fritas (chips) compra papas a dos proveedores. Las papas deben tener un tamaño mínimo caso contrario son consideradas de rechazo. En un muestreo se determina que el porcentaje de papas rechazadas del primer proveedor es igual a 20%, mientras que del segundo proveedor es el 10%. Si se sabe que al primer proveedor se le compra cinco veces más que al segundo, determine la probabilidad de que al encontrar una papa de rechazo en la línea de producción esta sea del segundo proveedor.

Resolución

Definamos los eventos:

$E_1 = \{\text{comprada al primer proveedor}\}$;

$E_2 = \{\text{comprada al segundo proveedor}\}$;

$F = \{\text{la papa es de rechazo}\}$.

Las probabilidades marginales se pueden calcular a partir de la cuota de cada proveedor. Si se supone que se compra n kilogramos de papa

$$n = n_1 + n_2$$

n_1 - es la cantidad de kilogramos comprados al primer proveedor;

n_2 - es la cantidad de kilogramos comprados al segundo proveedor;

$$p\left(E_1\right) = \frac{n_1}{n_1 + \frac{1}{5}n_1} = \frac{5}{6} = 0.833$$

$$p\left(E_2\right) = 1 - p\left(E_1\right) = \frac{1}{6} = 0.167$$

Las probabilidades condicionales de seleccionar una papa de rechazo de un proveedor determinado son:

$$p(F|E_1) = 0.2$$

$$p(F|E_2) = 0.1$$

La probabilidad total de seleccionar una papa de rechazo es:

$$p\left(F\right) = \sum_{i=1}^{n} p\left(F|E_i\right) p\left(E_i\right)$$

$$p\left(F\right) = 0.833 \times 0.2 + 0.167 \times 0.1 = 0.183$$

De acuerdo al teorema de Bayes:

$$p\left(E_2|F\right) = \frac{p(F|E_2)p(E_2)}{p\left(F\right)}$$

$$p\left(E_2|F\right) = \frac{0.1 \times 0.167}{0.183} = 0.091$$

El teorema de Bayes generalmente se interpreta en términos de causa - efecto. Los eventos E_i se consideran las causas y el evento F; el efecto. Un ejemplo de causa - efecto puede ser el hecho de que un auto no enciende. Se puede señalar las siguientes causas:

- la batería está agotada;
- se agotó el combustible;
- hay un desperfecto eléctrico, entre otros.

Se intenta encender el auto y este no enciende. ¿Cuál es la probabilidad de que el auto no encienda debido a que la batería está agotada? El teorema de Bayes permite evaluar la probabilidad de que la batería sea la causa por la cuál no enciende el auto.

El teorema de Bayes también se interpreta en términos de hipótesis. En este caso los eventos E_i son las hipótesis. Por ejemplo, si se considera una enfermedad que afecta a una población y se toma una persona al azar se puede considerar la hipótesis de que esté enfermo o su complemento de que se encuentre sano. En ausencia de evidencias esta probabilidad se toma en función de la prevalencia[1]. Es decir en base a las estadísticas de la población.

La determinación de si una persona está enferma o no se afecta drásticamente con la aportación de evidencias, por ejemplo, la realización de una prueba de laboratorio. Sin embargo, las pruebas de laboratorio no son perfectas[2] y a pesar de que la prueba dé positivo no se tiene la certeza de que la persona esté enferma.

[1]La prevalencia de una enfermedad es el número de personas en un grupo determinado que tuvieron la enfermedad en un momento específico.

[2]Las pruebas pueden dar falsos positivos y falsos negativos.

La información que se tiene antes de realizar la prueba se llama 'a priori'. Como resultado de la prueba se decide si la persona seleccionada está enferma con una probabilidad 'a posteriori'. Obviamente esta probabilidad está mejorada dada la aportación de evidencia (la prueba). Con el objetivo de incrementar la certeza en la hipótesis se puede realizar otra prueba. En este caso la probabilidad 'a posteriori' pasa a ser 'a priori'.

La probabilidad de que la persona esté enferma dado que la prueba da positivo de acuerdo al teorema de Bayes es:

$$p\left(E|F\right) = \frac{p\left(F|E\right)p\left(E\right)}{p\left(F\right)}$$

- $p\left(E|F\right)$ es la probabilidad 'a posteriori';
- $p(E)$ es la probabilidad 'a priori';
- $p\left(F\right)$ es la probabilidad total de que la evidencia sea positiva y
- $p\left(F|E\right)$ es la probabilidad de que la prueba de un verdadero positivo[1]

Ejemplo 49. Una enfermedad afecta al 1% de la población. Las pruebas identifican correctamente al 90% de los casos

[1]A la probabilidad de dar un verdadero postivo se le llama sensibilidad de la prueba y a la probabilidad de dar un verdadero negativo se la llama especificidad.

positivos e identifican incorrectamente al 5% de los negativos. Se realizan varios test a una persona. ¿Cuál es la certeza de que esta persona esté enferma después de cada test si estos dan positivo?

Resolución

De acuerdo al enunciado la prevalencia de la enfermedad es del 1%.

La sensibilidad es igual a 0.9 y

la especificidad es igual a 0.95.

Antes de realizar el primer test la certeza de que una persona esté enferma es igual a la prevalencia.

Para calcular la probabilidad después del primer test definamos los eventos:

$E = \{$está enfermo$\}$;

$F = \{$el test da positivo$\}$.

Supongamos que se observan 10000 personas para completar el cuadro de datos.

El número de afectados entre la población es:

$$N\left(E\right) = 0.01 \times 10000 = 100$$

El número de casos positivos detectados correctamente por los test es:

$$N\left(E \cap F\right) = 0.9 \times 100 = 90$$

El número de casos negativos detectados incorrectamente[1] por los test es:

$$N\left(E^{c}\cap F\right)=0.05\times 9900=495$$

	E	E^c	$\sum$
F	90	495	585
F^c	10	9405	9415
$\sum$	100	9900	10000

A partir del cuadro la certeza de estar enfermo si el test da positivo es:

$$p\left(E|F\right)=\frac{90}{585}=0.154$$

Para calcular la probabilidad de estar enfermo después del segundo test se debe considerar que el primer test dió positivo y por lo tanto la población analizada[2] es ahora $N\left(F\right)$ =585.

	E	E^c	$\sum$
F	81	24.75	105.75
F^c	9	470.25	479.25
$\sum$	90	495	585

[1] falso positivo

[2] Para la explicación se ha utilizado decimales a pesar de que el número de personas sea una cantidad discreta.

A partir del cuadro la certeza de estar enfermo si el segundo test da positivo es:

$$p\,(E|F) = \frac{81}{105.75} = 0.766$$

Después del tercer test la población es $N\,(F)$ =105.75

	E	E^c	$\sum$
F	72.9	1.24	74.14
F^c	8.1	23.51	31.61
$\sum$	81	24.75	105.75

A partir del cuadro la certeza de estar enfermo si el tercer test da positivo es:

$$p\,(E|F) = \frac{72.9}{74.14} = 0.983$$

Ahora verificaremos los cálculos utilizando el teorema de Bayes:

$$p\,(E|F) = \frac{p(F|E)p(E)}{p\,(F)}$$

Para el primer test:

$p\,(F|E) = 0.9$

$p\,(E) = 0.01$

$p\,(F) = p\,(F|E)\,p\,(E) + p\,(F|E^c)\,p\,(E^c)$

$p\,(F) = 0.9 \times 0.01 + 0.05 \times (1 - 0.01) = 0.0585$

$$p\left(E|F\right) = \frac{0.9 \times 0.01}{0.0585} = 0.154$$

Para el segundo test:
$p\left(F|E\right) = 0.9$
$p\left(E\right) = 0.154$
$p\left(F\right) = 0.9 \times 0.154 + 0.05 \times (1 - 0.154) = 0.1809$

$$p\left(E|F\right) = \frac{0.9 \times 0.154}{0.1809} = 0.766$$

Para el tercer test:
$p\left(F|E\right) = 0.9$
$p\left(E\right) = 0.766$
$p\left(F\right) = 0.9 \times 0.766 + 0.05 \times (1 - 0.766) = 0.701$

$$p\left(E|F\right) = \frac{0.9 \times 0.766}{0.701} = 0.983$$

Ejemplo 50. Escriba un código en Python para determinar el número de pruebas necesarias para tener una certeza superior al 95% de que una persona se encuentre afectada por una enfermedad si se sabe que la sensibilidad y la especificidad de las pruebas son 0.9 y 0.95 respectivamente y la prevalencia de la enfermedad es del 1%.

Resolución

Algoritmo 2.5 Teorema de Bayes

```
import numpy as np
import matplotlib.pyplot as plt
prevalencia = 0.01
sensibilidad = 0.9
especificidad = 0.95
p=prevalencia
P=np.array([])
while p<0.95:
    ptotal=sensibilidad*p+...
 (1-especificidad)*(1-p)
    p=sensibilidad*p/ptotal
    P=np.append(P,p)
plt.plot(range(3),P)
plt.xlabel('Número de pruebas',...
 fontsize=14)
plt.ylabel('Certeza de estar enfermo',...
 fontsize=14)
plt.locator_params(axis='x', nbins=3)
plt.grid(axis='y')
plt.show()
```

El resultado de la ejecución de este código es:

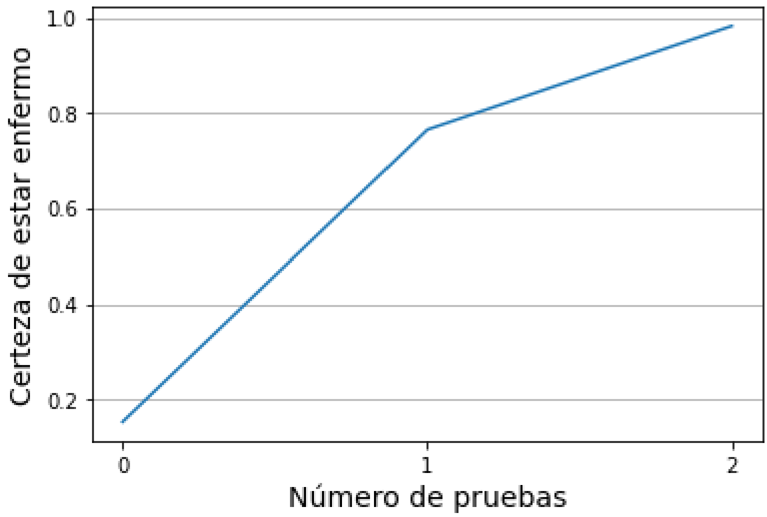


Figura 2.8: Teorema de Bayes

Capítulo 3

Variables aleatorias

3.1 Definición

El concepto de variable aleatoria es uno de los más importantes en la teoría de la probabilidad. Una variable aleatoria es una función que asigna un valor real a cada uno de los resultados de un experimento aleatorio. El dominio de una variable aleatoria es el espacio muestral del experimento y su rango un subconjunto de los números reales. Las variables aleatorias se denotan con letras mayúsculas, por ejemplo X y sus valores con letras minúsculas.

$$X\left(s\right) = x$$

s es un elemento del espacio muestral.

Ejemplo 51. Dado el experimento aleatorio que consiste en lanzar dos veces una moneda establezca la regla de correspondencia de la variable aleatoria ’el número de veces que sale cruz’.

Resolución

Definamos al espacio muestral:

S: sale cara

C: sale cruz

Ω = {CC, CS, SC, SS}

La regla de correspondencia queda de la siguiente manera:

$X(s)$	CC	CS	SC	SS
x	2	1	1	0

Ejemplo 52. Dado el experimento aleatorio que consiste en extraer dos bolas sin reposición de una urna que contiene 4 bolas rojas y 2 negras establezca la regla de correspondencia de la variable aleatoria ’el número de veces que sale una bola roja’.

Resolución

Definamos al espacio muestral:

R: sale una bola roja;

N: sale una bola negra

Ω = {RR, RN, NR, NN}

La regla de correspondencia queda de la siguiente manera:

$X(s)$	RR	RN	NR	NN
x	2	1	1	0

Ejemplo 53. Dado el experimento aleatorio que consiste en lanzar un dado establezca la regla de correspondencia de la variable aleatoria 'el puntaje de salida'.

Resolución

Definamos al espacio muestral:

C$_i$: cara i del dado

$\Omega = \{C_1, C_2, C_3, C_4, C_5, C_6\}$

La regla de correspondencia queda de la siguiente manera:

$X(s)$	C_1	C_2	C_3	C_4	C_5	C_6
x	1	2	3	4	5	6

Ejemplo 54. Dado el experimento aleatorio que consiste en que tres estudiantes, Pedro, Carlos y Juan, se acercan aleatoriamente a retirar sus exámenes de una pila de exámenes ordenados alfabéticamente; establezca la regla de correspondencia de la variable aleatoria 'el número de exámenes retirados correctamente'.

Resolución

Definamos al espacio muestral:

C: Carlos;

J: Juan;

P: Pedro.

El orden de las iniciales indica el orden de llegada.

$\Omega = \{\text{CJP, CPJ, JCP, JPC, PCJ, PJC}\}$

Por ejemplo en el elemento CJP los estudiantes llegaron en orden alfabético, por lo tanto, el número de exámenes retirados correctamente es igual a 3.

La regla de correspondencia queda de la siguiente manera:

$X(s)$	CJP	CPJ	JCP	JPC	PCJ	PJC
x	3	1	1	0	0	1

Relación entre eventos y variables aleatorias

Consideremos el lanzamiento doble de una moneda, ejemplo 51, y la variable aleatoria X: 'el número de veces que sale cruz'. Si x toma el valor 1 esto significa que en los lanzamientos 'cruz' sale una sola vez. A su vez esto corresponde al evento:

$$E = \{\text{'cruz' sale una vez}\}$$

Se puede generalizar que un valor real fijo x corresponde a un evento E_x

$$E_x = \{s|X(s) = x\} = [X = x]$$

Se puede definir algunos eventos en términos de una variable aleatoria para valores fijos de x, a, y b:

- $[X \leq x] = \{s|X(s) \leq x\}$
- $[X > x] = \{s|X(s) > x\}$
- $[a < X < b] = \{s|a < X(s) < b\}$

Ejemplo 55. De una caja que contiene 7 piezas estándar y 3 defectuosas se extrae secuencialmente sin reposición piezas hasta obtener una estándar. Considerando la variable aleatoria 'el número de piezas defectuosas extraídas' dé una interpretación del evento $[X < 2] = \{s|X(s) < 2\}$.

Resolución

La variable aleatoria X puede tomar los valores 0, 1, 2, 3.

- si $X(s) = 0$ el operario extrae una pieza estándar en el primer intento.

- si $X(s) = 1$ el operario extrae una pieza estándar en el segundo intento.

El evento $[X < 2] = \{s|X(s) < 2\}$ significa que el operador debe sacar una pieza estándar en no más de 2 intentos.

Tipos de variable aleatoria

La variable aleatoria puede ser de dos tipos:

1. discreta - toma valores múltiplos de un valor mínimo y que pueden numerarse. Por ejemplo:

   X: "el número de veces que sale 'cara' "

2. continua - puede tomar cualquier valor dentro de un intervalo. Por ejemplo:

   X: "la estatura de un estudiante "

3.2 Distribución de probabilidad

La distribución de probabilidad de una variable aleatoria es la función que establece una conexión entre los posibles valores de la variable aleatoria y las probabilidades correspondientes.

X	x_1	x_2	$\cdots$	x_n
$P(X = x_i)$	p_1	p_2	$\cdots$	p_n

Debido a que la variable aleatoria necesariamente toma uno de los valores x_i, entonces los eventos correspondientes forman un grupo exhaustivo y la suma de las probabilidades de su ocurrencia es igual a uno:

$$\sum_{i=1}^{n} p_i = 1 \tag{3.1}$$

Ejemplo 56. Establezca la función de distribución de probabilidad[1] de la variable aleatoria 'el número de veces que sale cruz' en tres lanzamientos de una moneda.

Resolución

La regla de correspondencia de la variable aleatoria X es la siguiente:

[1]La función de distribución de probabilidad de las variables aleatorias discretas también se conoce como función de masa de probabilidad. En el caso de las variables aleatorias continuas la distribución de probabilidad se caracteriza por la función de densidad de probabilidad.

$X(s)$	x
X (CCC)	0
X (CCS)	1
X (CSC)	1
X (CSS)	2
X (SCC)	1
X (SCS)	2
X (SSC)	2
X (SSS)	3

De la regla de correspondencia se puede ver que x toma valores entre 0 y 3. Por lo tanto, la función de distribución de probabilidad es:

x_i	0	1	2	3
p_i	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

x_i	0	1	2	3
p_i	0.125	0.375	0.375	0.125

Ejemplo 57. Un lote de 20 computadoras similares contiene 3 que están defectuosas. Establezca la función de distribución de probabilidad de la variable aleatoria 'el número de computadoras defectuosas' si alguien selecciona al azar 2 de estas computadoras.

Resolución

La variable aleatoria puede tomar valores de 0 a 2 que corresponden a los eventos:

- $[X = 0] = \{$ninguna computadora es defectuosa$\}$;
- $[X = 1] = \{$una computadora es defectuosa$\}$ y
- $[X = 2] = \{$dos computadoras son defectuosas$\}$.

Las probabilidades correspondientes de estos eventos se calcula de la siguiente manera:

$$p_1 = \frac{N([X = 0])}{N(\Omega)}$$

La cardinalidad de $[X = 0]$, de acuerdo al principio de la multiplicación, es el producto del número de alternativas de elegir 2 computadoras operativas de 17 y el número de alternativas de no elegir ninguna de tres defectuosas:

$$N([X = 0]) = C_2^{17} \times C_0^3$$

$$N([X = 0]) = \frac{17 \times 16 \times 15!}{2! \times 15!} \times \frac{3!}{0! \times 3!} = 136$$

El tamaño del espacio muestral es el número de alternativas de elegir 2 computadoras de entre 20

$$N(\Omega) = C_2^{20} = \frac{20 \times 19 \times 18!}{2! \times 18!} = 190$$

y por lo tanto,

$$p_1 = \frac{136}{190} = 0.716$$

Las otras probabilidades son:

$$p_2 = \frac{C_1^{17} \times C_1^3}{C_2^{20}} = \frac{17 \times 3}{190} = 0.268$$

$$p_3 = \frac{C_0^{17} \times C_2^3}{C_2^{20}} = \frac{1 \times 3}{190} = 0.016$$

La función de distribución de probabilidad es:

x_i	0	1	2
p_i	0.716	0.268	0.016

Ejemplo 58. Determine la función de distribución de probabilidad de la variable aleatoria 'la suma de los puntos obtenidos' en el lanzamiento de un dado 2 veces.

Resolución

La variable aleatoria X puede tomar los valores del 2 al 12.

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

Del cuadro se cuenta el número de ocurrencias y se establece la función de distribución de probabilidad de X:

x_i	2	3	4	5	6	7	8	9	10	11	12
p_i	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

3.3 Características de la variable aleatoria

La esperanza matemática

La esperanza matemática describe un valor o posición promedio de una variable aleatoria en el eje real. Por ejemplo, si se considera la variable aleatoria X 'la masa de un recién nacido' y que su esperanza matemática es 3 kilogramos, entonces los valores de la variable aleatoria se concentran en torno a este valor alejándose (dispersándose) de él en cierto grado.

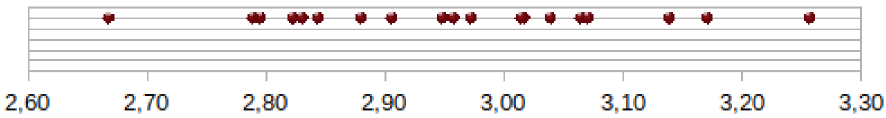


Figura 3.1: Esperanza matemática

La esperanza matemática de X se denota por $E\left[X\right]$ y se calcula de la siguiente manera:

$$E[X] = \sum_{i=1}^{n} x_i p_i \tag{3.2}$$

Ejemplo 59. Calcule la esperanza matemática de la variable aleatoria que está dada por la siguiente función de distribución de probabilidad:

x_1	0	1	2	3	4	5	6
p_i	0.05	0.15	0.15	0.10	0.25	0.20	0.10

<u>Resolución</u>
La esperanza matemática se obtiene directamente:

$$E\left[X\right] = 0 \times 0.05 + 1 \times 0.15 + 2 \times 0.15 + 3 \times 0.10 + 4 \times 0.25 + 5 \times 0.20 + 6 \times 0.10$$

$$E\left[X\right] = 3.35$$

Ejemplo 60. Calcule la esperanza matemática de la variable aleatoria que está dada por la siguiente función de distribución de probabilidad:

x_i	-2	-1	0	1	2
p_i	a	0.30	2a	0.15	0.20

Resolución

Debido a que los eventos $[X = x]$ forman un conjunto exhaustivo la suma de sus probabilidades es igual a 1.

$$\sum p_i = a + 0.30 + 2a + 0.15 + 0.20 = 1$$

y por lo tanto,

$$a = 0.117$$

La esperanza matemática es entonces:

$$E[X] = -2 \times 0.117 - 1 \times 0.30 + 0 \times 0.23 + 1 \times 0.15 + 2 \times 0.20$$

$$E[X] = 0.016$$

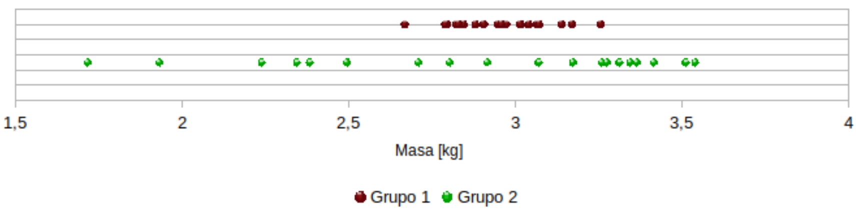
Propiedades de la esperanza matemática

Si C es una constante y X y Y son variables aleatorias discretas independientes se cumplen las siguientes propiedades:

1. $E\left[C\right] = C$;
2. $E\left[C \times X\right] = C \times E\left[X\right]$;
3. $E[X + Y] = E[X] + E[Y]$;
4. $E\left[X \times Y\right] = E\left[X\right] \times M\left[Y\right]$.

Varianza de una variable aleatoria

La varianza caracteriza la dispersión de una variable aleatoria alrededor de su esperanza matemática.



La varianza de una variable aleatoria es la esperanza matemática del cuadrado de la desviación de los valores de X con respecto a su valor medio.

$$Var\left(X\right) = E\left[\left(X - E\left[X\right]\right)^2\right] \tag{3.3}$$

La expresión anterior se la puede expandir tomando en cuenta la linealidad de la esperanza matemática

$$Var\left(X\right) = E\left[X^2 - 2XE\left[X\right] + E^2\left[X\right]\right]$$
$$Var\left(X\right) = E\left[X^2\right] - 2E\left[X\right]E\left[E\left[X\right]\right] + E^2\left[X\right]$$
$$Var\left(X\right) = E\left[X^2\right] - 2E^2\left[X\right] + E^2\left[X\right]$$

Finalmente se obtiene la expresión equivalente:

$$Var\left(X\right) = E\left[X^2\right] - E^2\left[X\right] \tag{3.4}$$

Ejemplo 61. Encuentre la varianza de la variable aleatoria X cuya función de distribución está dada por:

x_i	1	2	3	4
p_i	0.3	0.2	0.4	0.1

Resolución

La varianza de X se la puede obtener de forma tabular mediante la ecuación 3.4

x_i	p_i	x_ip_i	x_i^2	${x_i}^2p_i$
1	0.3	0.3	1	0.3
2	0.2	0.4	4	0.8
3	0.4	1.2	9	3.6
4	0.1	0.4	16	1.6
	$\sum$	2.3	$\sum$	6.3

La varianza de X es entonces:

$$Var\,(X) = 6.3 - (2.3)^2$$

$$Var\,(X) = 1.01$$

Ejemplo 62. Encuentre la esperanza matemática y la varianza de los puntos obtenidos en el lanzamiento de un dado.

Resolución

La función de distribución de probabilidad en este caso es:

x_i	1	2	3	4	5	6
p_i	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

La varianza de X se la puede obtener de forma tabular mediante la ecuación 3.4

x_i	p_i	x_ip_i	x_i^2	${x_i}^2p_i$
1	0.167	0.167	1	0.167
2	0.167	0.333	4	0.667
3	0.167	0.500	9	1.500
4	0.167	0.667	16	2.667
5	0.167	0.833	25	4.167
6	0.167	1.000	36	6
	$\sum$	3.500	$\sum$	15.168

La esperanza matemática se toma directamente del cuadro:

$$E[X] = 3.500$$

La varianza de X es entonces:

$$Var(X) = 15.168 - (3.500)^2$$

$$Var(X) = 2.917$$

Propiedades de la varianza

La varianza posee las siguientes propiedades:

1. $Var(X) \geq 0$;
2. $Var(C) = 0$;
3. $Var(X + C) = Var(X)$;
4. $Var(CX) = C^2 Var(X)$;
5. $Var(X \pm Y) = Var(X) + Var(Y)$.

Ejemplo 63. Demuestre la propiedad:
$Var(CX) = C^2 Var(X)$.
Resolución
Por definición $Var(CX) = E\left[(CX)^2\right] - E^2[CX]$

$$Var\left(CX\right) = E\left[C^2X^2\right] - C^2E^2\left[X\right]$$
$$Var\left(CX\right) = C^2E\left[X^2\right] - C^2E^2\left[X\right]$$
$$Var\left(CX\right) = C^2\left(E\left[X^2\right] - E^2\left[X\right]\right) = C^2Var\left(X\right)$$

La desviación estándar

La desviación típica o desviación estándar es el indicador más común de la dispersión de los valores de una variable aleatoria en relación con su esperanza matemática. La desviación estándar se denota con la letra griega σ y es igual a la raíz cuadrada de la varianza.

$$\sigma = \sqrt{Var\left(X\right)} \tag{3.5}$$

La desviación estándar es una medida de la variabilidad de los valores individuales de un conjunto de datos.

Ejemplo 64. Compare los siguientes proyectos de inversión e indique cuál de ellos es más conveniente.

		Beneficio/probabilidad		
N	1	1000/0.5	0/0.5	
	2	500/0.5	1000/0.25	0/0.25
	3	300/0.6	1500/0.05	700/0.35

Resolución

Los proyectos se pueden comparar tomando como criterio el beneficio esperado:

$$E[X_1] = 1000 \times 0.5 + 0 \times 0.5 = 500$$
$$E[X_2] = 500 \times 0.5 + 1000 \times 0.25 + 0 \times 0.25 = 500$$
$$E[X_3] = 300 \times 0.6 + 1500 \times 0.05 + 700 \times 0.35 = 500$$

En este caso los tres proyectos tienen el mismo beneficio esperado. Por lo que es necesario evaluar su variabilidad (riesgo) mediante la desviación estándar:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} \left(x_i - E[X]\right)^2}{n}}$$
$$\sigma[X_1] = 500$$
$$\sigma[X_2] = 408$$
$$\sigma[X_3] = 600$$

El proyecto 2 tiene el menor riesgo.

La función de distribución acumulada

La función de distribución acumulada caracteriza la probabilidad de que una variable aleatoria X tome un valor menor que x. La función de distribución acumulada determina completamente la variable aleatoria.

$$F(x) = P(X \leq x) = \sum_{y \leq x} p(y) \tag{3.6}$$

Ejemplo 65. Una peonza, cuyos rótulos se indican en la figura, se hace girar dos veces. Establezca la función de distribución de la variable aleatoria 'la suma de los puntos'. Calcule la esperanza matemática y la varianza.

Resolución

Los valores posibles que puede tomar la suma de los puntos son: 2, 3 y 4.

La probabilidad de que salga 1 en un lanzamiento es 0.25 y la probabilidad de que salga 2 es 0.75.

Las probabilidades $P(X = x_i) = p_i$:

$$P(X = 2) = 0.25^2 = 0.063$$

$$P(X = 3) = 0.25 \times 0.75 + 0.75 \times 0.25 = 0.375$$

$$P(X = 4) = 0.75^2 = 0.563$$

La distribución de probabilidad es:

x_i	2	3	4
p_i	0.063	0.375	0.563

La esperanza matemática es:

$$E[X] = 0.063 \times 2 + 0.375 \times 3 + 0.563 \times 4 = 3.500$$

La varianza es:

$$Var[X] = \frac{(2-3.5)^2 + (3-3.5)^2 + (4-3.5)^2}{3} = 0.917$$

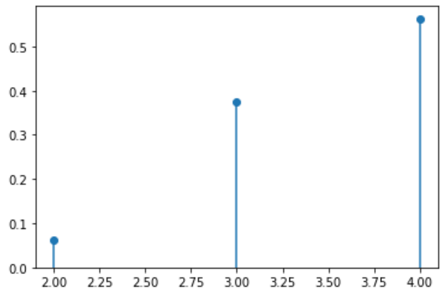
Ejemplo 66. Escriba un código en Python para graficar la función de probabilidad del ejemplo anterior.

Resolución

Algoritmo 3.1 Función de distribución

```
import matplotlib.pyplot as plt
X = [2, 3, 4]
p = [0.063, 0.375, 0.563]
plt.plot( X, p, 'o' )
plt.vlines( X, 0, p )
plt.ylim( bottom=0 )
plt.show()
```

El resultado de la ejecución de este código es:



3.4 Distribuciones de variable discreta

Algunas distribuciones de variable aleatoria discreta son:

- La distribución uniforme;
- la distribución binomial;
- la distribución geométrica;
- la distribución hipergeométrica;
- la distribución de Poisson, entre otras.

La distribución uniforme

Una variable aleatoria tiene una distribución uniforme si toma un número finito de valores con probabilidades iguales. La esperanza matemática de una distribución uniforme es

$$E[X] = \frac{1}{n}\sum_{i=1}^{n} x_i$$

si $x_1 = a$ y $x_n = b$ la esperanza matemática es:

$$E[X] = \frac{a+b}{2} \tag{3.7}$$

La varianza de la distribución uniforme es:

$$Var[X] = \frac{(b-a+1)^2 - 1}{12} \tag{3.8}$$

Ejemplos de procesos que obedecen la distribución uniforme discreta son:

- el lanzamiento de un dado;
- girar una ruleta;
- extraer una carta de una baraja;
- extraer un boleto en una rifa.

Ejemplo 67. Una caja contiene fichas numeradas del 1 al 10. Determine la función de probabilidad y la función acumulativa de probabilidad de la variable aleatoria correspondiente al puntaje obtenido al sacar una ficha. Calcule la esperanza matemática, la desviación estándar y la probabilidad de que el puntaje se encuentre entre 5 y 8.

Resolución

Todas las fichas tienen la misma probabilidad de ser extraídas, por lo tanto esta es una distribución uniforme discreta.

La esperanza matemática es:

$$E\left[X\right] = \frac{10+1}{2} = 5.5$$

La varianza es:

$$Var\left(X\right) = \frac{10^2-1}{12} = 8.25$$

La desviación estándar es:

$$\sigma = \sqrt{8.25} = 2.87$$

Las funciones de probabilidad y probabilidad acumulada las podemos expresar en forma tabular.

X	$P(X)$	$F(X)$
1	0.1	0.1
2	0.1	0.2
3	0.1	0.3
4	0.1	0.4
5	0.1	0.5
6	0.1	0.6
7	0.1	0.7
8	0.1	0.8
9	0.1	0.9
10	0.1	1.0

La probabilidad de que el puntaje se encuentre entre 5 y 8 se puede obtener a partir de la función acumulativa $F(X)$:

$$P(5 < X < 8) = F(8) - F(5)$$

$$P(5 < X < 8) = 0.8 - 0.5 = 0.3$$

Ejemplo 68. Escriba un código para graficar la función de distribución de probabilidad acumulada para el ejemplo anterior.

Algoritmo 3.2 Función de distribución de probabilidad acumulada

```
import matplotlib.pyplot as plt
F_x = 0
for x in range(10):
        plt.hlines(F_x, x, x+1,...
         color='red')
        plt.vlines(x+1, F_x, F_x+0.1,...
         color='blue', linestyles='dashed')
        F_x = F_x + 0.1
        plt.plot(x+1, F_x-0.1, 'D',...
         color='black')
plt.hlines(F_x, 10, 11, color='red')
plt.xlabel('X', fontsize=14)
plt.ylabel('F(X)', fontsize=14)
plt.grid()
plt.show()
```

El resultado de la ejecución de este código es:

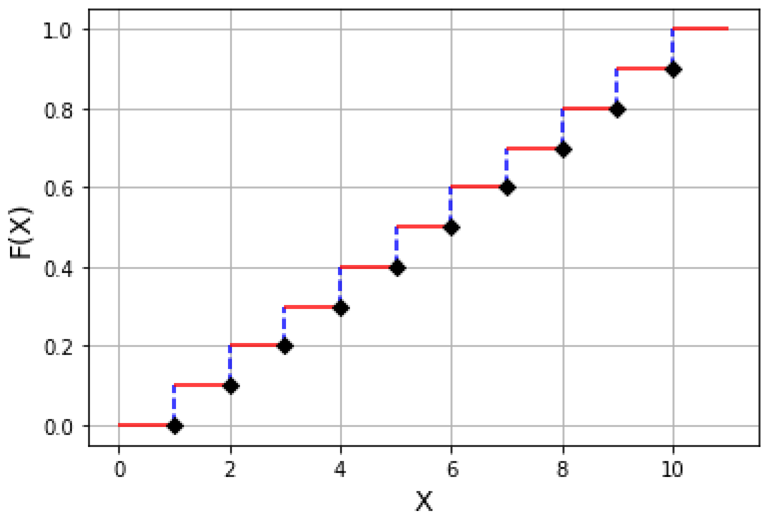


Figura 3.2: Función de distribución de probabilidad acumulada

---

La distribución binomial

El experimento de lanzar una moneda al aire tres veces y el experimento de disparar a un blanco en series de tres tiros tienen varias características en común:

1. se realizan varios intentos o ensayos;

2. solo existen dos posibles resultados 'acertar' o 'fallar';

3. la probabilidad de 'acertar' o de 'exito' no cambia en cada ensayo;

A este tipo de experimentos se les denomina ensayos de Bernoulli o experimento binomial. La variable aleatoria es 'el número de éxitos que se tiene en n ensayos'. El término éxito es relativo, por ejemplo, si en el experimento se registra las veces que sale 'cara' cuando salga 'cruz' se considera un fracaso. Por el contrario si se registra las veces que sale 'cruz' cuando sale 'cara' se considera un fracaso.

La probabilidad $P_n(x)$ de que un evento A ocurra exactamente x veces en n ensayos, cuando el evento A tiene una probabilidad de éxito p y de fracaso $q = 1–p$, se puede encontrar mediante el teorema de Bernoulli:

$$P_n(x) = C_x^n p^x q^{n-x} \tag{3.9}$$

Ejemplo 69. Encuentre la probabilidad de que cruz salga 3 veces en cinco lanzamientos de una moneda.

Resolución

La probabilidad de éxito es $p = 0.5$ y
la probabilidad de fracaso es $q = 1 - p = 0.5$.

La probabilidad buscada se puede encontrar mediante la aplicación de la ecuación 3.9:

$$P_5(3) = C_3^5 \, (0.5)^3 \, (0.5)^{5-3}$$

$$P_5(3) = \frac{5 \times 4 \times 3!}{3!\,(5-3)!}\,(0.5)^3\,(0.5)^{5-3}$$
$$P_5(3) = 10\,(0.5)^5 = 0.313$$

Ejemplo 70. Se lanza una moneda diez veces. ¿Cuál es la probabilidad de que salga 'cara' exactamente 8 veces?

Resolución

La probabilidad buscada es:

$$P_{10}(8) = C_8^{10}\,(0.5)^8\,(0.5)^{10-8}$$
$$P_{10}(8) = \frac{10 \times 9 \times 8!}{8!\,(10-8)!}\,(0.5)^{10}$$
$$P_{10}(8) = 0.044$$

Ejemplo 71. De una urna que contiene 2 bolas blancas y 6 negras, se elige una bola al azar y se devuelve 5 veces seguidas. Calcule la probabilidad de que aparezca una bola blanca 4 veces.

Resolución

Como la bola se devuelve a la urna la probabilidad de sacar una bola u otra no se ve afectada y los ensayos se pueden considerar independientes.

Se debe considerar como éxito sacar una bola blanca y como fracaso sacar una bola negra. Entonces:

$p = 0.333$ y
$q = 0.666$.
El número de ensayos es 5.
La probabilidad buscada es:

$$P_5(4) = C_4^5\,(0.333)^4\,(0.666)^{5-4}$$

$$P_5(4) = \frac{5 \times 4!}{4!\,(5-4)!}\,(0.333)^4\,(0.666)^1$$

$$P_5(4) = 0.041$$

Los coeficientes binomiales forman las filas del triángulo de Pascal:

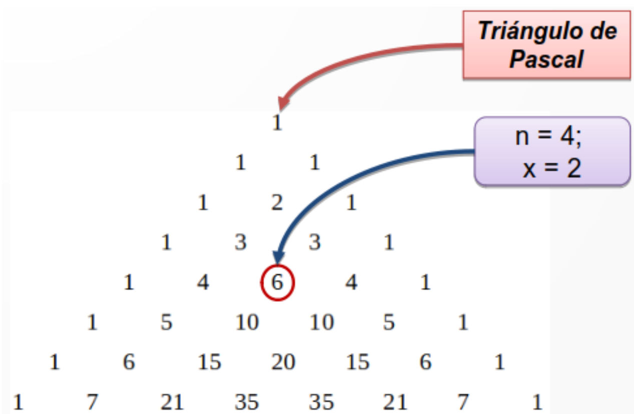


Figura 3.3: Selección del coeficiente binomial para 4 ensayos y 2 éxitos.

Ejemplo 72. Se lanza un dado 10 veces. ¿Cuál es la probabilidad de que el 3 salga no más de 4 veces?

Resolución

Como éxito se toma que sale el número 3 y como fracaso cualquier otro número. Entonces:

$p = \frac{1}{6} = 0.167$ y $q = \frac{5}{6} = 0.833$

Cualquiera de las siguientes alternativas satisface la condición dada:

$P_{10}(0)$, $P_{10}(1)$, $P_{10}(2)$, $P_{10}(3)$, $P_{10}(4)$.

Por el principio de adición la probabilidad buscada es:

$$P_{10}(X \leq 4) = P_{10}(0) + P_{10}(1) + P_{10}(2) + P_{10}(3) + P_{10}(4)$$

Algoritmo 3.3 Aplicación de la distribución binomial

```
from scipy.stats import binom
P = 0
for i in range(1):
        P = P + binom.pmf(k=i, n=10, p=1/6)
print(f'La probabilidad de que el tres...
 salga menos de 4 veces es {P:.4f}')
```

El resultado de la ejecución de este código es:

```
La probabilidad de que el tres salga...
 menos de 4 veces es 0.1615
```

Análisis del código

- scipy.stats - módulo de Python que contiene funciones estadísticas.
- binom - objeto de Python que tiene los atributos de una variable aleatoria binomial.
- binom.pmf - función que corresponde a la función de probabilidad binomial.

Ejemplo 73. Escriba un código para graficar la probabilidad de que salga cara en 10 lanzamientos de una moneda.

Resolución

Algoritmo 3.4 Probabilidad de salir 'cara' en 10 lanzamientos

```
from scipy.stats import binom
import numpy as np
import matplotlib.pyplot as plt
X = np.array(range(11)); P = []
for x in X:
    P = np.append(P, binom.pmf(k=x,...
      n=10, p=1/2))
plt.xlabel('X', fontsize=14)
plt.ylabel('P(X=x)', fontsize=14)
plt.bar(X,P);plt.grid(axis='y');plt.show()
```

El resultado de la ejecución de este código es:

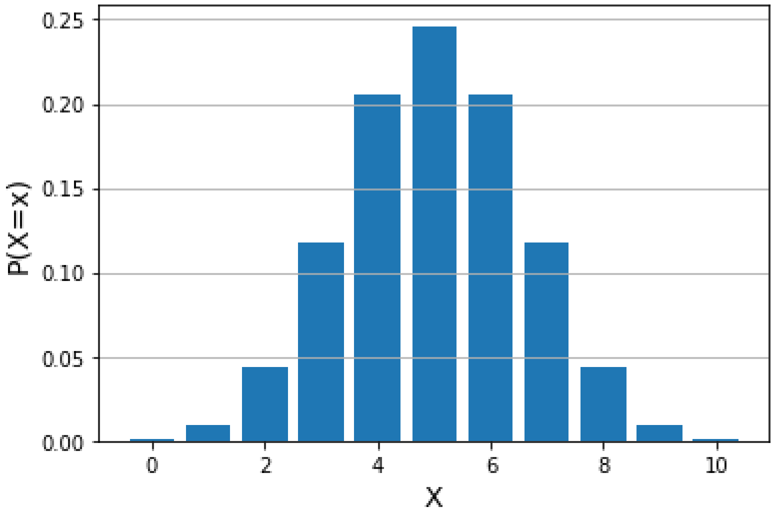


Figura 3.4: Número de veces que sale 'cara' en 10 lanzamientos

La esperanza matemática y la varianza de la distribución binomial se calcula de la siguiente manera:

$$E\left[X\right] = np \tag{3.10}$$

$$Var\left(X\right) = np\left(1 - p\right) \tag{3.11}$$

Ejemplo 74. Se compra un lote de cajas de tornillos a una fabrica. Se inspeccionan 10 cajas de 100 tornillos y se encuentra el siguiente número de tornillos defectuosos.

Caja	1	2	3	4	5	6	7	8	9	10
Defectuosos	3	7	0	0	3	3	2	5	5	1

Determine la probabilidad de encontrar no más de 5 tornillos defectuosos en una caja tomada al azar.

Resolución

Se puede considerar como éxito encontrar un tornillo defectuoso y la inspección de las cajas como ensayos independientes de seleccionar tornillos defectuosos de 100 tornillos. Por lo tanto, se puede encontrar la probabilidad de éxito a partir de la esperanza matemática

$$E[X] = np$$

$$E[X] = \frac{3+7+0+0+3+3+2+5+5+1}{10} = 2.9$$

$$p = \frac{2.9}{100} = 0.029$$

La probabilidad buscada es:

$$P = P_{100}(0) + P_{100}(1) + P_{100}(2) + P_{100}(3) + P_{100}(4) + P_{100}(5)$$

Esta probabilidad se puede calcular con el algoritmo del ejemplo 72

```
La probabilidad de encontrar menos...
 de 5 tornillos defectuosos es 0.9287
```

La distribución geométrica

Una variable aleatoria discreta tiene una distribución geométrica cuando se cuenta el número de fracasos antes de observar el primer éxito en un experimento aleatorio. Un ejemplo es contar 'el número de lanzamientos antes de que salga 'cara' ('sello') en el lanzamiento de una moneda. Otros ejemplos pueden ser:

- el número de disparos antes de acertar en el blanco;
- el número de pruebas de un aparato antes de la primera falla;
- el número de bolas antes de que aparezca una bola blanca.

De manera formal una variable aleatoria discreta X tiene una distribución geométrica con parámetro p, si toma los valores $x = [1, m]$ con probabilidades:

$$P(X = m) = pq^m \tag{3.12}$$

En forma tabular la distribución geométrica se escribe de la siguiente manera:

X	0	1	2	3	$\cdots$	m
p_i	p	pq	pq^2	pq^3	$\cdots$	pq^m

El nombre 'geométrico' refleja el hecho de que a medida que m aumenta, la probabilidad p_m disminuye en una progresión geométrica.

Las características numéricas de la ley geométrica de distribución de probabilidad están dadas por:

- la esperanza matemática

$$E\left[X\right] = \frac{1-p}{p} \tag{3.13}$$

- la varianza

$$Var\left(X\right) = \frac{1-p}{p^2} \tag{3.14}$$

Ejemplo 75. Un jugador de baloncesto novato acierta a 4 de 10 tiros al aro. Determine la probabilidad de fallar 5 tiros antes de encestar.

Resolución

Este es una ejemplo de la aplicación geométrica donde $p = 0.4$, por lo tanto, $q = 0.6$.

La probabilidad que acierte al aro en el sexto lanzamiento corresponde a que ocurren 5 fracasos y por lo tanto $X = 5$:

$$P\left(X = 5\right) = 0.4 \times 0.6^5$$

$$P\left(X = 6\right) = 0.4 \times 0.6^5 = 0.031$$

La probabilidad de que acierte antes del sexto lanzamiento es:

$$P(X < 6) = \sum_{i=0}^{5} P(X = x_i)$$

$$P(X < 6) = 0.4 \times \left(0.6^0 + 0.6^1 + 0.6^2 + 0.6^3 + 0.6^4 + 0.6^5\right) = 0.953$$

$$P(X < 6) = 0.953$$

Ejemplo 76. Según un fabricante de cada 100 dispositivos producidos una falla. Para verificar esta afirmación se realizan pruebas de confiabilidad en las que se registra el número de dispositivos operativos antes de encontrar un dispositivo defectuoso. Determine la experanza matemática, la varianza y la probabilidad de encontrar un dispositivo defectuoso en el quinto ensayo.

Resolución

En este caso como éxito se considera encontrar un equipo defectuoso. Entonces la esperanza matemática es:

$$E[X] = \frac{1 - 0.01}{0.01} = 99$$

En promedio se seleccionan 99 dispositivos antes de encontrar uno defectuoso.

La varianza es:

$$var(X) = \frac{1 - 0.01}{0.01^2} = 9900$$

La probabilidad de encontrar un dispositivo defectuoso después de cuatro operativos es:

$$P = 0.01 \times 0.99^4 \approx 0.01$$

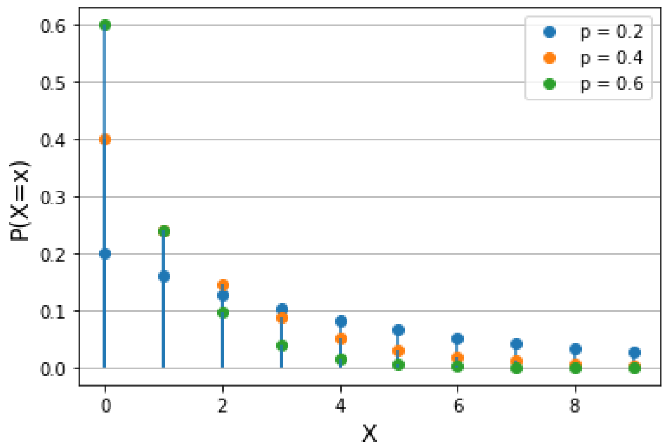
Ejemplo 77. Escriba un código para graficar la función de probabilidad de la distribución geométrica dada la probabilidad de éxito p.

Resolución

Algoritmo 3.5 Distribución geométrica

```
import numpy as np
import matplotlib.pyplot as plt
X = np.array(range(10))
P = [0.2,0.4, 0.6]
for p in P:
    P_g = p*(1-p)**X
    plt.plot(X, P_g, 'o',...
        label = 'p = '+ str(p))
    plt.vlines( X, 0, P_g )
plt.legend()
plt.xlabel('X', fontsize=14)
plt.ylabel('P(X=x)', fontsize=14)
plt.grid(axis='y')
plt.show()
```

El resultado de la ejecución de este código es:



La distribución geométrica es un caso especial de la distribución binomial negativa.

La distribución binomial negativa

En la distribución binomial negativa[1] la variable aleatoria representa el número k de fracasos que ocurren antes de obtener un número dado de éxitos n en una serie de ensayos independientes con una probabilidad de éxito p constante. Los parámetros de esta distribución son el número esperado de éxitos y la probabilidad de éxito. La función de distribución de probabilidad viene dada por:

[1]Esta distribución también se conoce como distribución de Pascal en honor a Blaise Pascal y distribución de Polya , en honor a George Pólya.

$$P(X=k) = f(k,n,p) = C_k^{k+n-1}(1-p)^k p^n \tag{3.15}$$

La esperanza matemática de la distribución binomial negativa es:

$$E[X] = \frac{n(1-p)}{p} \tag{3.16}$$

La varianza:

$$Var(X) = \frac{n(1-p)}{p^2} \tag{3.17}$$

Ejemplo 78. Con respecto a los dispositivos del ejemplo 76 cuál es la probabilidad de encontrar 90 dispositivos que funcionen correctamente antes de encontrar 2 defectuosos.

Resolución

$k = 90$;

$n = 2$;

$p = 0.01$.

La probabilidad buscada es:

$$f(k,n,p) = C_{90}^{90+2-1}(1-0.01)^{90}(0.01)^2$$

$$f(k,n,p) = \frac{91 \times 90!}{90!\,(91-90)!}(1-0.01)^{90}(0.01)^2$$

$$f(k,n,p) = 91 \times (0.99)^{90}(0.01)^2$$

$$f(k,n,p) = 0.004$$

Ejemplo 79. Con respecto al ejemplo 76 determine la probabilidad de que en 200 ensayos se encuentren 2 dispositivos defectuosos.

Resolución

El número de ensayos es la suma de la cantidad de dispositivos defectuosos y operativos $N = k + n$, por lo tanto la probabilidad buscada[1] es:

$$f(N, n, p) = C_{N-n}^{(N-n)+n-1} (1-p)^{N-n} p^n$$
$$f(N, n, p) = C_{N-n}^{N-1} (1-p)^{N-n} p^n$$
$$f(N, n, p) = C_{200-2}^{200-1} (1-0.01)^{200-2} (0.01)^2$$
$$f(N, n, p) = \frac{199 \times 198!}{198!\,(199-198)!} (1-0.01)^{198} (0.01)^2$$
$$f(N, n, p) = 199 \times (0.99)^{198} (0.01)^2$$
$$f(N, n, p) = 0.003$$

La distribución hipergeométrica

En la distribución hipergeométrica se considera una población, compuesta de N elementos de dos tipos A y B, de la cuál

[1] Esta es una definición alternativa de la distribución binomial negativa. La variable aleatoria corresponde al número de ensayos necesarios que se deben realizar para ontener un número fijo de éxitos.

se extrae sin reposición una muestra de n elementos. La distribución hipergeométrica da la probabilidad de que la muestra tenga un cierto número de elementos del tipo A. Como ilustración de la distribución hipergeométrica se puede tomar al problema de las urnas, ejemplo 30. En este contexto la variable aleatoria X es el número de bolas del primer tipo en la muestra.

La probabilidad de obtener X éxitos está dada por:

$$P\left(X = x\right) = \frac{C_x^k C_{n-x}^{N-k}}{C_n^N} \tag{3.18}$$

donde:

N es el tamaño de la población;

n es el tamaño de la muestra extraída;

k es el número de éxitos en la población.

La esperanza matemática en la distribución hipergeométrica

$$E\left[X\right] = \frac{kn}{N} \tag{3.19}$$

La varianza es:

$$D\left(X\right) = \frac{k}{N} \times n \times \frac{N-n}{N} \times \frac{N-k}{N-1} \tag{3.20}$$

Ejemplo 80. En una bodega hay 5 computadoras de 14" y 3 de 12". Elabore la distribución de probabilidad de que en una muestra de 4 computadoras tomadas al azar n sean de 14" y construya su gráfico.

Resolución

En este ejemplo los parámetros son:
el tamaño de la población $N = 8$;
el tamaño de la muestra es $n = 4$ y
el número de éxitos puede tomar los valores $k = \{1, 2, 3, 4\}$.

Las probabilidades respectivas se obtienen a partir de la ecuación 3.18

$$P(X = 1) = \frac{C_1^5 C_3^3}{C_4^8}$$

$$P(X = 1) = 0.071$$

El resto de resultados los expresamos en forma tabular:

X	1	2	3	4
$P(X)$	0.071	0.429	0.429	0.071

El gráfico se puede construir con el siguiente código

Algoritmo 3.6 Probabilidad hipergeométrica

```
import math
import numpy as np
import matplotlib.pyplot as plt
N = 8
K = 5
n = 4
X = [1,2,3,4]
P = []
for x in X:
    N_A = math.comb(K, x)
    N_B = math.comb(N-K, n-x)
    N_T = math.comb(N, n)
    p = N_A* N_B/N_T
    P = np.append(P,p)
plt.bar(X,P)
plt.xticks(range(5))
plt.xlabel('X', fontsize=14)
plt.ylabel('P(X)', fontsize=14)
plt.grid(axis='y')
plt.show()
```

El resultado de la ejecución de este código es:

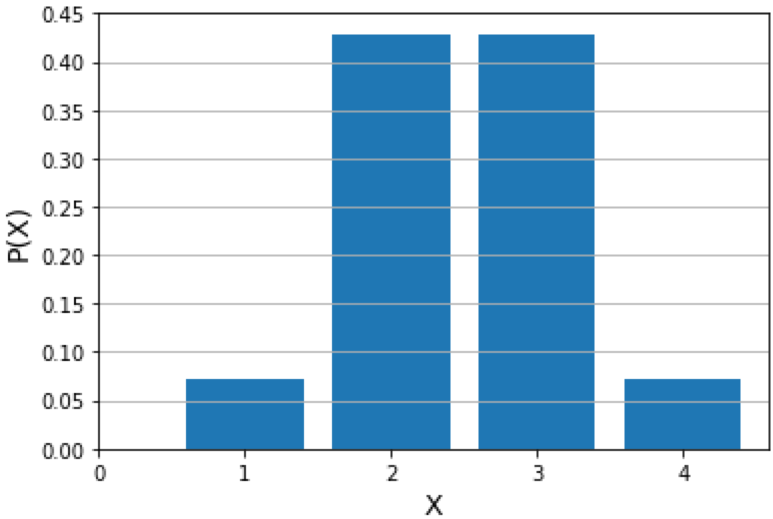


Figura 3.5: Probabilidad hipergeométrica

La distribución de Poisson

La distribución de Poisson describe el número de eventos que ocurren con una frecuencia o intensidad fijas, por ejemplo:

- el número de llamadas recibidas por la central telefónica durante el tiempo t;
- el número de defectos en una pieza de tela por unidad de longitud;

- el número de personas que padecen una enfermedad por cada 100000 habitantes, etc.

Si el número promedio de ocurrencias en un intervalo de tiempo fijo es λ, entonces la probabilidad de que haya exactamente x ocurrencias está dada por:

$$P\left(X=x\right)=\frac{\lambda^{x}e^{-\lambda}}{x!} \tag{3.21}$$

La esperanza matemática y la varianza de X son iguales a λ.

$$E\left[X\right]=Var\left(X\right)=\lambda \tag{3.22}$$

La distribución de Poisson es una aproximación de la distribución binomial en situaciones en las que la probabilidad de éxito es pequeña y n es un número grande. Una buena aproximación se obtiene si $n>50$ y $n\times p<5$.

Ejemplo 81. En un parque automotor se cambian los rodamientos de embrague a 150 vehículos. La probabilidad de falla de un rodamiento dentro de un mes es 0.03. Encuentre la probabilidad de que falle exactamente 1 rodamiento en un mes.

Resolución

El número de rodamientos es mayor que 50 y la probabilidad es menor 0.1. Por lo tanto, se puede considerar que la distribución de probabilidad es una distribución de Poisson.

$$\lambda = 0.03 \times (150) = 4.5$$
$$P\left(X = 1\right) = \frac{4.5^{1}e^{-4.5}}{1!}$$
$$P\left(X = 1\right) = 0.05$$

Ejemplo 82. En una policlínica atienden en promedio a 2 pacientes por hora en el servicio de tomgrafía. Calcule la probabilidad de que se atienda de 6 a 8 personas en 2 horas.

Resolución

El promedio de pacientes atendidos en 2 horas es:

$$\lambda = 2 \times 2 = 4$$

La probabilidad de atender de 6 a 8 pacientes es:

$$P\left(6 \leq x \leq 8\right) = P_6 + P_7 + P_8$$
$$P\left(6 \leq x \leq 8\right) = \frac{4^6}{6!}e^{-6} + \frac{4^7}{7!}e^{-7} + \frac{4^8}{8!}e^{-8}$$

$$P\left(6 \leq x \leq 8\right) = 0.017$$

3.5 Distribuciones de variable continua

En una distribución de probabilidad continua la variable aleatoria X puede tomar un número infinito de valores dentro de un intervalo. En consecuencia la probabilidad de tomar un valor específico es cero. Por lo tanto, la probabilidad se asigna por rangos de valores, por ejemplo, la probabilidad de que X sea mayor que cero:

$$P(X > 0) = 0.50$$

La probabilidad de que una variable aleatoria tome un valor en un intervalo a, b se describe por la densidad de distribución de probabilidad $f(x)$.

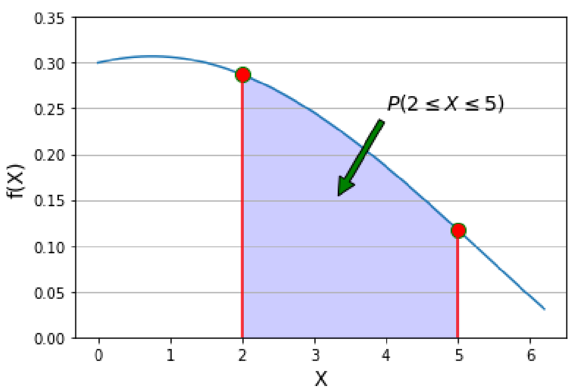


Figura 3.6: Densidad de probabilidad

El área debajo de la densidad de probabilidad es la probabilidad de que la variable aleatoria se encuentre entre a y b.

Las propiedades de la densidad de probabilidad son:

1. $f\left(x\right) \geq 0$
2. $\int_{-\infty}^{\infty} f\left(x\right) = 1$
3. $P\left(a \leq x \leq b\right) = \int_{a}^{b} f\left(x\right) dx = F\left(b\right) - F\left(a\right)$

A partir de la densidad de probabilidad se puede determinar la esperanza matemática de la variable aleatoria:

$$E\left[X\right] = \int_{-\infty}^{\infty} x f\left(x\right) dx \tag{3.23}$$

Las distribuciones de variable continua más importantes son las siguientes:

- Distribución uniforme continua;
- distribución exponencial;
- distribución normal;
- distribución F;
- distribución Gamma;

- distribución ji cuadrado;
- distribución t de Student.

La distribución continua uniforme

La distribución más simple es la distribución continua uniforme en el intervalo $[a, b]$. Su densidad de probabilidad es:

$$f(x) = \begin{cases} \frac{1}{b-a}, & x \in [a, b] \\ 0, & x \notin [a, b] \end{cases}$$

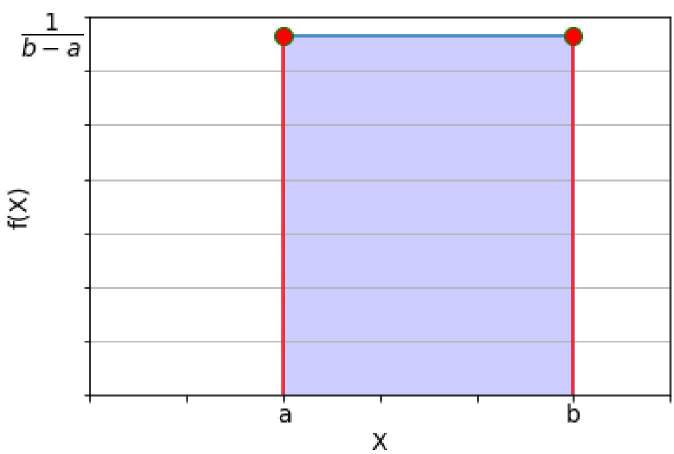


Figura 3.7: Densidad de probabilidad de la distribución uniforme

La esperanza matemática

$$E[X] = \frac{a+b}{2}$$

La varianza

$$Var(X) = \frac{(b-a)^2}{12}$$

Ejemplo 83. Las observaciones mostraron que el peso de la caja, destinado al transporte de hortalizas, se distribuye uniformemente en el rango de 985 a 1025. Se selecciona aleatoriamente una caja. Encuentre el porcentaje de cajas cuyos pesos se alejan de la media a una desviación estándar.

Resolución

La esperanza matemática:

$$E[X] = \frac{985+1025}{2} = 1005$$

La varianza:

$$Var(X) = \frac{(1025-985)^2}{12} = 133.33$$

La desviación estándar:

$$\sigma = \sqrt{133.33} = 11.547$$

La probabilidad de que el peso de la caja se encuentre a una desviación de estándar de la media es el área del rectángulo de ancho igual a 2σ y de alto igual a $\frac{1}{b-a}$, por lo tanto la probabilidad buscada es:

$$P\left(993.453 \leq x \leq 1016.547\right) = \frac{2\left(11.547\right)}{1025 - 985}$$

El porcentaje de cajas que se encuentran en este rango es 57.7%

La distribución exponencial

La distribución exponencial modela el tiempo entre dos ocurrencias sucesivas de un evento caracterizado por el número promedio de ocurrencias de un evento por unidad de tiempo λ. Esta distribución permite describir:

- la duración de la atención a un cliente;
- la vida útil de un aparato hasta su falla;
- el intervalo de tiempo entre averías, etc.

La densidad de probabilidad de la distribución de X viene dada por la fórmula:

$$f(x) = \begin{cases} 0, & x < 0 \\ \lambda e^{-\lambda x}, & x \geq 0 \end{cases} \tag{3.24}$$

La función de distribución tiene la forma:

$$F(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-\lambda x}, & x \geq 0 \end{cases} \tag{3.25}$$

La esperanza matemática de la distribución uniforme es:

$$E[X] = \lambda^{-1} \tag{3.26}$$

La varianza es:

$$Var(X) = \lambda^{-2} \tag{3.27}$$

Ejemplo 84. El tiempo medio de funcionamiento sin fallos de un dispositivo es de 80 horas. Suponiendo que el tiempo de actividad del dispositivo tiene una ley de distribución exponencial, encuentre: a) expresión de su densidad de probabilidad y función de distribución; b) la probabilidad de que dentro de 100 horas el dispositivo no falle.

Resolución

$$E[X] = \lambda^{-1}$$

$$\lambda = \frac{1}{80}$$

La densidad de probabilidad es entonces:

$$f(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{80}e^{-\frac{1}{80}x}, & x \geq 0 \end{cases}$$

La función de distribución tiene la forma:

$$F(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-\frac{1}{80}x}, & x \geq 0 \end{cases}$$

La probabilidad de que dentro de 100 horas el dispositivo no falle, es

$$P(100 \leq x \leq \infty) = F(\infty) - F(100)$$

$$P(100 \leq x \leq \infty) = \left(1 - e^{-\frac{\infty}{80}}\right) - \left(1 - e^{-\frac{100}{80}}\right)$$

$$P(100 \leq x \leq \infty) = e^{-\frac{100}{80}} - e^{-\frac{\infty}{80}}$$

$$P(100 \leq x \leq \infty) = 0.287$$

La distribución normal

La distribución normal es una de las distribuciones de probabilidad más importantes. Distribución de probabilidad de una variable aleatoria real X se llama normal si tiene una densidad de probabilidad igual a

$$p(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-a)^2}{2\sigma^2}} \tag{3.28}$$

La esperanza matemática coincide con a y la varianza con σ^2.

El gráfico de la densidad de probabilidad es simétrico a la línea vertical $x = a$, y alcanza su valor máximo en ese valor. La forma del gráfico depende de σ. Para valores menores el gráfico es mas angosto.

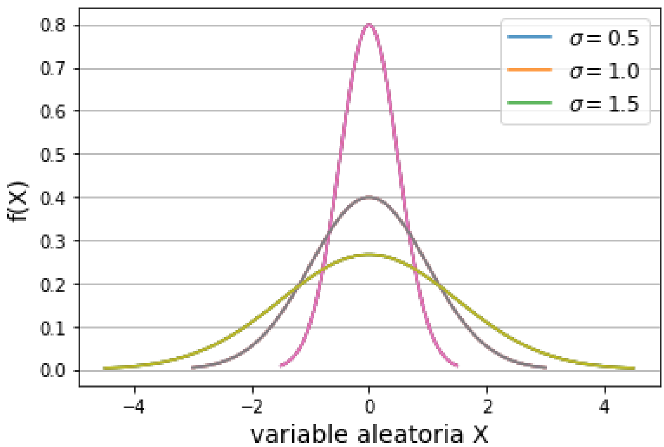


Figura 3.8: Función de densidad de probabilidad con diferentes valores de la desviación estándar.

Si $a = 0$ y $\sigma = 1$ la distribución se llama normal estándar.

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du \tag{3.29}$$

En el caso general, la función de distribución de $F(x; a, \sigma)$ se puede calcular usando la fórmula

$$F(x; a, \sigma) = \Phi(Z)$$

$$Z = \frac{x - a}{\sigma}$$

Se han compilado tablas extensas para la función $\Phi(x)$ y varias de sus derivadas.

Ejemplo 85. El crecimiento de los niños en el grupo de edad de 15 años es una variable aleatoria X normalmente distribuida con parámetros $a = 161\,cm$ y $\sigma = 4\,cm$. ¿Qué proporción de disfraces para niños con una altura de 152 a 158 cm debe proporcionarse en el volumen de producción para este grupo de edad?

Resolución

Encontremos qué proporción de trajes para niños con una altura de 152 a 158 cm debe proporcionarse en el volumen de producción para este grupo de edad. Usamos la fórmula para encontrar la probabilidad de que una variable aleatoria normal caiga en el intervalo:

$$P(152 \leq X \leq 158) = \Phi\left(\frac{158 - 161}{4}\right) - \Phi\left(\frac{152 - 161}{4}\right)$$

$$P\,(152 \leq X \leq 158) = 0.2266 - 0.0122 = 0.2144.$$

Regla de las tres sigmas

Si una variable aleatoria se distribuye normalmente, entonces el valor absoluto de su desviación de la expectativa matemática no excede tres veces la desviación estándar. La probabilidad de que la desviación en valor absoluto sea menor que tres veces la desviación estándar es 0.9973

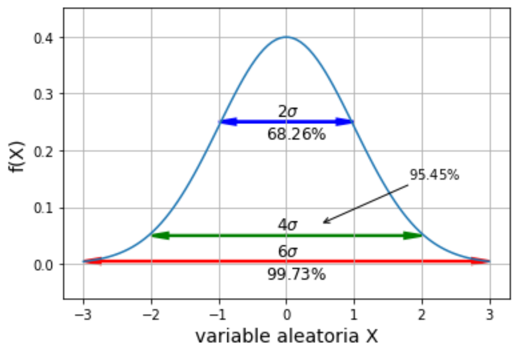


Figura 3.9: Regla de las tres sigmas

En la práctica: si se desconoce la distribución de una variable aleatoria, pero se cumple la condición especificada en esta regla, entonces hay razón para suponer que la variable aleatoria se distribuye normalmente.

Referencias

[1] Aidara, N. (2018). Introduction to probability and statistics.

[2] Debnath, L., & Basu, K. (2015). A short history of probability theory and its applications. International Journal of Mathematical Education in Science and Technology, 46(1), 13-39.

[3] Devore, J. (2007). Probabilidad Y Estadistica.

[4] Downey, A. B. (2011). Think Stats Probability and Statistics for Programmers. Version 1.6. 0.

[5] Duchesnay, E., Lofstedt, T., & Younes, F. (2021). Statistics and Machine Learning in Python.

[6] Forsyth, D. (2018). Probability and statistics for computer science (pp. 36-42). UK: Springer International Publishing.

[7] George, C. C. (1988). Probabilidad y estadística, aplicaciones y métodos. McGrawHill/Interamericana de México, SA. México.

[8] Giri, N. C. (2019). Introduction to probability and statistics. CRC Press.

[9] Gmurman, V. E., & Akoptr, G. (1974). Teoría de las probabilidades y estadística matemática.

[10] Hamming, R. W. (2004). Methods of mathematics applied to calculus, probability, and statistics. Courier Corporation.

[11] Hansen, N. R. (2005). Probability theory and statistics. University of Copenhagen, Denmark.

[12] Johnson, J. L. (2011). Probability and statistics for computer science. John Wiley & Sons.

[13] Montgomery, D. C., & Runger, G. C. (2018). Applied statistics and probability for engineers (p. 710). Hoboken, NJ: Wiley.

[14] O'Leary, M. L. (2015). A first course in mathematical logic and set theory. John Wiley & Sons.

[15] Rao, B. P. (2009). A first course in probability and statistics. World Scientific.

[16] Ross, S. M. (2004). Introduction to probability and statistics for engineers and scientists. Elsevier.

[17] Soong, T. T. (2004). Fundamentals of probability and statistics for engineers. John Wiley & Sons.

[18] Statistics and probability for engineering applications with Microsoft Excel. W. Decoursey Published 2003 Computer Science, Engineering

[19] Unpingco, J. (2016). Python for probability, statistics, and machine learning (Vol. 1). Springer International Publishing.

[20] Walpole, R. E., Myers, R. H., Myers, S. L., & Ye, K. (2007). Probabilidad y estadística para ingeniería y ciencias (No. TA340. P76 2007.). México: Pearson educación.

-Editorial-
CILADI
Centro de Investigación Latinoamericano
para el Desarrollo e Innovación

ISBN: 978-9942-7078-2-6
9 789942 707826